CODEN: JASMAN

ISSN: 0001-4966

The Journal of the Acoustical Society of America

Vol. 110, No. 1 July 2001

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 110, NO. 1, JULY 2001

CONTENTS—*Continued from preceding page*

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 110, NO. 1, JULY 2001

CONTENTS—*Continued from preceding page*

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 110, NO. 1, JULY 2001

CONTENTS—*Continued from preceding page*

ACOUSTICAL NEWS—USA

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

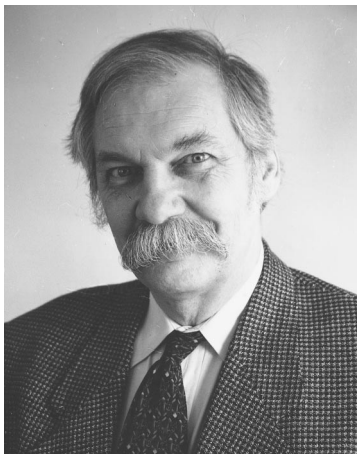
Editor's Note: *Readers of this Journal are encouraged to submit news items on awards, appointments, and other activities about themselves or their colleagues. Deadline dates for news items and notices are 2 months prior to publication.*

New Fellows of the Acoustical Society of America

Rene Causse—For contributions to the understanding of the acoustics of string and wind instruments.

Stan E. Dosso—for contributions to ocean and Arctic acoustic inverse problems.

Jun-ichi Kushibiki—For contributions to ultrasonic metrology and to acoustic microscopy.



Thomas L. Szabo—For contributions to diagnostic ultrasound imaging.

 0001-4966/2001/110(1)/1/5/$18.00

ASA Fellow Elected to the National Academy of Engineering

H. Vincent Poor, Professor of Electrical Engineering at Princeton University, was elected a Fellow of the National Academy of Engineering in 2001 ''for contributions to signal detection and estimation and their applications in digital communications and signal processing.'' He was among the 74 engineers and 8 foreign associates elected this year. Dr. Poor is a Fellow of the Acoustical Society of America.

Leo Beranek receives Honorary AIA Membership

Leo L. Beranek was among ten people to be named honorary members of the American Institute of Architects. Dr. Beranek was acknowledged for his 50 years as an acoustical consultant on projects that ranged from the Tokyo City Opera to the Mann Auditorium in Tel Aviv, as an author and leader in acoustics research, and as a mentor to the next generation of acoustics engineers, notably at Harvard University and the Massachusetts Institute of Technology.

Leo Beranek is an Honorary Fellow of the Acoustical Society of America and has received the ASA's R. Bruce Lindsay Award, the Wallace Clement Sabine Award, and the Gold Medal. He has served ASA in a wide variety of roles including President and Vice President.

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.

2001

15–19 Aug. ClarinetFest 2001, New Orleans, LA [Dr. Keith Koons, ICA Research Presentation Committee Chair, Music Dept., Univ. of Central Florida, P.O. Box 161354, Orlando, FL 32816-1354; Tel.: 407-823-5116; E-mail: kkoons@pegasus.cc.ucf.edu].

19–24 Aug. Asilomar Conference on Implantable Auditory Prostheses, Pacific Grove, CA [Michael Dorman, Dept. of Speech and Hearing Science, Arizona State Univ., Tempe, AZ 85287-0102; Tel.: 480-965-3345; Fax: 480-965-0965; E-mail: mdorman@asu.edu].

4–6 Oct. Ninth Annual Conference on the Management of the Tinnitus Patient, Iowa City, IA [Rich Tyler, Tel.: 319-356-2471; E-mail: rich-tyler@uiowa.edu; WWW: www.medicine.uiowa.edu/otolaryngology/news/news].

7–10 Oct. 2001 IEEE International Ultrasonics Symposium Joint with World Congress on Ultrasonics, Atlanta, GA [W. O'Brien, Electrical and Computer Engineering, Univ. of Illinois, 405 N. Mathews, Urbana, IL 61801; Fax: 217-244-0105; WWW: www.ieee-uffc.org/2001].

29–31 Oct. NOISE-CON 01, The 2001 National Conference and Exposition on Noise Control Engineering, Portland, ME [Institute of Noise Control Engineering, P.O. Box 3206 Arlington Branch, Poughkeepsie, NY 12603; Tel.: 914-462-4006; Fax: 914-462-4006, E-mail: omd@ince.org; WWW: users.aol.com/inceusa/ince.html].

15–18 Nov. American Speech Language Hearing Association Convention, New Orleans, LA [American Speech-Language-Hearing Association, 10801 Rockville Pike, Rockville, MD 20852; Tel.: 888-321-ASHA; E-mail: convention@asha.org; WWW: professional.asha.org/convention/abstracts/welcome.asp].

3–7 Dec. 142nd Meeting of the Acoustical Society of America, Ft. Lauderdale, FL [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]. Deadline for submission of abstracts: 3 August 2001.

2002

21–23 Feb. National Hearing Conservation Association Annual Conference, Dallas, TX [NHCA, 9101 E. Kenyon Ave., Ste. 3000, Denver, CO 80237; Tel.: 303-224-9022; Fax: 303-770-1812; E-mail: nhca@gwami.com; WWW: www.hearingconservation.org/index.html].

10–13 March Annual Meeting of American Institute for Ultrasound in Medicine, Nashville, TN [American Institute of Ultrasound in Medicine, 14750 Sweitzer Lane, Suite 100, Laurel, MD 20707-5906; Tel.: 301-498-4100 or 800-638-5352; Fax: 301-498-4450; E-mail: conv_edu@aium.org; WWW: www.aium.org].

3–7 June 143rd Meeting of the Acoustical Society of America, Pittsburgh, PA [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].

2–6 Dec. Joint Meeting: 144th Meeting of the Acoustical Society of America, 3rd Iberoamerican Congress of Acoustics and 9th Mexican Congress on Acoustics, Cancun, Mexico [Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; WWW: asa.aip.org/cancun.html].

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

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

Some indexes are out of print as noted below.

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

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

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

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

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

Volumes 36–44, 1964–1968: Contemporary Literature. Classified by subject and indexed by author and inventor. 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

Akhtar, Shama Y., ACG, LLC, 10005 Coehorn Ct., Ellicott City, MD 21042
Banks, Michael W., NuTech Engineering Systems, Santo Domingo Office, 10400 NW 33rd St., Ste. 280, Miami, FL 33172
Bemman, Ya-Juan J., Eberspaecher North America, Inc., 33533 W. Twelve Mile Rd., Ste. 160, Farmington Hills, MI 48331
Borchardt, Steven G., Lydall Thermal Acoustical Group, 1241 Buck Shoals Rd., Hamptonville, NC 27020
Boxhorn, George R., Banker, C. I. Wire & Iron Works, Inc., S84 W19120 Enterprise Dr., Muskego, WI 53150
Bugaj, Stephan V., Webmind, Inc., 50 Broadway, Ste. 2100, New York, NY 10004
Cann, Richard G., Grozier Technical Systems, Inc., 157 Salisbury Rd., Brookline, MA 02445-2032
Cervenka, Pierre, Group Acoustique Physique & Ultrasound, Univ. Pierre et Marie Curie, 2 Place de la Pare de Ceinture, Sabt-Cyr-L'Ecole 78210, France
Chen, Yingzhi, Motorola Lab., Motorola China Research Ctr., Room 402, Central Pl., 16 Henan Rd., S., Shanghai 200002, China
Ciccozzi, Marco G., Environment Protection Agency, P.O. Box 2607, Adelaide SA 5001, Australia
Dalrymple, Glenn V., 11305 William Plaza, Omaha, NE 68144-1879
de Lima, Washington J., 7121 Hart Ln., 2089, Austin, TX 78731
Distler, Mark A., ENSR International, 6601 Kirkville Rd., East Syracuse, NY 13057
English, Gary E., 1850 Columbia Pike, #426, Arlington, VA 22204
Francine, Jon K., SRS Technologies, 105 N. ''H'' St., Lompoc, CA 93436
Giner, Jose C., Giner Audio Solutions, Rue Bartira 1430 Ste. 63, Sao Paulo SP 05009-000, Brazil
Grimault, Nicolas, Speech and Sciences, Psychoacoustics Lab., Arizona State Univ., Community Service Ctr., 200 East Curry Rd., Tempe, AZ 85287-1908
Hartynenko-Kresse, Olga M., 8836 York Rd., S., Bloomington, MN 55431
Hiramatsu, Kozo, Dept. of Human Environment Studies, Mukogawa Womens Univ., Ikebirakicho 6-46, Nishinomiya, Hyogo 663-8558, Japan
Ho, Wai-Lun, Arup Acoustics, Ove Arup & Partners, Level 5, Festival Walk, 80 Tat Chee Ave., Kowloon Tonh, Kowloon, Hong Kong
Hodges, Richard P., 20 Butlertown Rd., Waterford, CT 06385
Ives, Terri E., 4930 National Ave., #16, San Jose, CA 95124
Jayachandran, Vijay, United Technologies Research Ctr., 411 Silver Ln., MS 129-17, East Hartford, CT 06108
Jeong, Dae-Up, Dept. of Architectural Eng., Faculty of Architecture, Chonbuk National Univ., Jeonju, Cholabuk-do, 561–756, South Korea
Kim, Mi-Ran, 1180 Chaddwyck Dr., Athens, GA 30606
Kirk, Des L., Dept. of Physiology, Univ. of Western Australia, 35 Stirling Hwy., Crawley WA 6009, Australia
Laverty, Richard R., 4835 W. Braddock Rd., #102, Alexandria, VA 22311
Lazzari, Ricardo O., Pioneer Speakers, Inc., 132 Mile of Cars Way, National City, CA 91950
Mantakas, Marios, Galimis 6, Prokonisos, Stamata Attikis 14575, Greece
Masahiko, Komatsu, 4-4-17-108 Nishiaoki, Kawaguchi, Saitama 332-0035, Japan
Mays, Brian T., 420 Montemar Ave., Baltimore, MD 21228
Meyer, Alan W., 665 Dunhill Dr., Danville, CA 94506
Middel, Trevor A., Ontario Ministry of Natural Resources, Aquatic Ecosystem Science, 300 Water St., Peterborough, ON K9J 8M5, Canada
Mimnagh, Michael L., Naval Surface Warfare Ctr., Dept. 841, 1000 Kitty Hawk Ave., Philadelphia, PA 19112
Morrone, Ivano, Via F. Simonetta 1A, Cosenza CS I-87100, Italy
Moyer, David L., Riverbank Acoustical Laboratories, IITRI, 1512 S. Batavia Ave., Geneva, IL 60134
Nucci, Adam, J., 7 Knight St., Narragansett, RI 02882
Parente, Carlos E., COPPE/UFRJ, Ocean Engineering, Ilha do Fundao, C.T. P.O. Box 68508, Rio de Janeiro 21945-970, Brazil
Parthasarathy, Teralandur K., 344 Westminster Dr., Glen Carbon, IL 62034
Paulini, Heliton L., Syncrotape, rua Paes de Araujo 29-CJ 11, Sao Paulo 04531-090, Brazil
Pulli, Jay J., BBN Technologies, Sensor Systems and Technologies, 1300 N. 17th St., Arlington, VA 22209
Sabransky, Philip, 838 Park Ave., Wilmette, IL 60091
Schwartz, Gregory L., Schulmerich Bells, Engineering, 1 Carillon Hill, Sellersville, PA 18960
Shimizu, Yasushi, 64 Mansion Boulevard, Apt. J, Delmar, NY 12054
Smith, Steven J., Exponent Failure Analysis Associates, 2 N. Riverside Plaza, Ste. 1400, Chicago, IL 60606
Stevens, John W., Kelman Seismic Processing, 16000 Barker's Point Ln., Ste. 100, Houston, TX 77079
Terra, Maria E., Duke Engineering and Services, 1111 N. Forest St., Belingham, WA 98225
Torres, Juan A., Vocal Technologies, Ltd., 200 John James Audubon Pkwy., Buffalo, NY 14228
Uhlman, James S., 99 John St., Newport, RI 02840
Underwood, Ronald L., IBM, 2K6A, 060/T116, 3039 Cornwallis Rd., Research Triangle Park, NC 27709
Upton, Zachary, 12821 Bushey Dr., Silver Spring, MD 20906
Van Dyke, Paul R., 6754 Bellview Pines Rd., Pensacola, FL 32526
Van Wijngarden, Sander J., TNO Human Factors, Kampweg 5, Soesterberg 3769 DE, The Netherlands
von Muggenthaler, Elizabeth K., P.O. Box 1126, Hillsborough, NC 27278
Wisdom, Sheyna S., Environmental Branch, URS Corporation, 1615 Murray Canyon Rd., Ste. 1000, San Diego, CA 92108
Zhao, Xiang, Mechanical Engineering, Wayne State Univ., 5050 Anthony Wayne Dr., Room 2124, Detroit, MI 48202
Zu, Yiqing, Motorola Lab., Motorola China Research Ctr., Room 402, Central Place, 16 Henan Rd., S., Shanghai 200002, China

New Students

Ahkuputra, Visarut, Digital Signal Processing Research Lab., Dept. of Electrical Engineering, Faculty of Engineering, Chulalongkorn Univ., Patumwan, Bangkok 10330, Thailand
Ahmad, Sabri, National Univ. of Malaysia, Mechanical and Materials Eng., Bangi Selangor D.E., Selangor 43600, Malaysia
Al-Bataineh, Osama M., 15C Graduate Cir., State College, PA 16801
Alabi, Halimat I., 218-05 Nimitz Dr., W. Lafayette, IN 47906
Albanese, Anne M., 350159 Georgia Tech Station, Atlanta, GA 30332
Alberts II, William C. K., NCPA, Univ. of Mississippi, 1 Coliseum Dr., University, MS 38677
Ament, Jacob E., 3105 Woodland St., Ames, IA 50014
Andersen, Troy M., P.O. Box 724, Gloucester Point, VA 23062
Anderson, Michael D., 119 W. Fowler Ave., Apt. 2A, W. Lafayette, IN 47906
Auriemmo, Jane, 10 Waterside Plaza, 9C, New York, NY 10010
Beniwal, Ravinder S., 5200 Anthony Wayne Dr., Apt. 805, Detroit, MI 48202
Bent, Tessa C., Linguistics Dept., Northwestern Univ., 2016 Sheridan Rd., Evanston, IL 60208
Bharadwaj, Sneha V., 4401 Creekstone Dr., Plano, TX 75093
Billon, Alexis J., LEPTAB, Pole Sciences et Technologies, La Rochelle 17042, France
Bonde, Thomas L., 680 Hathaway St., Apt. 1, Platteville, WI 53818
Brown, Daniel, 112 Skyline Dr., Apt. A, Oxford, MS 38655
Browne, Shannon T., 108 W. 39th St., Apt. 32, Baltimore, MD 21210
Buedenbender, Niklas, Hovestadtstrasse 20, Hagen NRW 58093, Germany
Chaparro, Gabriel E., 2767 Grove Way, Castro Valley, CA 94546
Clopper, Cynthia G., 800 N. Smith Rd., 1C, Bloomington, IN 47408
Cone, Michael T., 313 N. High, Kirksville, MO 63501-3034
Crystal, Amy, 800 Ave. ''H,'' Apt. 2H, Brooklyn, NY 11230
Cummins, Andrea E., 934 Montecito Dr., Los Angeles, CA 90031
Deecke, Volker B., School of Biology, Inst. of Environmental and Evolutionary Biology, D31 Bute Bldg., Univ. of St. Andrews, St. Andrews KY16 9TS, United Kingdom
Della Chiesa, Andrea, via Walter Tobagi 8, Gavirate, Varesa 21026, Italy
Delory, Eric, Acoustic Research Lab., TMSI, Natl. Univ. of Singapore, Singapore 119260, Singapore
Dittberner, Andrew B., 1805 5th St., #1, Coralville, IA 52241
Douville, Hugo, 883 Larocque, Sherbrook, PQ J1H 4R2, Canada
Evans, Michele E., School of Psychology, Univ. of Birmingham, Edgbaston, Birmingham, W. Midlands B15 2TT, Great Britain
Feist, Jeff P., 1222 Burberry Dr., East, Apt. 185, Lafayette, IN 47905
Freeburg, Michael J., Sound, Acoustical Design & Sound Contracting, Colubia College Chicago, 676 N. LaSalle St., Chicago, IL 60610

Hansen, Scott, Engineered Aeroacoustics, Inc., 5207 Douglas Dr., N., Minneapolis, MN 55429
Harte, James M., Signal Processing and Control Group, Inst. of Sound and Vibration Research, Highfield, Southampton, Hampshire SO17 1BJ, United Kingdom
Harwell, Ross M., 321 Aberdeen Ave., Apt. 2, Hamilton, ON L8P 2R6, Canada
Hayiou Thomas, Marianna, Dept. of Experimental Psychology, Oxford Univ., South Parks Rd., Oxford OX1 3UD, United Kingdom
Helms, Zacharay M., P.O. Box 16299, Raleigh, NC 27607
Horn, Jennifer H., 401 Amberson Ave., #143, Pittsburgh, PA 15232
House, Sara M., Geological Sciences, Southern Methodist Univ., 3225 Daniel Ave., Dallas, TX 75275
Huang, Chen-Fen, Marine Physical Lab., Scripps Inst. of Oceanography, 9500 Gilman Dr., La Jolla, CA 92093-0238
Huss, Martina M. E. F., Dept. of Experimental Psychology, Univ. of Cambridge, Downing St., Cambridge CB2 3EB, United Kingdom
Judge, James D., Dept. of Life Sciences, Nottingham Univ., Nottingham NG7 2RD, United Kingdom
Kaouri, Katerina, Oxford Univ., Oxford Ctr. for Industrial and Applied Mathematics, 24-29 St. Giles', Oxford OX2 6HD, England
Kawai, Norimune, 4570 Comanche Dr., Boulder, CO 80303
Kim, Jong-Hwa, Gruberzeile 73, Berlin D-13593, Germany
Knight, Derrick P., 25 Lake Shore Dr., Apt. 2B, Watervilet, NY 12189
Kock, Johan F. W., R & D Audiology, JCH Developments, 562 Jaqueline Dr., Garsfontien, Pretoria, Gauteng 0042, South Africa
Kwon, Youngmin, 400 McChesney Ave. Extension, Troy, NY 12180
Lauer, Amanda M., Psychology Dept., Univ. of Maryland, College Park, MD 20742
Lee, Byung Kwon, SeLim Apt. 8-1102, MaJang-Dong, SungDong-Gu, Seoul 133-752, Korea
Lee, Kelvin, Blk. 850 Woodlands St. 82, #08-225, 730850, Singapore
Lee, Sunghoon, 601 N. Eutaw St., #212, Baltimore, MD 21201
Liu, Chang, Dept. of Speech and Hearing Sci., Indiana Univ., 200 S. Jordan Ave., Bloomington, IN 47405
Liu, Li, Aerospace and Mechanical Eng., Boston Univ., 110 Cummington St., Boston, MA 02215
Lutz, Steven D., Penn State/ARL, Graduate Program in Acoustics, P.O. Box 30, N. Atherton St., State College, PA 16804-0030
McLaughlin, Daniel A., 8126 Middlesex Rd., Mentor, OH 44060
Menezes, Caroline M., Speech and Hearing Sciences, The Ohio State Univ., 1070 Carmack Rd., Columbus, OH 43210
Milanesi, Veronica, Teodoro Vilardebo 3053, Buenos Aires 1417, Argentina
Monroy, Bonnie, 815 Park Ln., Hoffman Estates, IL 60195
Nakajima, Yasuo, Linguistics Dept., Univ. of Colorado at Boulder, Campus Box 295, Boulder, CO 80309
Nock, Allen B., NCPA, Univ. of Mississippi, 1 Coliseum Dr., University, MS 38677
Nuttakorn, Thubthong, 39/166 M.1 Rama 2 Rd., Bangkhunthein, Bangkok 10150, Thailand
Ordonez, Carlos, 9320 Fontainebleau Boulevard, B-410, Miami, FL 33172
Padilla, Monica, 12038 207th St., Lakewood, CA 90715
Paredes, Tomas S., Austria 2332, Valdivia, Chile
Parn, Paavo, Staadioni 14-20, Tartu Maakond 51008, Estonia
Pelletier, Martin E., 5018 Gabriel, Rock Forest, PQ J1N 2H2, Canada
Pilon, Dominic, 4657 Nuckle, Pierrefonds, PQ H9J 3P8, Canada
Renaud, Stephanie C., 3830 McDowell Ln., Baltimore, MD 21227
Rickter, Elizabeth A., P.O. Box 802, CM 1073-Drew, Madison, NJ 07940-0802
Rinehart, Hank S., 200 Hartness St., Apt. B-3, Starkville, MS 39759
Saifuddin, Kazi, 902, Kinoshita, 2-10-15 Kano-cho, Chuo-Ku, Kobe 650-0001, Japan
Saleh, Khaldon, 7-C Graduate Cir., State College, PA 16801
Seney, Kali A., 1223 N. Wolcott, #2, Chicago, IL 60622
Serafin, Stefania, Music Dept., CCRMA, Stanford Univ., Stanford, CA 94305-8180
Shimizu, Kazushi, Motonakayama 3-8-16, Funabashi-City, Chiba 273-0035, Japan
Smith, Alan K., 2420 #16-D Marchbanks Ave., Anderson, SC 29621
Solis, Andres, 90 rue Saint Louis en L'ile, Paris 75004, France
Tebbens, Christopher J., OCMR 1469, Oberlin, OH 44074
Tong, Angus C., Claycroft 1, Room 4, Flat 16, Univ. of Warwick, Coventry CV4 7AL, United Kingdom
Tsai, Chen-Gia, Martinstr. 3, Berlin 12167, Germany
Vliegen, Joyce, Sonseweg 7, Eindhoven, Noord Brabant 5632 NA, The Netherlands
Warren, Willis J., 1900 Burton Dr., #121, Austin, TX 78741
Wheeler, Howard B., Physics Dept., Univ. of Mississippi, 1 Coliseum Dr., University, MS 38677
Whitten, Julie L., 3459 N. Kolmar Ave., Chicago, IL 60641
Wong, Hon Yiu, Room 12, 17/F, Hiu Kwong Ct., Hiu Kwong St., Kowloon, Hong Kong
Yerikalapudi, Aparna V., HIG-44 M.V.P. Colony Sector-1 (Venkojipalem), Vsakhapatnam AP 530017, India
Zhang, Xinya, 1206 University Village, Apt. B, East Lansing, MI 48823
Zvonik, Elena, Computer Science, Univ. College Dublin, Belfield, Dublin 4, Ireland

Members Elected Fellows

P. J. Abbas, S. L. Broschat, A. J. Campanella, R. E. Causse, M. F. Dorman, S. E. Dosso, E. C. Everbach, M. P. Gorga, J. H. Grose, T. Hidaka, P. C. Hines, C. K. Holland, J.-I. Kushibiki, N. C. Martin, C. F. Moss, V. M. Richards, M. A. Ruggero, C. Soize, T. L. Szabo, E. A. Tobey, W. A. Watkins, B. R. Tittmann

Associates Elected Members

J. B. Au-Yeung, J. W. Benson, C. S. Campbell, D. D. B. Chan, F. Chong, U. Eysholdt, N. A. Gumerov, M. J. Hammer, F. H. Kerr, O. N. Larsen, T. W. Leishman, Y. T. Y. T. Leung, E. A. Lopez-Poveda, S. N. Makarov, S. A. Rizzi, K.-I. Sakakibara, M. A. Sumbatyan, S. Suresh, N. Tanaka, J. D. Teel, R. H. Telling, S. Vera, C. Wang, J. Zeegers, P. J. Zoccola, L. M. Zurk

Students to Associates

J. S. Bevan, G. J. Cannon, P. A. Clements, S. C. Conlon, A. M. Crawford, P. M. Daly, J. A. Daniels, V. F. Dias, H. E. Farris, C. H. Frazier, N. M. Gage, B. T. Hefner, M. G. Heinz, G. A. Herrera, A. W. Howitt, P. J. B. Jackson, J. A. Johnson, G. Kapodistrias, S. R. Khosla, J. C. Krause, E. H. Langendijk, C. J. Long, W.-L. D. Ma, J. A. Marler, M. Mellody, J. M. Meyer, S. Muhdramli, A. E. Moreton, H. Muesch, T. B. Neilsen, D. A. Najolia, M. L. Oelze, T. Painter, P. M. Pavlov, N. Prodi, N. E. Rayess, D. T. Ries, M. A. Roch, P. Rutherford, T. O. Samuels, S. J. Sharp, J. L. Slifka, A. M. Thode, D. A. Vickers, K. A. Weatherly, S. H. Yonak

Associates to Students

F. J. Gallun, D. O. Ludwigsen, P. Torre

Reinstated

G. J. Quentin, M. B. Sachs—*Fellows*
R. H. Gilkey, Q. Lin—*Members*
H. C. Robinson—*Associate*

Resigned

A. Freedman, S. M. Khanna, R. C. Potter, B. J. Uscinski—*Fellows*

A. Edwards, A. B. Kolpakov, M. W. McNeary, R. V. Murray, R. L. O'Toole, J. Penrose, J. G. Pierce, M. Watari, J. C. Weibler, J. Yan—*Members*

K. A. Barrett, K. N. Bates, R. Bayly, J. F. Borst, D. G. Druker, R. G. Frodl, C. Galmarini, E. Jones, N. Polish, A. Rochet, F. B. Seeley, A. Seither-Preisler, R. E. Stone, Jr., H. Strik, P. H. H. M. van Lieshout, T. M. Young, S. E. Zeller—*Associates*

C. DeSellier, M. R. Eldredge, L. Hinkelman, O. S. Jahromi, J. Kocbach, G. T. Stapleton, L. Von-Paumgartten, P. W.—*Students*

Deceased

Paul Rosenberg, L. K. Stein, B. Sturtevant—*Members*
B. Case, R. Stark, T. L. Walker, Jr.—*Associates*

Dropped

J. N. Decarpigny, F. R. DiNapoli, V. A. Krasilnikov, D. Lee, R. L. McKinley, C. W. Nixon, S. Tjotta—*Fellows*

L. Amundsen, S. Asvadurov, J. R. Bain, J. B. Bairstow, B. Bajic, F. M. Becker, K. J. Berhardt, R. C. Berry, P. M. Bradshaw, J. R. Britt, B. F. Brown, M. S. Brussieux, E. L. Butler, L. Chang, D. S. Clark, N. D. Cogger, M. G. Cote, T. J. Degroot, V. Dutt, D. Eden, T. H. Ensign, J. P. Ertel, C. Y. Espy-Wilson, D. G. Fagen, Z. Feng, M. Fink, R. P. Flanagan, F. K. Forster, A. T. Fry, W. T. Frysinger, G. P. Gibbs, G. Gimenez, A. K. Goble, M. D. Hauser, R. S. Heffner, E. G. Henneke, J. L. Hieronymus, D. Houben, N. I. Ivanov, D. A. James, P. H. S. Jen, H. E. Jones, J. J. Kassal, D. E. Kidd, J. S. Kim, J. T. Kim, A. N. Kings, A.-R. M. Koliani, S. F. Kooney, A. Lakomyj, H. W. Laney, P. C. Laux, L. M. Long, I.-T. Lu, I. Magrin-Chagnolleau, D. L. Mcpherson, P. F. Millot, C. A. Moore, P. D. Mourad, H. Ogawa, R. J. Orlowski, D. L. Paine, H. V. L. Patrick, D. M. Photiadis, T. W. Picton, Y. Qi, N. N. Reddy, J. S. Robertson, W. I. Roderick, R. J. Roeser, C. Schweitzer, S. M. Sessler, C. Smith, W. Soedel, S. A. Steele, F. J. M. Sullivan, G. S. Sylvester, Y. M. Szymko, X. Tang, H. M. Tavossi, R. T. Trochta, P. K. Weber, K. Weissman, H. Y. Woo, H. Zhang, L. J. Ziomek—*Members*

V. P. Adamson, A. M. Ambrose, B. T. Arnason, K. Avval, T. G. Barker, S. Barrera-Figueroa, B. P. Basnett, N. H. Bay, E. Benarrous, V. A. Benavides, K. M. Berg, M. Bernard, L. Biru, D. G. Blasko, J. C. Bogus, G. L. Bradshaw, V. Cai, A. M. Calabrese, M. Cappelli, N. M. Carbone, Q. Chen, B. Cherek, K. O. Chun, J. S. Chung, L. T. Cohen, T. S. Collins, M. P. Cooke, N. T. Cooley, B. R. Crabtree, C. H. Craig, K. N. Crayme, G. B. Cunningham, M. E. De Billy, D. L. Debruin, D. D. Deliyski, P. L. De Leon, L. D. De Mendonca, E. Douaze, R. K. Durabb, Jr., R. M. Ead, C. H. Echols, A. A. Elshintinawy, B. A. Engel, R. J. Epstein, W. E. Evans, R. Eyraud, J.-F. Feiteira, J. Flores, J. R. Ford, S. Gaona, M. D. Garrett, J. R. Gavin, G. K. Goacher, D. L. Gochenour, S. Goguet, L. R. Goldberg, A. J. Goodall, R. A. Goodlin, F. J. Goodman, K. C. Goyal, J. S. Gravel, C. H. Greene, M. Grenie, J. Guasca, S. K. Gupta, R. Gunda, H. S. Haggerty, D. A. Hamilton, W. Han, R. K. Hansen, D. R. Hansmann, X. Hao, T. Haresign, S. L. Hargus, J. A. Hartikainen, C. Herman, C. J. Hickey, R. B Hill, J. W. Hillenbrand, D. A. Hock, L. D. Holden-Pitt, W. Hong, B. V. Honein, W. Huang, R. Hukin, K. Inagaki, J. W. Irza, B. H. Juang, A. J. Juengst, G. Kaduchak, H. Kaneko, R. Kanta, K. D. Karlson, M. Kato, G. J. Kenehan, H.-W. Kim, J. H. Kim, Y. H. Kim, Y.-N. Kim, M. E. King, G. Klimanis, E. I. Knudsen, J. J. Kowalewski, R. L. Kranz, S. Krishnamurti, D. H. Kruger, D. M. Labrozzi, M. P. Lamoureux, C. F. H. Lawrence, F. Le Her, A. R. Leyman, D. Li, Jay A. Lightfoot, Y. Liu, C. C. Locqueteau, R. M. Logan, J. Lovetri, D. Margoliash, B. E. Marlin, B. J. Martin, M. A. Martin, S. A. Martin, D. C. Massie, M. Mastoianni, T. C. Mathews, C. M. Mayhugh, S. Mazzola, R. M. McCafferty, B. R. McLennan, S. Mehedincu, D. R. Melton, J. M. Mendoza, D. Menemenlis, R. K. Menoche, T. A. Meyer, F. J. Miranda, J. G. Mitchell, C. W. Morton, M. Naguib, S. Narayan, C. L. Nicholson, P. Nihalani, B. A. Oakley, R. M. Oba, M. T. Ochs, H. Oguz, A. Oliver, S. Omata, E. J. O'Neill, J. L. Orr, J. W. Osborne, G. Papanikolaou, L. R. Parma Salazar, T. N. Patro, R. A. Pawlowicz, C. Peil, N. R. Pennington, D. E. Perkins, J. N. Phillips III, L. L. Pierson, W. W. Piroth, P. Polson, J. E. Preminger, D. Pressnitzer, E. Rajcan, D. T. Raphael, M. Riera-Gaxiola, L. N. Robinette, I. E. Rodas, L. F. Rogers, P. Rossi, S. D. Roth, L. S. Ryden, T. Sandralodabai, A. Sakuta, J. S. Salmetno, D. P. Saomon, A. M. Sampath, C. M. Sapienza, H. Schloegel, R. M. Schmitt, S. K. Scott, J.-P. Sessarego, L. S. Sheiba, D. L. Siurna, P. D. Slack, M. J. Smith, B. Somek, E. Sonnenschein, G. A. J. Soulodre, W. G. Spears, Jr., L. S. Straus, W. Sun, M. D. Szymanski, C. G. Thurman, P. Tierce, A. Tsuchida, S. M. Unger, N. Uosaki, R. N. Urella, R. A. Wallander, S. R. Wallin, S. D. Wampler, W. Wang, J. C. Wendoloski, M. A. Whitley III, T. M. Whitney, M. L. Wilson, D. P. Wissing, M. A. Wolfson, J. P. Woolley, L. Wyse, W. Yang, G. Zhang—*Associates*

M. E. M. Aanerud, R. J. Abrams, T. B. Alder, E. A. Andalaft, D. V. Anderson, R. B. Arslan, J. D. Atkinson, M. J. Baldridge, C. W. Bartlett, I. M. Berk, R. J. Berrington III, M. A. Bertolasi, R. Billerey, S. D. Billingsley, L. E. Bischel, G. I. Block, N. D. Bohne, C. P. Brown, J. W. Bruce III, A. D. Buckthought, J. C. Carter, D. Chelidze, S. M. Christian, C. J. Christianson, B. R. Clement, M. A. Cordel, J. C. Couche, F. P. Curra, A. G. Deloach, J. M. Denton, U. Di Mondo, A. M. Diaz, S. R. Diehl, J. L. Dimaggio, R. K. Duncan, E. H. El Moutia, A. A. Ergin, P. J. Evers, C. F. Famoso, J. A. Felix, L. E. N. Fonseca, M. Fortin, J. B. Franklin, J. J. Freeman, C. G. Fuller, J. D. Gaffney, J. S. Gerber, A. Grant, Q. Guo, H. Habibzadeh-Vaneghi, T. R. Hahn, J. M. Hajda, M. Hampson, F. Han, J. E. Hardiman, C. R. Herr, A. Holobinko, J. K. Hopkins, A. Hou, D. M. Houston, X. Huang, F. V. Hutcheson, J.-I. Im, J. M. Isorna, S. Jannedy, M. A. Karahalios, D. M. Kaufman, M. C. Kelly, C. D. Kestel, K. T. Khilnani, N. Kibre, Y. K. Kim, W. M. King, D. W. Klem, A. Kimar, J. A. Lackpour, I. Ladabaum, R. G. Lanyon, Z. I. Lazar, S.-H. Lee, D. A. Lemieux, D. W. Lemonds, D. Lin, C. J. Link, J. F. Lopez, C. S. E. Malliopoulos, I. N. Mann, G. Manthei, V. C. Marino, M. J. Marr-Lyon, K. D. Martin, K. S. Martzall, K. I. Matveev, K. M. McArthur, S. B. Mehta, J. A. Miele, M. L. Miller, E. Mitchell, M. A. S. Mohammad, A. M. Monsalve, M. L. Morrison, P. E. Murphy, S. O. Murray, S. E. Murray, H. G. Myung, C. E. O'Connell, E. C. Odgaard, N. D. Okonak, A. A. Olchowski, J. G. Olgin, R. A. Osorio, V. Panaso, K. Parham, R. Patel, J. K. Poulsen, D. S. Prasad, E. A. Pyper, A. Ratle, M. A. Redford, R. D. Reich, R. A. Riebau, L. J. Romano, K. M. Rosen, L. G. Rozelle, A. J. Sach, D. J. Schram, P. N. Schwartz, M. I. Scott, L. Shaffer, T. R. Shaffer, J. E. Shaklee, D. D. W. Sihn, T. M. Skar, A. J. Sottile III, C. W. Southworth, T. C. Steele, M. Stepura, A. Taff, A. Taglia, C.-T. Tan, T. S. Tierney, P. A. Torres, O. X. Uribe, D. Urueta-Mazzilli, I. Vaynshteyn, J. J. Venditti, J. A. Viator, R. G. Villacorte, M. Wakabayashi, E. Wang, L. Wang, M. K. West, J. K. Wickiser, L. Wu, S. C. Wyatt, Q. Xie, S. R. Yoder, I. Yun, M. Zanartu, B. A. Zawaydeh—*Students*

Fellows	982
Members	2692
Associates	2420
Students	782
	6876

BOOK REVIEWS

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

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

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

Science of Percussion Instruments

Thomas D. Rossing

World Scientific, New Jersey, 2000.
224 pp. hardcover (160 pp. paperback) Price: $75.00 hardcover, $17.00 paperback ISBN: 9810241585 hardcover (9810241593 paperback).

This book is concerned with percussion instruments. It can be viewed as a substantial extension of the chapters devoted to percussion instruments in the book *The Physics of Musical Instruments* by N. H. Fletcher and the author. It also contains new material: acoustics of steelpans, bells, and glass instruments, in particular. Readers with long-term experience in the musical acoustics literature will recognize that some chapters are extended versions of previous material published in 1976 and 1977 in *The Physics Teacher* by the author. As mentioned in the Preface, this book is written primarily for musicians. It is largely based on the research performed by the author himself on percussion instruments, together with students and colleagues, for more than 25 years.

Chapter 1 briefly presents the percussion family. The author insists, with reason, on the difficulty of classifying percussion instruments. This classification mainly depends on whether the purpose is to emphasize physical or perceptual differences or to highlight the use of a given instrument in an orchestra.

Chapter 2 is devoted to ''Drums with definite pitch.'' It starts with two short paragraphs on the vibrations of strings and membranes in order to present the concept of normal mode, which will be used extensively throughout the book. The author then presents in greater detail the acoustics of timpani where the air–membrane coupling substantially modifies the vibration of the membrane, compared to the *in vacuo* case. In my opinion, two papers on timpani, which were recently published in JASA: Rhaouti *et al.* (J. Acoust. Soc. Am. **105**, 3545–3562, 1999) and Sullivan (J. Acoust. Soc. Am. **101**, 530–538, 1997), and which usefully complement this presentation, could have been summarized in the text and inserted in the references. The chapter ends with an interesting presentation of Indian drums.

One innovation of the present book is the introduction of three chapters called ''Interludes,'' the aim of which is to help the reader become more familiar with the basic concepts of sound perception and vibrations of structures. Chapter 3 is devoted to ''sound and hearing.'' It starts with a brief presentation of sound waves, followed by the definition of the decibel scale and of the concept of loudness. This short interlude ends with considerations on sound power level, on masking sounds, and on the dependence of loudness on duration. This selection is coherent with percussion instruments since, as written by the author, a number of percussion instruments ''are able to mask any instrument in the orchestra.'' In addition, some percussion instruments, such as the cymbal or the xylophone, are characterized by a very loud impulsive starting transient which can cause severe damage to the ear. Other short ''interludes'' appear as particular sections in other chapters of the book. Interludes on pitch and sound radiation, for example, can be found in Chapter 2, where the presentation of these concepts is directly connected to the acoustics and psychoacoustics of drums.

In contrast with Chapter 2, Chapter 4 is devoted to the acoustics of drums with indefinite pitch. This chapter also contains two interludes dealing with the problem of pitch glides in membranes, and with the presentation of modal coupling in a system with two degrees of freedom (a ''two-mass vibrator''). The former phenomenon is musically important. It is the consequence of a nonlinear effect due to the fact that the tension of the membrane increases with the magnitude of the transverse displacement. The rest of the chapter is almost entirely concerned with the presentation of a large variety of drums from all over the world: tom-toms, snare drums, bass drums, congas, bongos, rototoms, African and Japanese drums, and so on. In each case, the description and use of the instrument under study is followed by a discussion of its vibrational and/or acoustical behavior. This discussion is essentially based on experimental results. A substantial part of these results was obtained by the author himself, together with colleagues and students. The reader will certainly be impressed by the large amount of experimental data, including particularly beautiful holographic interferograms of mode shapes.

Chapter 5 is an interlude in which fundamental results on the vibrations of bars and air columns are briefly presented. This chapter serves as an introduction for Chapter 6, where xylophones and marimbas are presented in detail. Chapter 7 describes metallophones. These two families of instruments correspond to the general class of mallet percussion instruments where rigid elastic bodies are set into vibration by the impact of a mallet. The ''physical'' content of Chapters 6 and 7 is similar to Chapter 19 of *The Physics of Musical Instruments*. The primary originality of the present book is the description of lesser known instruments, such as tubaphones, gamelan chimes, and African lamellaphones. In addition, a number of interesting musical references are given at the end of Chapter 7.

Following the same scheme, Chapter 8 is an interlude which describes the vibrations of shells and plates in order to help the reader to understand the remaining seven chapters of the book. Chapter 9 is concerned with cymbals, gongs, and plates. As in the previous chapters, the author describes each instrument and its musical use in detail and makes an effort to highlight the connections between the physical observations and their musical implications. The collection of electronic TV holograms of cymbals and Karen bronze drums shown in this chapter is particularly impressive.

Chapter 10, devoted to steelpans, contains 21 pages, which is many more compared with the previous book: *Physics of Musical Instruments*. This illustrates the current deep interest of the author in this family of instruments. As stated in the introduction, ''the Carribean steelpan is probably the most important new acoustical instrument to develop in the 20th century.'' Acousticians who are not familiar with optical measurements of vibrations will certainly appreciate the very clearly written interlude on holographic interferometry. The presentation of the instruments is certainly well adapted to musicians. People with more developed scientific and mathematical backgrounds will benefit from reading the various papers written by A. Achong, whose references are given at the end of the paper. The Carribean steelpan is not only musically fascinating, but is also a nonlinear mechanical system of considerable difficulty whose behavior is far from being completely understood at this time.

The three following chapters (Chapters 11–13) describe three families of bells. These chapters are entitled: ''Church bells and carillons,'' ''Handbells, choirchimes, crotals and cowbells,'' and ''Eastern bells.'' This extensive part of the book clearly reflects a strong interest of the author and his long-time experience with these instruments. From a musical point of view, one should appreciate the number and variety of illustrations, as well as the sheer quantity of measurements.

Glass musical instruments are described in Chapter 14. These instruments produce a very charming sound and their acoustic behavior is very similar to that of bells. To a first approximation, models of thin shells can be applied to the vibrations of glasses. As in the previous chapters, the author provides interesting information relevant to music and perception, in connection with the physics. The numerous illustrations and descriptions of various instruments contribute to the pleasantness of the book for both musicians and acousticians. The book ends with Chapter 15, with the presentation of some new or exotic percussion instruments such as the anklung, the Deagan organ chimes, the ''cristal Baschet,'' or the ceramic instruments of Ward Hartenstein. This trip through the world of percussion instruments seems to be unfinished since, as mentioned by the author, ''there are literally thousands of other percussion instruments;'' a nice perspective for researchers who would like to enter this fascinating world.

The book is very clear and easy to read. The author demonstrates particular skill and enthusiasm for presenting rigorously complex phenomena using simple words and short sentences. Although the number of pages is rather small, this book contains a high density of information. It should certainly become a reference book for percussionists, music teachers, and physicists who are interested in this family of instruments. Professional acousticians will probably like to complete their knowledge with the book *The Physics of Musical Instruments* and by reading musical acoustics papers in journals.

ANTOINE CHAIGNE
Ecole Nationale Supérieure des Techniques Avancées (ENSTA)
Unité d'Enseignement et de Recherche en Mécanique (UME)
Chemin de la Hunière
91761 Palaiseau cedex
France

 0001-4966/2001/110(1)/15/1/$18.00

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:

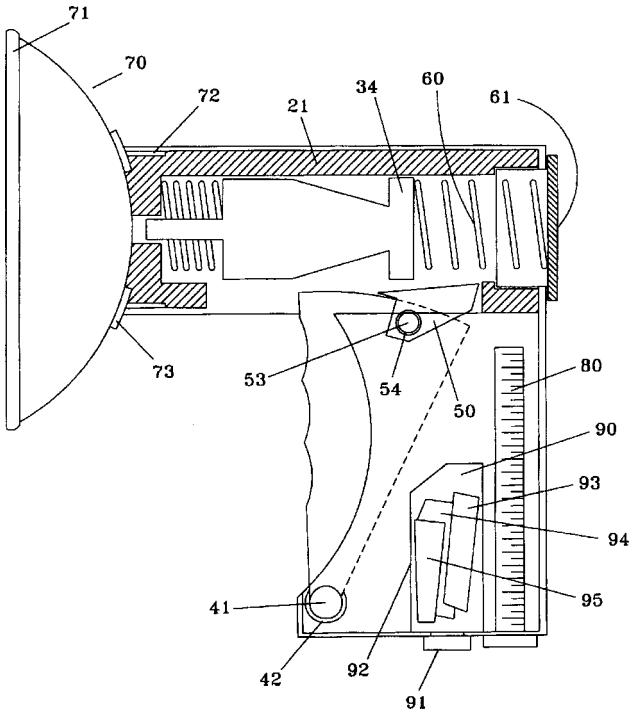
GEORGE L. AUGSPURGER, *Perception, Incorporated, Box 39536, Los Angeles, California 90039*
IBRAHIM M. HALLAJ, *Wolf Greenfield & Sacks P.C., 600 Atlantic Avenue, Boston, Massachusetts 02210*
DAVID PREVES, *Songbird Hearing, Inc., 5 Cedar Brook Drive, Cranbury, New Jersey 08512*
WILLIAM THOMPSON, JR., *601 Glenn Road, State College, Pennsylvania 16803*

6,160,760

43.30.Jx UNDERWATER SIGNALING DEVICE

Mark Rayner, Calgary, Canada
12 December 2000 (Class 367/141); filed 11 June 1999

A hand-held underwater signaling device, configured like a hand gun, comprises a pivoted trigger mechanism (not numbered in this figure) which causes a pivoting spur **50** that bears against a hammer (item **34** is the rearmost portion of this hammer) to force the hammer backwards, thereby compressing the primary spring **60**. When the spur **50** clears the flange **34** of the hammer, the hammer is thrust forward by the spring whereupon its tip



strikes a bell **70** which radiates into the surrounding medium. A secondary spring at the front end of the hammer urges it away from the bell after the initial contact, thereby preventing the hammer from damping the vibrations of the bell. The spur **50** pivots back to its resting position engaged against flange **34** of the hammer so that the sequence can be repeated.—WT

5,991,236

43.30.Pc METHOD OF MEASURING BURIED OBJECTS, GEOLOGICAL FORMATIONS AND SEDIMENT PROPERTIES

Tokuo Yamamoto, assignor to Tokuo Yamamoto; Kawasaki Steel Corporation
23 November 1999 (Class 367/41); filed 3 April 1997

A method is described for locating buried objects or for assessing geological features and variation of sediment properties such as density and sound speed. An array of receivers and an acoustic source are placed in the body of water or in the sediment layer. The signals received by the array of sensors are recorded. These received signals are correlated with the transmitted signal to measure arrival time. A space–time wave field diagram is then generated from which any scatterers can be located.—WT

6,160,756

43.30.Pc SEABED SONAR MATRIX SYSTEM

Jacques Y. Guigne, assignor to Guigne International Limited
12 December 2000 (Class 367/88); filed 15 June 1998

A system for mapping a seabed comprises a towed linear array, or planar matrix array, of transducers wherein each is sequentially energized separately at a high frequency to produce a narrow beam and its echoes are detected to produce one pixel of an X-Y display before the next transducer is energized and its echoes detected to produce another pixel of the display corresponding to an adjacent area of the seabed. If a large change in echo amplitude is detected between adjacent sensors, scanning at these locations can be accentuated by sonification at different frequencies.—WT

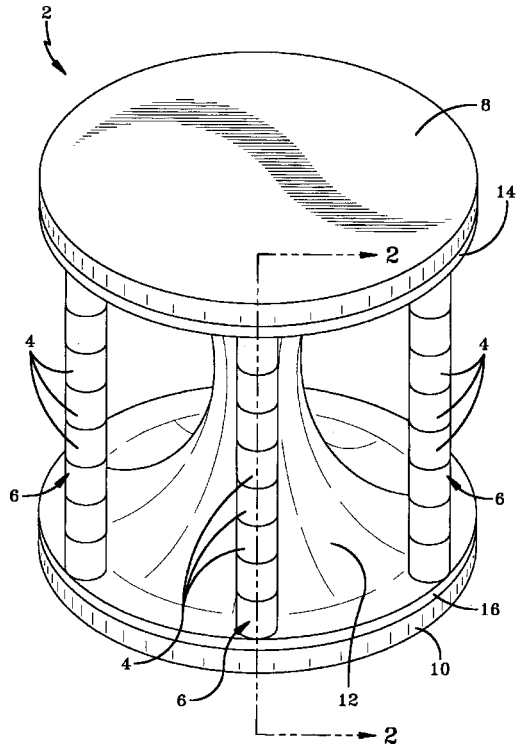
5,875,154

43.30.Yj BARREL STAVE FLEXTENSIONAL PROJECTOR

Robert A. Dechico, assignor to the United States of America as represented by the Secretary of the Navy
23 February 1999 (Class 367/163); filed 13 November 1997

This projector consists of a concave radiating surface **12**, two attached end masses **8** and **10**, and a series of stacks **6** of piezoceramic discs or rings **4** that also join the two end masses but which are located external to the radiator **12**. A minimum of three such stacks is required while six is the suggested number. The concave surface **12** is either fashioned from a number of individual staves, to which a water impervious boot is applied, or is preformed as a continuous surface from a high strength composite material

 0001-4966/2001/110(1)/17/10/$18.00



such as Kevlar or Teflon. Upon electrical excitation, the stacks alternately elongate and shorten causing the concave surface **12** to flex outwardly and inwardly at the same rate, thereby radiating sound into the surrounding medium. The cavity within surface **12** can house tuning coils and electronics. Enhanced cooling of the (presumably encapsulated) stacks **6** of piezoceramic should be realized since they are immersed in the surrounding water rather than interior to the radiating surface as in other designs.—WT

6,160,763

43.30.Yj TOWED ARRAY HYDROPHONE

Jerry C. Cole, assignor to SeaLandAire Technologies, Incorporated
12 December 2000 (Class 367/173); filed 28 December 1998

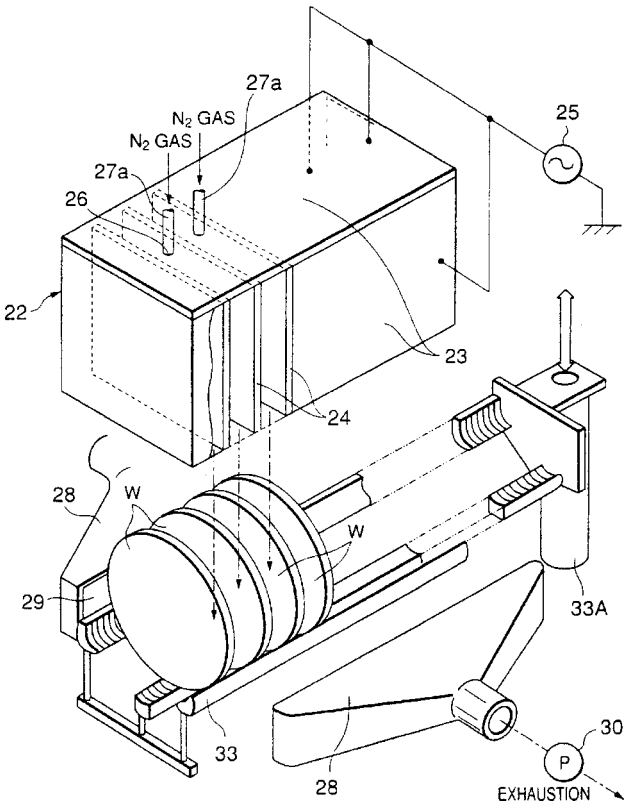
The individual sensor units for a towed line array consist of a pair of circular UNIMORPHS mounted at opposite ends of a cylindrical tube of nylon so as to seal the ends of the tube. The assembly is then suitably potted. The polarities and interconnection of the UNIMORPHS are such that their electrical outputs are additive in response to an incident low-frequency acoustic wave but their presumably equal outputs cancel in response to an axial acceleration of the tube. The units are insensitive to transverse acceleration of the tube. The length of the tube is chosen from buoyancy considerations. Numerous such units are connected in various series-parallel electrical arrangements within the line array to achieve desired electrical characteristics.—WT

6,119,367

43.35.Zc SYSTEM FOR DRYING SEMICONDUCTOR WAFERS USING ULTRASONIC OR LOW FREQUENCY VIBRATION

Yuji Kamikawa and Satoshi Nakashima, assignors to Tokyo Electron Limited
19 September 2000 (Class 34/401); filed in Japan 10 March 1998

Vibrating plates **24** are interleaved between wet semiconductor wafers



W for drying the wafers. Gas can be introduced into the apparatus through inlets **26** and exhausted through suction **28**.—IMH

6,125,704

43.35.Zc ULTRASONIC TECHNIQUE FOR INSPECTION OF WELD AND HEAT-AFFECTED ZONE FOR LOCALIZED HIGH TEMPERATURE HYDROGEN ATTACK

Weicheng David Wang, assignor to Shell Oil Company
3 October 2000 (Class 73/602); filed 16 January 1998

Spectral analysis of pulse-echo signals from piping welds is described for detecting high-temperature hydrogen attack in welds. A signal from a weld which has high-temperature hydrogen attack symptoms is characterized by an amplitude which increases with frequency, while other welds and defects have the same frequency dependence as a reference spectrum.—IMH

6,128,092

43.35.Zc METHOD AND SYSTEM FOR HIGH RESOLUTION ULTRASONIC IMAGING OF SMALL DEFECTS OR ANOMALIES

Daniel Levesque *et al.*, assignors to National Research Council of Canada
3 October 2000 (Class 356/432T); filed 13 July 1999

The patent describes a system for detecting and imaging small defects at or near the surface of objects (possibly in conjunction with laser excitation or interferometry) using synthetic aperture focusing.—IMH

6,094,495

43.38.Ja HORN-TYPE LOUDSPEAKER SYSTEM

Jeffrey A. Rocha, assignor to Eastern Acoustic Works, Incorporated
25 July 2000 (Class 381/340); filed 24 September 1998

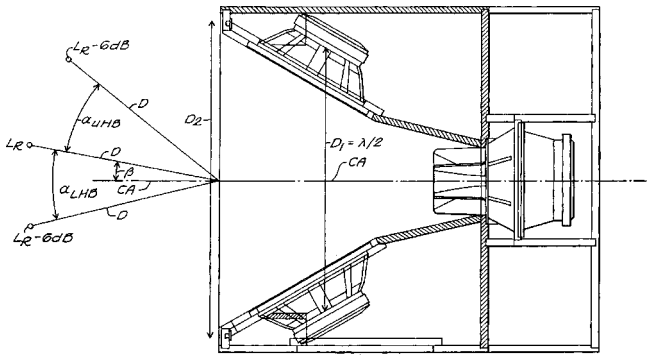
A horn driver has a W-shaped cone or diaphragm mated to an orange slice type phasing plug. In contrast to a simple single-element phasing plug, this geometry operates effectively over a wider frequency range and can generate a plane wavefront in the throat of the horn, even at relatively high frequencies.—GLA

6,118,883

43.38.Ja SYSTEM FOR CONTROLLING LOW FREQUENCY ACOUSTICAL DIRECTIVITY PATTERNS AND MINIMIZING DIRECTIVITY DISCONTINUITIES DURING FREQUENCY TRANSITIONS

Jeffrey A. Rocha, assignor to Eastern Acoustic Works, Incorporated
12 September 2000 (Class 381/387); filed 24 September 1998

The current hot item in touring sound systems is a line array made of identical high-power, wide-range modules. A number of variants are commercially available, but most of these combine a rectangular high-frequency horn with a pair of cone drivers. In several cases, flared horn mouth sections



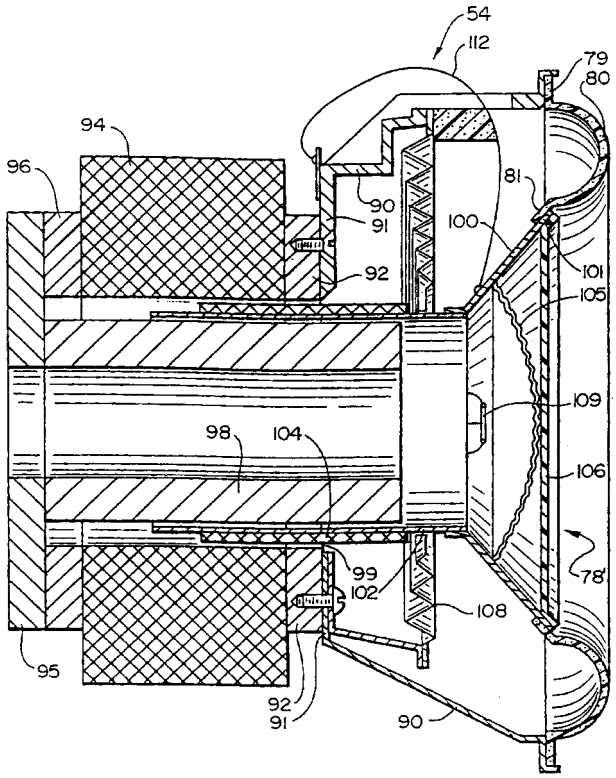
are used as mounting baffles for the cone loudspeakers. If this basic geometry is optimized in relation to the frequency dividing network, then relatively constant beamwidth can be maintained through the crossover region.—GLA

6,130,954

43.38.Ja HIGH BACK-EMF, HIGH PRESSURE SUBWOOFER HAVING SMALL VOLUME CABINET, LOW FREQUENCY CUTOFF AND PRESSURE RESISTANT SURROUND

Robert W. Carver, Snohomish, Washington
10 October 2000 (Class 381/398); filed 2 January 1996

This lengthy patent is the latest in a series of continuations of divisions of earlier patents. The basic invention is a very small, high-power subwoofer which has proved to be commercially successful. A good part of the patent



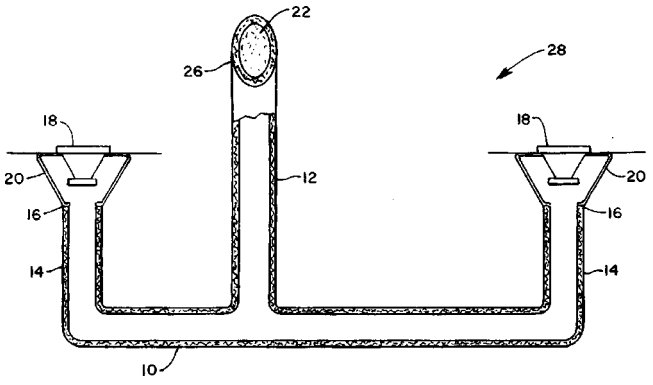
document describes a sophisticated power amplifier which is part of the subwoofer package. The remaining pages are devoted to the inventor's distinctive small-bore, long-stroke theory of loudspeaker design.—GLA

6,141,428

43.38.Ja AUDIO SPEAKER SYSTEM

Chris Narus, Southington, Connecticut
31 October 2000 (Class 381/338); filed 28 October 1993

A bifurcated transmission line loudspeaker system uses two loudspeakers **18**, coupled through individual intermediary ducts **14** to common exit duct **12**. Individual ducts may be any size, any shape, any length, and



lined or not lined with absorptive material. With properly chosen dimensions this scheme can probably deliver smoother response than an equivalent single speaker on an undamped pipe.—GLA

6,141,425

43.38.Lc SOUND QUALITY ADJUSTMENT CIRCUIT

Kazuo Murayama and Toshiya Murakami, assignors to Sony Corporation
31 October 2000 (Class 381/98); filed in Japan 30 September 1996

Conventional $\frac{1}{3}$-octave graphic equalizers have about 30 individual filters on fixed center frequencies. If each filter allowed limited adjustment of center frequency and Q, the number could be greatly reduced while retaining the general look and feel of a graphic equalizer. Such an equalizer was manufactured by Crown more than 10 years ago. This new design by Sony makes use of state-variable filters controlled by variable-conductance amplifiers.—GLA

6,130,870

43.38.Md RECORDING MEDIUM, RECORDING APPARATUS, REPRODUCING METHOD, AND REPRODUCING DEVICE

Teppei Yokota and Ayataka Nishio, assignors to Sony Corporation
10 October 2000 (Class 369/59); filed in Japan 31 May 1995

Standards for digital storage on optical disks restrict audio to 16-bit coding. Encoding schemes using longer word lengths can be stored on CD-ROM, but cannot be reproduced by a standard CD player. The object of the invention is to ''...provide a recording medium which can maintain compatibility with the conventional CD player, and can be reproduced in the state of high sound quality at the reproducing apparatus of high rank.'' In other words, the same CD can be played at 16-bit precision on your auto stereo system or at higher fidelity on a special ''master'' reproducer.—GLA

6,125,175

43.38.Si METHOD AND APPARATUS FOR INSERTING BACKGROUND SOUND IN A TELEPHONE CALL

Randy G. Goldberg *et al.*, assignors to AT&T Corporation
26 September 2000 (Class 379/207); filed 18 September 1997

For those whose lives do not contain enough sound, the inventor (presumably not related to Rube Goldberg) allows a telephone user to add a background sound to a phone call. The user's phone is coupled to an intermediate ''network node'' which can play one of a number of background sounds during the call (e.g., giving the impression that the user is calling from a busy airport or tavern).—IMH

6,130,949

43.38.Si METHOD AND APPARATUS FOR SEPARATION OF SOURCE, PROGRAM RECORDED MEDIUM THEREFOR, METHOD AND APPARATUS FOR DETECTION OF SOUND SOURCE ZONE, AND PROGRAM RECORDED MEDIUM THEREFOR

Mariko Aoki *et al.*, assignors to Nippon Telegraph and Telephone Corporation
10 October 2000 (Class 381/94.3); filed in Japan 18 September 1996

Those not familiar with the field probably don't realize that audio teleconference systems have become elaborate, computer-based processors capable of triggering microphones, cancelling room echo, and suppressing feedback—all in real time. The patent describes a sophisticated system for identifying, coding, and reconstructing individual sound sources in a multi-microphone teleconference setup.—GLA

6,157,680

43.38.Si AUDIO DISTORTION CANCELLER METHOD AND APPARATUS

William L. Betts and Gordon Bremer, assignors to Paradyne Corporation
5 December 2000 (Class 375/285); filed 5 March 1997

This telephone distortion and echo canceller is intended for use on a phone line which also carries data as well as voice, such as a digital subscriber line (DSL). The canceller reduces audible harmonic distortion such as is caused by the connection of a standard telephone instrument to a DSL line. This eliminates the need for the costly and troublesome device known as a POTS splitter.—DLR

6,118,875

43.38.Vk BINAURAL SYNTHESIS, HEAD-RELATED TRANSFER FUNCTIONS AND USES THEREOF

Henrik Møller *et al.*, all of Aalborg, Denmark
12 September 2000 (Class 381/1); filed in Denmark 25 February 1994

This is really a research paper disguised as a patent document. Head-related transfer functions, first investigated more than 30 years ago, are an important part of current sound localization techniques. By processing stereo audio signals with appropriate transfer functions, virtual sound images can be placed almost anywhere. The problem is that HRTFs are like fingerprints—no two are alike and they vary widely between individuals. The patent describes a set of ''general'' HRTFs that are said to satisfy almost all potential users in terms of accurate image placement. The transfer functions themselves, and the means used to derive them, depart significantly from prior art.—GLA

6,118,880

43.38.Vk METHOD AND SYSTEM FOR DYNAMICALLY MAINTAINING AUDIO BALANCE IN A STEREO AUDIO SYSTEM

George Kokkosoulis and Daniel Anthony Temple, assignors to International Business Machines Corporation
12 September 2000 (Class 381/303); filed 18 May 1998

Earlier patents have utilized electro-acoustic schemes to measure and adjust the loudness and delay of multi-channel stereo speakers in relation to a particular listening location. The invention described here differs from prior art in at least three respects: First, all measurements and adjustments are accomplished via a hand-held radio frequency remote control. Second, only relative levels are changed; equalization and delay are not included. Third, the actual method of level adjustment is a lossy sort-of-L-pad between each amplifier output and its associated loudspeaker. The stated application is home theater, yet how many home theater owners would be willing to insert L-pads into their spiral-wound, noninductive, gravity-cancelling, oxygen-free speaker cables?—GLA

6,130,948

43.38.Vk SOUND FIELD PROCESSOR WITH SOUND FIELD EXPANDING APPARATUS

Masamitsu Hirano and Noro Masao, assignors to Yamaha Corporation
10 October 2000 (Class 381/17); filed in Japan 27 September 1996

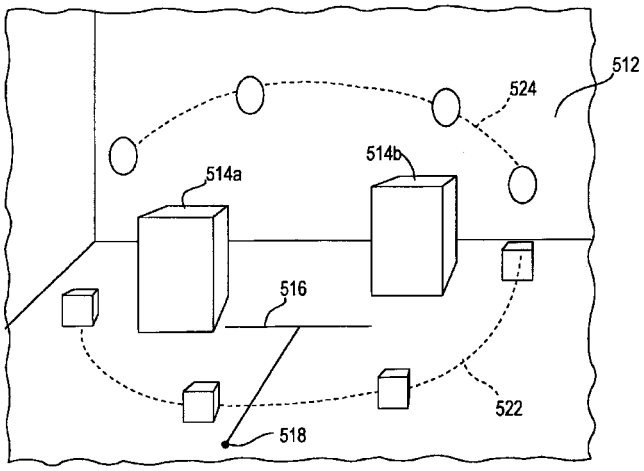
By this time Yamaha must own more than two dozen patents carrying the general title of ''Sound Field Expanding Apparatus.'' The patent describes improved analog filter circuitry to provide a wider range of adjustment and less sensitivity to component tolerances than prior art.—GLA

6,125,115

43.60.Sx TELECONFERENCING METHOD AND APPARATUS WITH THREE-DIMENSIONAL SOUND POSITIONING

Gerard Dirk Smits, assignor to QSound Labs, Incorporated
26 September 2000 (Class 370/389); field 12 February 1998

The patent describes a system for assisting a user participating in a teleconference (carried over a computer network) in discerning the identity of various participants. Spatialization cues are provided by 3-D audio local-



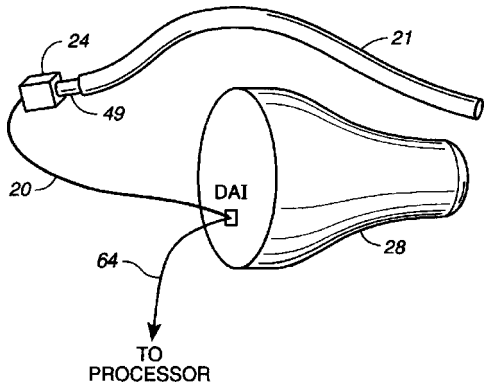
ization to give the user the impression that each unique teleconference participant's voice is coming from a unique location in the surroundings of the user.—IMH

6,154,546

43.66.Ts PROBE MICROPHONE

Bohumir Uvacek, assignor to Resound Corporation
28 November 2000 (Class 381/60); filed 18 December 1997

The output of a calibrated probe microphone is connected to the direct audio input of a digital hearing aid to implement a real ear measurement system, thereby permitting the hearing aid to be tested and calibrated *in situ* without requiring complex external equipment. The digitized representation



of the audio waveform is processed by the digital hearing aid itself or is transmitted via bi-directional link to a remote processor to make fitting computations.—DAP

6,157,727

43.66.Ts COMMUNICATION SYSTEM INCLUDING A HEARING AID AND A LANGUAGE TRANSLATION SYSTEM

Valentin Chapero Rueda, assignor to Siemens Audiologische Technik GmbH
5 December 2000 (Class 381/312); filed in Germany 26 May 1997

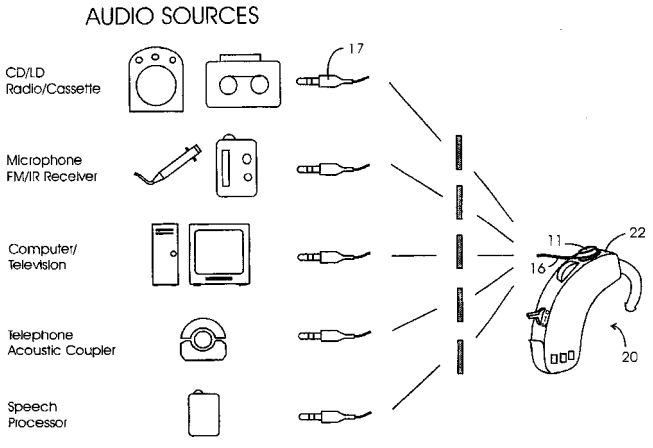
Speech signals are translated into a selected language as well as being processed to correct for the hearing loss of the hearing aid wearer. A communication link facilitates bidirectional data transmission between the hearing aid and a language translation system. The communication link may be wireless via infrared light, ultrasound, induction or high-frequency radio waves. The translated and amplified audio signals are supplied to the hearing aid wearer through the earphone of the hearing aid in the selected second language.—DAP

6,157,728

43.66.Ts UNIVERSAL SELF-ATTACHING INDUCTIVE COUPLING UNIT FOR CONNECTING HEARING INSTRUMENT TO PERIPHERAL ELECTRONIC DEVICES

Yit Chow Tong and Joseph Sylvester Chang, assignors to Multitech Products (PTE) Limited
5 December 2000 (Class 381/331); filed in Singapore 25 May 1996

External electronic devices are connected via an inductive coupling unit to any hearing aid with an internal telecoil. The inductive coupling unit is placed on the outer surface of the hearing aid and contains a transmitting



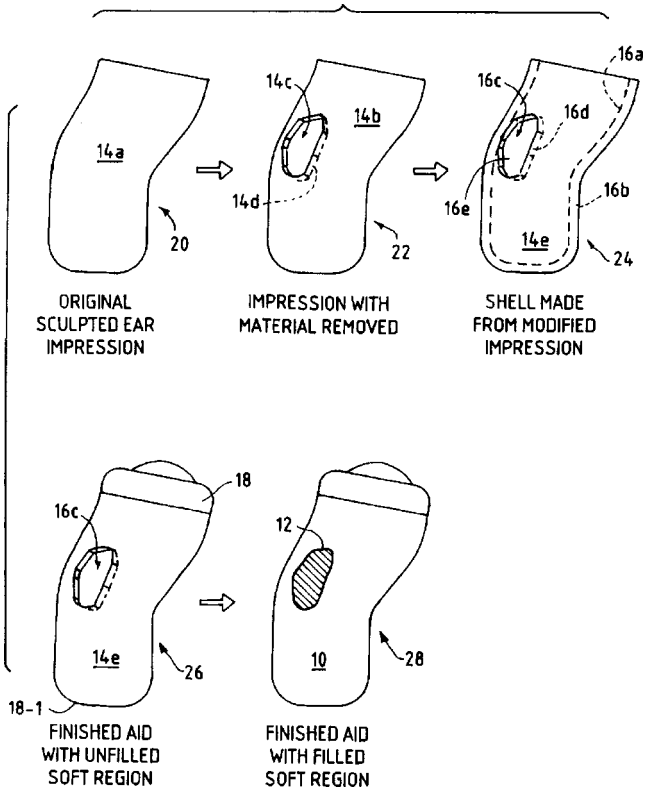
coil that is connected via a cable to the external electronic device. Signals from the external electronic device are induced into the telecoil inside the hearing aid.—DAP

6,167,141

43.66.Ts MULTIMATERIAL HEARING AID HOUSING

Robert S. Yoest, assignor to Beltone Electronics Corporation
26 December 2000 (Class 381/322); filed 30 April 1998

Methodology is described to form a rigid plastic hearing aid housing with a channel for a compliant elastomer inset. The elastomeric material is located on the housing to contact the more sensitive regions of a user's ear



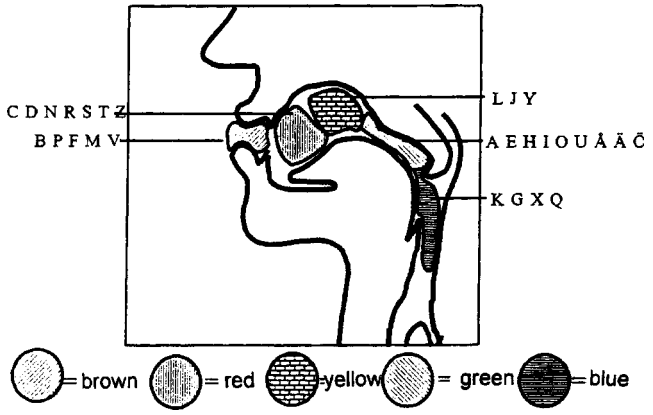
canal for providing improved seal and comfort under both static and dynamic ear canal conditions. Additionally, or alternately, compliant annular seals may be formed around the housing to achieve higher gain without acoustic feedback problems.—DAP

6,151,577

43.70.Dn DEVICE FOR PHONOLOGICAL TRAINING

Ewa Braun, assignor to Ewa Braun
21 November 2000 (Class 704/276); filed in Sweden 27 December 1996

This patent makes the argument that persons with certain types of dyslexic conditions can benefit from a visual display of the vocal tract showing the tongue shape corresponding to the sound being spoken. For example, in a matching mode, a certain speech symbol is entered into the



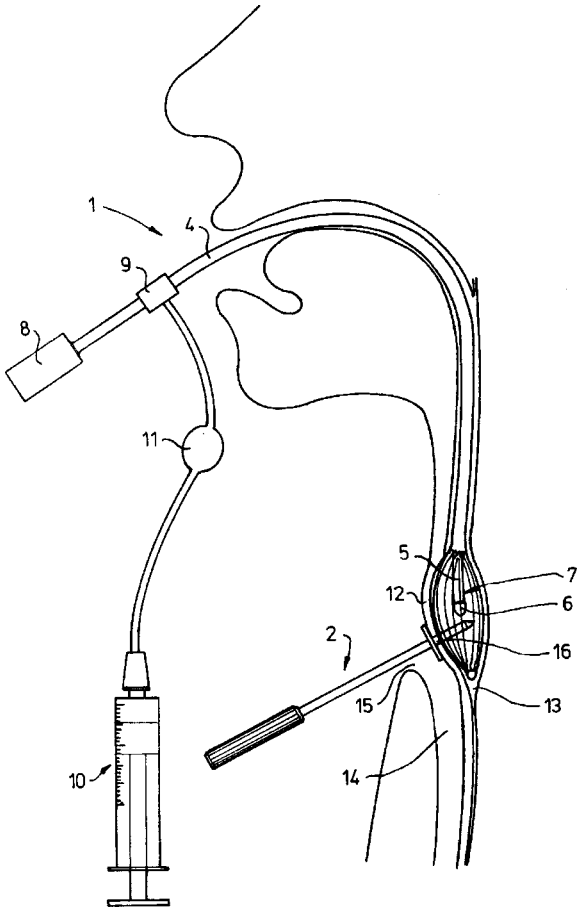
system, which is used to define a target tongue shape. As the person speaks various sounds, the tongue regions are indicated which are most important for the selected sound. Lip shapes and acoustic data displays are also available.—DLR

6,159,243

43.70.Dn KIT FOR IMPLANTATION OF A VOICE PROSTHESIS IN PATIENTS ON WHOM A LARYNGECTOMY HAS BEEN PERFORMED

Paul Ferdinand Schouwenburg, Aerdenhout, The Netherlands
12 December 2000 (Class 623/9); filed in The Netherlands 6 May 1996

A kit is described by which a patient who has previously received a tracheotomy can be fitted with a voice prosthesis without requiring a general anesthesia for the second cut, nor a second period of hospitalization. The device, inserted through the mouth, includes a trocar, or cutting tool, and the



additional parts needed to implant the prosthesis into the fresh cut. A balloon arrangement is said to reduce the chance of cutting into the esophageal wall. There is no discussion of how bleeding into the lungs is to be controlled.—DLR

6,154,721

43.72.Ar METHOD AND DEVICE FOR DETECTING VOICE ACTIVITY

Estelle Sonnic, assignor to U.S. Philips Corporation
28 November 2000 (Class 704/233); filed in France 25 March 1997

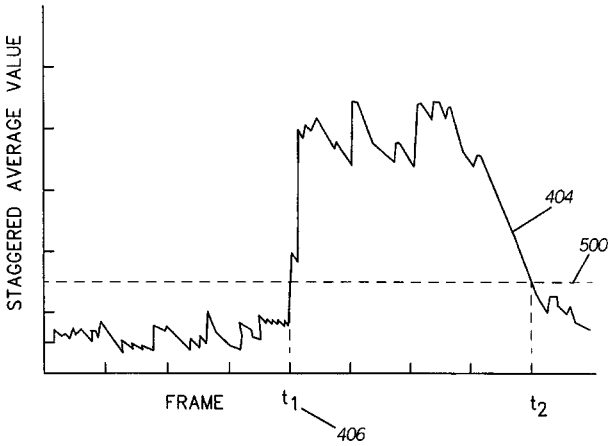
This three-state classifier determines whether frames of the speech input signal consist of noise only or speech plus noise. The energy level and center-clipped zero crossing count are smoothed and compared to thresholds to trigger state changes. The accuracy is affected by the possible presence of data signals coded into a portion of the voice band. The zero-crossing measure is intended to address that question.—DLR

6,157,906

43.72.Ar METHOD FOR DETECTING SPEECH IN A VOCODED SIGNAL

Richard Brent Nicholls *et al.*, assignors to Motorola, Incorporated
5 December 2000 (Class 704/214); filed 31 July 1998

This speech presence detector takes as input the frames of a vocoded speech signal. The design is intended primarily for use with the vector sum (VSELP) vocoding technique. The frame energy is processed by a method called staggered averaging. By this is meant that the input is used directly without delay when that input value increases, but is filtered by an exponential smoothing filter when the input value decreases. The figure shows the result of this filtering. The staggered average signal is multiplied by the voicing level to further improve the discrimination between voiced and unvoiced frames.—DLR



6,157,913

43.72.Ar METHOD AND APPAPARATUS FOR ESTIMATING FITNESS TO PERFORM TASKS BASED ON LINGUISTIC AND OTHER ASPECTS OF SPOKEN RESPONSES IN CONSTRAINED INTERACTIONS

Jared C. Bernstein, Palo Alto, California
5 December 2000 (Class 704/275); filed 25 November 1996

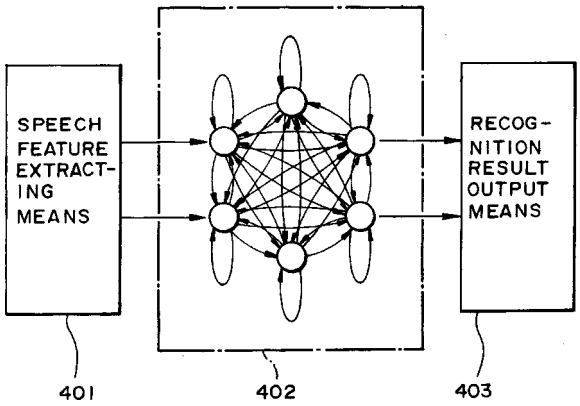
This patent describes a computer program for interacting with a person, to elicit verbal responses from that person, such as answering specific questions, repeating a specific word, phrase or sentence, or identifying objects presented visually or verbally. The computer system then extracts linguistic or nonlinguistic information from the person's responses in order to evaluate the person's psychological or physical state. The stated purpose of the system is to assist in judging the suitability of an applicant for a particular job.—DLR

6,151,592

43.72.Bs RECOGNITION APPARATUS USING NEURAL NETWORK, AND LEARNING METHOD THEREFOR

Mitsuhiro Inazumi, assignor to Seiko Epson Corporation
21 November 2000 (Class 706/16); filed 26 March 1993

This patent is a veritable catalog of neural network architectures, culminating in a fully connected feedback network as shown in the figure. In addition, each neuronlike element contains an internal state representing the past history of all of its inputs. The Kullback–Leibler distance, a measure of the similarity between two probability distributions, is cited as applicable for learning in such a network. Noise inputs are used in various ways to facilitate network learning.—DLR



6,161,090

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

Dimitri Kanevsky and Stephane Herman Maes, assignors to International Business Machines Corporation
12 December 2000 (Class 704/246); filed 11 June 1997

The patent covers a procedure for using voice interaction to gain access to secure facilities or services. A two-pronged approach is described which uses analyses of all submitted voice samples, but also asks the applicant a question from a user-specific and possibly random set and uses the response to help judge the authenticity of the applicant.—DLR

6,161,094

43.72.Fx METHOD OF COMPARING UTTERANCES FOR SECURITY CONTROL

Sherrie Adcock and Gerald S. Schwartz, assignors to Ann Adcock Corporation
12 December 2000 (Class 704/273); filed 14 February 1997

The technique covered here, described as a speaker verification system, performs a simple comparison of spectral data between test and reference utterances. FFT frames of the input signal are tested for nonzero en-

ergy, then divided into *N* bands. Points in the frame/band matrix are differenced between test and reference utterances. The frames being compared may slide back or forward in time for the best match. A set of utterances must all compare within some threshold in order for the matrix average to be stored as a reference for that speaker. The approach is reminiscent of techniques widely used in the 1970s.—DLR

6,163,769

43.72.Ja TEXT-TO-SPEECH USING CLUSTERED CONTEXT-DEPENDENT PHONEME-BASED UNITS

Alejandro Acero *et al.*, assignors to Microsoft Corporation
19 December 2000 (Class 704/260); filed 2 October 1997

This concatenative speech synthesizer uses triphone or larger speech segments as a basis for natural-sounding speech output. An automated system for creating the speech segment database uses a decision tree and builds a hidden Markov model to identify and cluster similar phonetic elements within similar contexts. An unlabeled corpus of speech from a target speaker may thus be processed to build up a set of segments for use by the speech synthesizer.—DLR

6,157,909

43.72.Kb PROCESS AND DEVICE FOR BLIND EQUALIZATION OF THE EFFECTS OF A TRANSMISSION CHANNEL ON A DIGITAL SPEECH SIGNAL

Laurent Mauuary and Jean Monne, assignors to France Telecom
5 December 2000 (Class 704/228); filed in France 22 July 1997

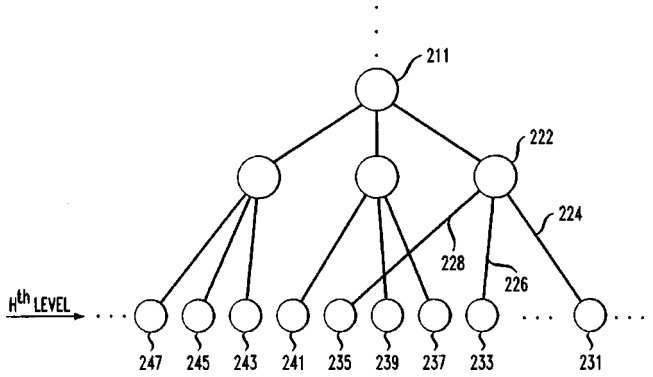
The patent describes a method equalizing a communication channel using a cepstral analysis of the received signal. An average of the short-term cepstral vectors is computed and compared to a stored reference cepstrum derived from a long-term analysis of clear speech. The difference between the cepstrum of the received signal and the reference is used to compute an equalization filter to compensate for line distortions.—DLR

6,151,574

43.72.Ne TECHNIQUE FOR ADAPTATION OF HIDDEN MARKOV MODELS FOR SPEECH RECOGNITION

Chin-Hui Lee and Koichi Shinoda, assignors to Lucent Technologies, Incorporated
21 November 2000 (Class 704/256); filed 5 December 1997

A method is presented for adapting the HMM speaker model of a speech recognizer to a particular speaker's voice without a specific training period. An initial (factory) training provides a generalized speaker-independent set of HMM models. During recognition, successive portions of



the spectral distributions are removed according to the acoustic feature data. Any remaining residual in the distribution of the classified model is used to adapt the corresponding factory HMM.—DLR

6,151,575

43.72.Ne RAPID ADAPTATION OF SPEECH MODELS

Michael Jack Newman *et al.*, assignors to Dragon Systems, Incorporated
21 November 2000 (Class 704/260); filed 28 October 1996

Collections of speech data to be used in training a recognizer are grouped according to similar characteristics, which may include common patterns such as gender, dialect patterns, or age. Working in the acoustic feature domain, linear interpolators are constructed between like elements from different domains. Various forms of averaging may be used to estimate the parameters for missing data. A model may be constructed for a particular speaker by estimating the distance to move along each linear measure.—DLR

6,154,526

43.72.Ne DATA ACQUISITION AND ERROR CORRECTING SPEECH RECOGNITION SYSTEM

Karl Dahlke *et al.*, assignors to Intellivoice Communications, Incorporated
28 November 2000 (Class 379/88.03); filed 4 December 1996

This speech recognizer offers a proposed solution to the problem of recognition accuracy in a voice dialing system. The system controller maintains a statistical database for each user, including numbers typically called by area code and by time of day. This information is considered along with the recognition confidence to determine whether to suggest a ''correction'' to the user before dialing.—DLR

6,154,722

43.72.Ne METHOD AND APPARATUS FOR A SPEECH RECOGNITION SYSTEM LANGUAGE MODEL THAT INTEGRATES A FINITE STATE GRAMMAR PROBABILITY AND AN *N*-GRAM PROBABILITY

Jerome R. Bellegarda, assignor to Apple Computer, Incorporated
28 November 2000 (Class 704/257); filed 18 December 1997

This speech recognition system uses an *n*-gram model (typically, with *n* equals 2) but having the transition probabilities adjusted according to a finite state grammar of the input language. These two constraints are weighted such that an utterance need not be strictly matched by the finite state grammar and still recognized as valid as long as the word sequence probabilities are reasonable.—DLR

6,157,705

43.72.Ne VOICE CONTROL OF A SERVER

Jeffrey Perrone, assignor to E*Trade Group, Incorporated
5 December 2000 (Class 379/88.01); filed 5 December 1997

This application of a speech recognition system would allow a user to control an Internet server using a standard telephone connection to the server site. The phone connection may also be made over the network in an alternative arrangement. All of the speech recognition software is located at the server, rather than at the user's site. It is assumed without further discussion that a typical commercial voice recognition system will be able to provide sufficiently accurate results given unconstrained, natural speech via a telephone connection.—DLR

6,157,910

43.72.Ne DEFERRED CORRECTION FILE TRANSFER FOR UPDATING A SPEECH FILE BY CREATING A FILE LOG OF CORRECTIONS

Kerry A. Ortega, assignor to International Business Machines Corporation
5 December 2000 (Class 704/231); filed 31 August 1998

A method is described by which multiple users of a speech recognition system can make updates to the recognition vocabulary or word pronunciations without interfering with each other's corrections. When any user makes corrections to the system, those changes are not applied directly to the system but are instead maintained in a log file of corrections. Any user can then keep a copy of the original master system and his or her own log of changes. These changes will be applied each time that user runs the system on any machine.—DLR

6,157,911

43.72.Ne METHOD AND A SYSTEM FOR SUBSTANTIALLY ELIMINATING SPEECH RECOGNITION ERROR IN DETECTING REPETITIVE SOUND ELEMENTS

Masaru Kuroda, assignor to Ricoh Company, Limited
5 December 2000 (Class 704/251); filed in Japan 28 March 1997

When using a typical speech recognition system, short repetitive words will often not be correctly recognized unless isolated by a short silent interval. This causes an unacceptable delay in the system's response. One alternative is to use a word spotting method, which can locate a word without regard to word or speech boundaries. This creates a problem, however, for words such as Japanese ''nana'' for ''seven.'' Two repetitions of ''nana'' are seen by the word spotter as three utterances of the word. The solution described here is to force a specific minimum interval between word-spotting results. Thus, the word spotter can still respond quickly, but will fail if a response is too close in time to a previous response.—DLR

6,157,912

43.72.Ne SPEECH RECOGNITION METHOD WITH LANGUAGE MODEL ADAPTATION

Reinhard Kneser *et al.*, assignors to U. S. Philips Corporation
5 December 2000 (Class 704/270); filed in Germany 28 February 1997

This speech recognition patent deals with the question of adapting an *n*-gram probabilities model based on a large text sample to the more specific patterns found in particular usage areas. The latter patterns may not provide nearly as large a text sample as the broader training sample. The method used is to compute the individual word frequencies of the smaller corpus (the unigram model), and then use these word frequency values to adjust the large-corpus *n*-gram probabilities to make them better fit the small corpus.—DLR

6,160,877

43.72.Ne METHOD OF SCREENING AND PRIORITIZING AN INCOMING CALL

Gregory R. Tatchell *et al.*, assignors to Stentor Resource Centre, Incorporated
12 December 2000 (Class 379/197); filed 19 November 1996

This speech recognition system is incorporated into the telephone system to allow the subscriber to selectively control the various phone system services in a convenient and understandable manner. The primary goal seems to be a more comprehensive control of a caller ID service than the typical options of simply rejecting all calls or rejecting all calls which have the ID signal blanked out. This system accepts names defined for the voice dialing service and allows calls to be accepted selectively according to the caller ID signal.—DLR

6,161,087

43.72.Ne SPEECH-RECOGNITION-ASSISTED SELECTIVE SUPPRESSION OF SILENT AND FILLED SPEECH PAUSES DURING PLAYBACK OF AN AUDIO RECORDING

Colin W. Wightman and Joan Bachenko, assignors to Lernout & Hauspie Speech Products N.V.
12 December 2000 (Class 704/215); filed 5 October 1998

This playback system for audio speech recordings provides a means of automatically suppressing pauses during which the speaker has uttered fill sounds, such as ''um'' or ''uh.'' An indication is provided to the listener of when such filled pauses have been suppressed and the user may turn such suppression on or off at any time.—DLR

6,163,765

43.72.Ne SUBBAND NORMALIZATION, TRANSFORMATION, AND VOICENESS TO RECOGNIZE PHONEMES FOR TEXT MESSAGING IN A RADIO COMMUNICATION SYSTEM

Oleg Andric *et al.*, assignors to Motorola, Incorporated
19 December 2000 (Class 704/204); filed 30 March 1998

This telephone message system converts a caller's message to text for display on the phone of the paged person. The caller may listen to a synthesized version of the converted text to verify correct recognition. The analysis process involves generating cepstral difference envelope vectors and transforming the envelope data to get voicing energy. These two measures are used to generate Gaussian mixture probability density functions to identify the likely phoneme sequence for the utterance.—DLR

6,163,767

43.72.Ne SPEECH RECOGNITION METHOD AND SYSTEM FOR RECOGNIZING SINGLE OR UN-CORRELATED CHINESE CHARACTERS

Donald T. Tang *et al.*, assignors to International Business Machines Corporation
19 December 2000 (Class 704/231); filed in China 19 September 1997

Because each Chinese character typically has a large number of homophones, correct recognition of a spoken, isolated character is essentially impossible. The solution proposed here is to describe the character using

any of a variety of more or less standard phrases, collectively referred to as a character description language (CDL). The recognizer interprets the sentence in CDL, from which the correct character may be extracted. Much of the patent is concerned with details of the available CDL techniques and the grammars by which they may be analyzed.—DLR

6,163,768

3.72.Ne NON-INTERACTIVE ENROLLMENT IN SPEECH RECOGNITION

Stefan Sherwood *et al.*, assignors to Dragon Systems, Incorporated
19 December 2000 (Class 704/235); filed 15 June 1998

It is fairly common to use recognition techniques to process speech inputs intended for updating a speaker model for a particular speaker. This system uses such techniques to allow the user to record an extended selection from the enrollment material without being especially concerned that the recording contain the exact word sequence. When the recording is played back into the recognizer for updating purposes, corresponding words and phrases are recognized and located in the enrollment material. Once located, the new words may be used for the update.—DLR

5,850,372

43.80.Nd METHOD OF ALERTING SEA COWS OF THE DANGER OF APPROACHING MOTOR VESSELS

Joseph E. Blue, assignor to Leviathan Legacy, Incorporated
15 December 1998 (Class 367/139); filed 9 December 1994

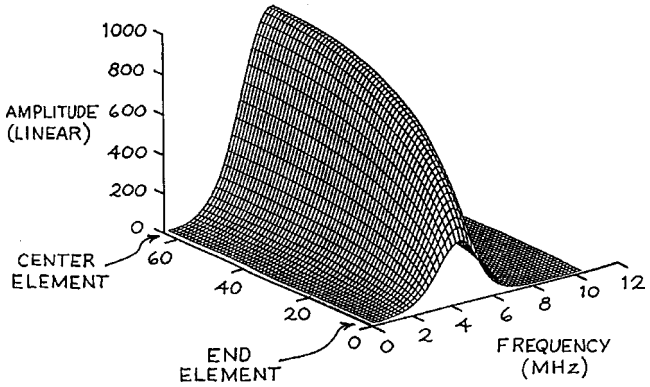
This patient discusses the concept of generating underwater acoustic signals in the frequency range from 3 to 26 kHz, either with linear sonar transducers or nonlinear parametric sonar projectors, and of projecting these signals in the direction of travel of a motor vessel, expressly for the purpose of alerting submerged manatees near the ocean's surface to the danger of that approaching vessel.—WT

6,122,222

43.80.Vj ULTRASONIC TRANSMIT AND RECEIVE SYSTEM

John A. Hossack and Jian-Hua Mo, assignors to Acuson Corporation
19 September 2000 (Class 367/7); filed 2 March 1995

A transmit beamformer is provided, having an array whose elements are excited with at least two frequencies. The result produces a multi-focal beam with progressively higher frequency foci located at progressively shallower focal distances. The invention claims to use the frequency-dependent



line focus to increase the rate at which an image can be generated while maintaining multiple focal points and eliminating the need for multiple bursts at multiple steering positions.—IMH

ADVANCED-DEGREE DISSERTATIONS IN ACOUSTICS

Editor's Note: *Abstracts of Doctoral and Master's theses will be welcomed at all times. Please note that they must be limited to 200 words, must include the appropriate PACS classification numbers, and formatted as shown below. If sent by postal mail, note that they must be double spaced. The address for obtaining a copy of the thesis is helpful. Submit abstracts to: Acoustical Society of America, Thesis Abstracts, Suite 1N01, 2 Huntington Quadrangle, Melville, NY 11747-4502, e-mail: asa@aip.org*

On the subjective responses based on the auditory-brain model in relation to the factors extracted from the interaural cross-correlation mechanism and the auto-correlation mechanism of sound fields [43.55.Gx, 43.55.Hy, 43.66.Lj]—Shin-ichi Sato, *Graduate School of Science and Technology, Kobe University, Kobe, Japan, March 1999 (Ph.D.).* Based on the auditory-brain model for subjective responses, the author investigated the relationship between several important subjective attributes of the sound field and the orthogonal factors that are extracted from both the interaural cross-correlation function and the autocorrelation function. There are three significant findings. (1) The apparent source width (ASW) is described by two orthogonal spatial factors: the IACC and the width of the interaural cross-correlation function (W_{IACC}). The scale value of ASW is formulated by superpositioning two terms: the $\frac{3}{2}$ power of the IACC and the $\frac{1}{2}$ power of the W_{IACC}. (2) The most preferred delay time of a single reflection for a cello soloist can be calculated by the amplitude of the reflection and the minimum value of the effective duration of the running autocorrelation function of the music motifs played by the cellist. (3) From subjective preference tests by paired-comparison method, for different source locations on the stage in an actual concert hall, the subjective preference can be calculated using four orthogonal factors and the interaural time delay of the IACC ($_{\mathrm{IACC}}$). When the IACC is not obtained at $_{\mathrm{IACC}}=0$, the preference scores decrease rapidly due to an image shift or an unbalanced sound field. Results agree well with the subjective preference obtained by use of simulated sound fields when $_{\mathrm{IACC}}=0$.

Thesis advisor: Yoichi Ando

Copies of this thesis written in English can be obtained from Shin-ichi Sato, Graduate School of Science and Technology, Kobe University, Rokkodai, Nada, Kobe 657-8501, Japan. E-mail address: satos@kobe-u.ac.jp.

Computational systems for sound fields, as tools in design and diagnosis [43.50.Rq, 43.55.Hy, 43.55.Ka, 43.55.Mc, 43.58.Gn, 43.58.Ta]—Masatsugu Sakurai, *Graduate School of Science and Technology, Kobe University, Kobe, Japan, March 2000 (Dr. Eng.).* Computational systems for design and diagnosis of sound field were developed. The concept of these systems are based on the model of human auditory-brain system [Y. Ando, *Architectural Acoustics-Blending Sound Sources, Sound Fields, and Listeners* (AIP/Springer-Verlag, New York, 1998)], including the autocorrelation mechanisms, interaural cross-correlation mechanism, and the specialization of the human cerebral hemispheres. Thesis consists of the following four parts: (1) Computer simulation system for calculating the four orthogonal factors (LL, t_1, T_{sub} and IACC) at each seat using architectural drawings were developed. The accuracy of this system as examined by comparing the calculated factors with the measured ones at Kirishima International Concert Hall. (2) The diagnostic system to measure the four orthogonal factors and additional factors ($_{\mathrm{IACC}}$, W_{IACC} and A-value) was developed. This system can be used for the measurement in the scale model. The effect of the reflectors' array above stage which cannot calculate at the design stage was evaluated at the scale model measurement of Tsuyama music cultural hall. (3) As an application of the subjective preference theory, seat selection system to maximize the preference of each individual was developed. This system, which is installed in Kirishima International Concert Hall, can find the seat where the preference is maximized by the paired comparison test with respect to the four orthogonal factors. (4) Noise measurement system to measure and identify the environment noise using the factors extracted from the autocorrelation function ((0), $_e$, $_1$ and $_1$) and the factors extracted from the interaural cross-correlation function (IACC, $_{\mathrm{IACC}}$, and W_{IACC}) was developed. As an example, noises from different car engines were characterized by these factors measured by this system.

Thesis advisor: Yoichi Ando

Copies of this thesis written in English can be obtained from Masatsugu Sakurai, Yoshimasa Electronic, Inc., Daiichi-Nishiwaki Bldg., 1-58-10 Yoyogi, Shibuya, Tokyo 151-0053, Japan. E-mail address: saku@ymec.co.jp.

Reverberance and its control in relation to the physical factors of sound fields in halls [43.55.Hy, 43.55.Fw]—Shigeo Hase, *Graduate School of Science and Technology, Kobe University, Kobe, Japan, March 2001 (Dr. Eng.).* Based on the theory of subjective preference, the author investigated the effects of the set of four orthogonal physical factors (listening level, initial time delay gap between the direct sound and the first reflection, subsequent reverberation time, and IACC) on the reverberance of sound fields. The study clarifies how much both temporal and spatial factors contribute to reverberance. Experiments were conducted to determine the effects of the physical factors on reverberance in simulated sound fields and in actual halls. Scale values of reverberance for both music and speech signals were obtained by the paired-comparison method. The results of these experiments show that these four factors affect the scale values of reverberance independently, and that the scale values of a sound field at any seat in a room can be calculated by the linear combinations of these factors. On the basis of the results from this and previous studies, some methods for controlling the reverberance of sound fields in an actual hall are proposed. The results of this study will allow the design of various reverberances for sound fields that are optimal for various source signals with its effective duration of autocorrelation function, $(_e)_{\min}$.

Thesis advisor: Yoichi Ando

Copies of this thesis written in English can be obtained from Shigeo Hase, Theater Design Laboratory Co., 2-13-26 Makinohonmachi, Hirakata, Osaka 573-1144, Japan. E-mail address: shigeo.hase@nifty.ne.jp.

Blending architectural and acoustic factors in designing an event hall [43.55.Br, 43.55.Fw, 43.55.Gx, 43.55.Hy]—Akio Takatsu, *Graduate School of Science and Technology, Kobe University, Kobe, Japan, March 2001 (Dr. Eng.).* In order to synthesize both architectural design and acoustic design consisting of temporal and spatial design, a design process in which these elements are all considered is proposed in this study. Using this design process, a multi-purpose event hall, which is part of a larger architectural complex, was designed. This hall was measured for both temporal and spatial orthogonal factors after construction, with favorable results. One goal of this study was to solve the acoustic problems caused by the rounded shape of the event hall, where the architectural design had already been determined by the theme and concept of the complex. In a worst-case scenario, had the acoustic problems been solved without unduly affecting the architecture of the hall, the process would have been considered to be successful. Remarkably, some knowledge about methods to solve acoustic problems caused by the round shaped architecture was obtained through the design process blending architectural and acoustic factors.

Thesis advisor: Yoichi Ando

Copies of this thesis written in English can be obtained from Akio Takatsu, Showa Sekkei Co., 1-2-1-800 Benten, Minato-ku, Osaka 552-0007, Japan.

Acoustic measurements and the individual subjective evaluation for sound fields [43.55.Br, 43.55.Gx, 43.55.Hy]—Hiroyuki Sakai, *Graduate School of Science and Technology, Kobe University, Kobe, Japan, March 2001 (Ph.D.).*Acoustical measurements and individual subjective evaluations for sound fields were performed based on a model of the human auditory-brain system and subjective preference theory. In the acoustic measurement section, orthogonal factors in historical opera houses and outdoor sound fields in forests are discussed. In the subjective evaluation section, individual differences and intra-individual changes of subjective preference are described in terms of the orthogonal factors. The main conclusions are as follows: (1) In acoustic measurements in a forest and in a bamboo forest, it is found that multiple scattering effects between trees or bamboo improve IACC. (2) Results of a preference test in an existing opera house show that subjective preference theory can be applied to opera houses, including the orchestra pit and the box seats. (3) Coefficient is introduced for individual preference in a sound field. (4) The investigations of preference tests in relation to listening level (LL) and subsequent reverberation time (T_{sub}) for a simulated sound field show that subjects with a large value of α have small intra-individual changes, so the range of its preferred value is small. On the other hand, subjects with small value have minor preference as its parameter is changed.

Thesis advisor: Yoichi Ando

Copies of this thesis written in English can be obtained from Hiroyuki Sakai, Graduate School of Science and Technology, Kobe University, Rokkodai, Nada, Kobe 657-8501, Japan. E-mail address: sakaih@kobe-u.ac.jp.

The role of the bandwidth-duration product WT in the detectability of diotic signals [43.66.Ba, 43.66.Dc, 43.66.Fe]—Judi A. Lapsley Miller, *Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand (Ph.D.).*The bandwidth-duration product, WT, is a fundamental parameter in most theories of aural amplitude discrimination of Gaussian noise. These theories predict that detectability is dependent on WT, but not on the individual values of bandwidth and duration. Due to the acoustical uncertainty principle, it is impossible to completely specify an acoustic waveform with both finite duration and finite bandwidth. An observer must decide how best to trade off information in the time domain with information in the frequency domain. As Licklider (1963) states, ''The nature of [the ear's] solution to the time-frequency problem is, in fact, one of the central problems in the psychology of hearing.'' This problem is still unresolved, primarily due to observer inconsistency in experiments, which degrades performance making it difficult to compare models. The aim was to compare human observers' ability to trade bandwidth and duration, with simulated and theoretical observers. Human observers participated in a parametric study where the bandwidth and duration of 500 Hz noise waveforms was systematically varied for the same bandwidth-duration products (WT=1, 2, and 4, where W varied over 2.5–160 Hz, and T varied over 400–6.25 ms, in octave steps). If observers can trade bandwidth and duration, detectability should be constant for the same WT. The observers replicated the experiments six times so that group operating characteristic (GOC) analysis could be used to reduce the effects of their inconsistent decision making. Asymptotic errorless performance was estimated by extrapolating results from the GOC analysis, as a function of replications added.

Three simulated ideal observers: the energy, envelope, and full-linear (band-pass filter, full-wave rectifier, and true integrator) detectors were compared with each other, with mathematical theory and with human observers. Asymptotic detectability relative to the full-linear detector indicates that human observers best detect signals with a bandwidth of 40–80 Hz and a duration of 50–100 ms, and that other values are traded off in approximately concentric ellipses of equal detectability. Human detectability of Gaussian noise was best modeled by the full-linear detector using a nonoptimal filter. Comparing psychometric functions for this detector with human data shows many striking similarities, indicating that human observers can sometimes perform as well as an ideal observer, once their inconsistency is minimized. These results indicate that the human hearing system can trade bandwidth and duration of signals, but not optimally. This accounts for many of the disparate estimates of the critical band, rectifier, and temporal integrator, found in the literature, because (a) the critical band is adjustable, but has a minimum of 40–50 Hz, (b) the rectifier is linear, rather than square-law, and (c) the temporal integrator is either true or leaky with a very long time constant.

Thesis advisor: Dr. John K. Whitmore

Copies of this thesis can be obtained from Judi Lapsley Miller, NSMRL, Subase NLON Box 900, Groton, CT 06349-5000, E-mail address: judi@psychophysics.org or by dowloading from <www.psychophysics.org>.

Energy radiation from point sources whose duration of accelerated motion is finite

H. Levine
Mathematics Department, Stanford University, Stanford, California 94305-2125

G. C. Gaunaurd[a)]
Carderock Division, Naval Surface Warfare Center, West Bethesda, Maryland 20817-5700

(Received 1 October 2000; accepted for publication 16 April 2001)

A detailed analysis is given of the energy and power measures associated with point sources which proceed along a rectilinear path in a prescribed and continuous manner that involves a limited period of acceleration. Attention is focused on monopole source types either singly or in pairs. The consideration of a single monopole precedes that of a ''colliding'' pair of such sources that have either equal or opposite source strengths. In all cases the angular and the joint angular and frequency distribution of the radiated energy, as well as the total radiated energy, are found at subsonic Mach numbers, and, for the paired sources, estimates are contingent on small Mach numbers. Numerical plots of the angular distributions are given, and applications of the analysis include the elementary modeling of the sound fields radiated by new aircrafts capable of completely reversing their course in minute distances. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1379728]

PACS numbers: 43.20.Px, 43.20.Rz, 43.28.Ra [ANN]

I. INTRODUCTION

The analytical aspects of radiation generated by sources in unsteady motion have been described comprehensively,[1–7] with numerous examples relating to monopole, dipole and quadrupole models already contained in the literature. The present article furnishes new results for pointlike sources that travel at constant speed along a straight line, apart from a finite time interval after which their direction of motion is reversed. In addition to cases of isolated source motion, those of ''colliding'' sources which initially approach and then recede from each other also receive attention.

The analysis presupposes a linear inhomogeneous wave equation:

$$\left(\nabla^2-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\phi(\mathbf{r},t)=-\frac{1}{\rho_0}Q(\mathbf{r},t) \tag{1}$$

for the scalar quantity $\phi(\mathbf{r},t)$ representing the velocity potential in a homogeneous medium with sound speed c and equilibrium density ρ_0; here a general source function $Q(\mathbf{r},t)$ is present. Expressions for the power and energy that find direct use in the subsequent development are known.[5–10] Thus adopting the notations $\mathbf{n}$, $d\Omega_n$, δ for an arbitrarily oriented unit vector, the element of solid angle about the latter direction and the symbolic delta function, the relations

$$P(t)=\int Q(\mathbf{r},t)\frac{\partial\phi}{\partial t}d\mathbf{r}=\int P(\mathbf{n},t)\,d\Omega_n \tag{2}$$

and

$$P(\mathbf{n},t)=-\frac{1}{16\rho_0 c\pi^2}\int_{-\infty}^{\infty}Q(\mathbf{r},t)Q(\mathbf{r}',t')\times\delta''\left(t'-t+\frac{\mathbf{n}}{c}\cdot(\mathbf{r}-\mathbf{r}')\right)d\mathbf{r}\,d\mathbf{r}'\,dt' \tag{3}$$

specify the overall instantaneous power $P(t)$ radiated from the source into its surroundings and the amount $P(\mathbf{n},t)$ in a specific direction. The corresponding measures

$$E(\mathbf{n})=\int_{-\infty}^{\infty}P(\mathbf{n},t)\,dt \tag{4}$$

$$=-\frac{1}{16\rho_0 c\pi^2}\int_{-\infty}^{\infty}Q(\mathbf{r},t)Q(\mathbf{r}',t')\times\delta''\left(t'-t+\frac{\mathbf{n}}{c}\cdot(\mathbf{r}-\mathbf{r}')\right)d\mathbf{r}\,d\mathbf{r}'\,dt\,dt' \tag{5}$$

$$=\int_{-\infty}^{\infty}E(\mathbf{n},\omega)\,d\omega \tag{6}$$

and

$$E(\mathbf{n},\omega)=\frac{\omega^2}{2\pi\rho_0 c(4\pi)^2}\left|\int_{-\infty}^{\infty}Q(\mathbf{r},t)e^{i\omega(t-\mathbf{n}\cdot\mathbf{r}/c)}d\mathbf{r}\,dt\right|^2 \tag{7}$$

pertain to the directional and also to the frequency distribution of the radiated energy.

The particular source function

$$Q(\mathbf{r},t)=f(t)\,\delta[x-x_s(t)]+g(t)\frac{\partial}{\partial x}\delta[x-x_s(t)] \tag{8}$$

characterizes a pair of point sources in motion along the x axis, one constituting a monopole and the other a dipole with collinear axis. In this circumstance, the power output $P(\mathbf{n},t)$ is given by

[a)]Address of new affiliation: Army Research Laboratory, Code AMSRL-SE-RU, 2800 Powder Mill Road, Adlephi, MD 20783-1197. Electronic mail: ggaunaurd@arl.army.mil

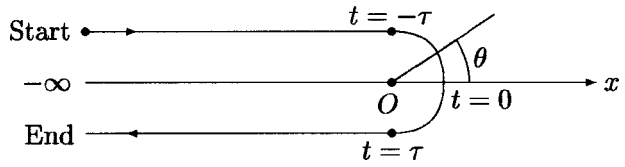


FIG. 1. Rectilinear path of the single source in Sec. II A showing its change of direction around $t=0$.

$$P(\mathbf{n},t)=\frac{1}{16\rho_0 c\pi^2[1-M(t)\cos\theta]}\left[\frac{d}{dt}\left\{\frac{f(t)}{1-M(t)\cos\theta}-\frac{1/c}{1-M(t)\cos\theta}\frac{d}{dt}\left(\frac{g(t)\cos\theta}{1-M(t)\cos\theta}\right)\right\}\right]^2, \quad (9)$$

where

$$M(t)=\frac{1}{c}\frac{d}{dt}x_s(t)<1, \quad (10)$$

and the polar angle θ is measured from the x direction. For an isolated monopole of fixed strength q, Eq. (9) reduces to Eq. (7):

$$P(\mathbf{n},t)=\frac{q^2\cos^2\theta}{16\rho_0 c\pi^2[1-M(t)\cos\theta]^5}\left(\frac{dM}{dt}\right)^2. \quad (11)$$

The preceding expressions are employed in the next section to determine the radiation characteristics of isolated monopoles which undergo reversals in direction during an otherwise steady linear motion. Sections II B and II C are concerned with colliding monopoles of either equal or opposite strengths.

II. ANALYSIS OF THREE CONFIGURATIONS

A. A monopole in unsteady motion

Suppose that the position $x_s(t)$, speed $v_s(t)$, and acceleration $\dot{v}_s(t)$ of a point which moves along the half-line $-\infty<x<0$ are given by

$$x_s(t)=\begin{cases} v_0(t+\tau-2\tau/\pi), & -\infty<t<-\tau,\\ 2v_0(\tau/\pi)[\cos(\pi t/2\tau)-1], & -\tau<t<\tau,\\ v_0(-t+\tau-2\tau/\pi), & \tau<t<\infty,\end{cases} \quad (12)$$

$$v_s(t)=\begin{cases} v_0, & -\infty<t<-\tau,\\ -v_0\sin(\pi t/2\tau), & -\tau<t<\tau,\\ -v_0, & \tau<t<\infty,\end{cases} \quad (13)$$

$$\dot{v}_s(t)=\begin{cases} 0, & -\infty<t<-\tau,\\ -v_0\pi/(2\tau)\cos(\pi t/2\tau), & -\tau<t<\tau,\\ 0, & \tau<t<\infty,\end{cases} \quad (14)$$

as shown in Fig. 1. Let these be the kinematical parameters of a coincident monopole with fixed strength q. On substituting the pertinent source function,

$$Q(x,t)=q\,\delta[x-x_s(t)], \quad (15)$$

into the integral, Eq. (7), for $E(\mathbf{n},\omega)$, it is found that

$$\omega\int_{-\infty}^{\infty}Q(\mathbf{r},t)e^{i\omega(t-\mathbf{n}\cdot\mathbf{r}/c)}\,d\mathbf{r}\,dt$$

$$=\omega q\int_{-\infty}^{\infty}e^{i\omega[t-\cos\theta x_s(t)/c]}\,dt \quad (16)$$

$$=\omega q\int_{-\infty}^{-\tau}e^{i\omega[t-M\cos\theta(t+\tau-2\tau/\pi)]}\,dt+\int_{-\tau}^{\tau}e^{i\omega(t-(2M\tau/\pi)\cos\theta(\cos(\pi t/2\tau)-1))}\,dt+\int_{\tau}^{\infty}e^{i\omega[t-M\cos\theta(-t+\tau-(2\tau/\pi))]}\,dt \quad (17)$$

$$=\omega q e^{(2i\omega\tau/\pi)M\cos\theta}\left\{e^{-i\omega\tau}\left[\pi\delta(\omega(1-M\cos\theta))-\frac{i}{\omega(1-M\cos\theta)}\right]+e^{i\omega\tau}\left[\pi\delta(\omega(1+M\cos\theta))+\frac{i}{\omega(1+M\cos\theta)}\right]+(2\tau/\pi)\int_{-\pi/2}^{\pi/2}e^{(2i\omega\tau/\pi)(\sigma-M\cos\theta\cos\sigma)}\,d\sigma\right\}, \quad (18)$$

where $M=v_0/c$.

The relation $\omega\delta(\omega)=0$ serves to eliminate the delta functions above, and, after an integration by parts that is initiated by writing

$$(2\omega\tau/\pi)e^{(2i\omega\tau/\pi)(\sigma-M\cos\theta\cos\sigma)}=\frac{-i}{1+M\cos\theta\sin\sigma}d_\sigma[e^{(2i\omega\tau/\pi)(\sigma-M\cos\theta\cos\sigma)}], \quad (19)$$

it is found that

$$q\omega\int_{-\infty}^{\infty}e^{i\omega[t-\cos\theta x_s(t)/c]}\,dt=-qie^{(2i\omega\tau/\pi)M\cos\theta}\int_{-\pi/2}^{\pi/2}\frac{M\cos\theta\cos\sigma}{[1+M\cos\theta\sin\sigma]^2}\times e^{(2i\omega\tau/\pi)[\sigma-M\cos\theta\cos\sigma]}\,d\sigma. \quad (20)$$

Accordingly, the joint angular and frequency distribution of the radiated energy is

$$E(\mathbf{n},\omega)=\frac{q^2M^2\cos^2\theta}{2\pi\rho_0 c(4\pi)^2}\times\int_{-\pi/2}^{\pi/2}\frac{\cos\sigma\cos\mu}{[1+M\cos\theta\sin\sigma]^2[1+M\cos\theta\sin\mu]^2}\times\exp[(2i\omega\tau/\pi)(\sigma-\mu)-(2i\omega\tau/\pi)M\cos\theta(\cos\sigma-\cos\mu)]\,d\sigma\,d\mu, \quad (21)$$

and, after integration over all frequencies, the directional spectrum of the radiated energy turns out to be

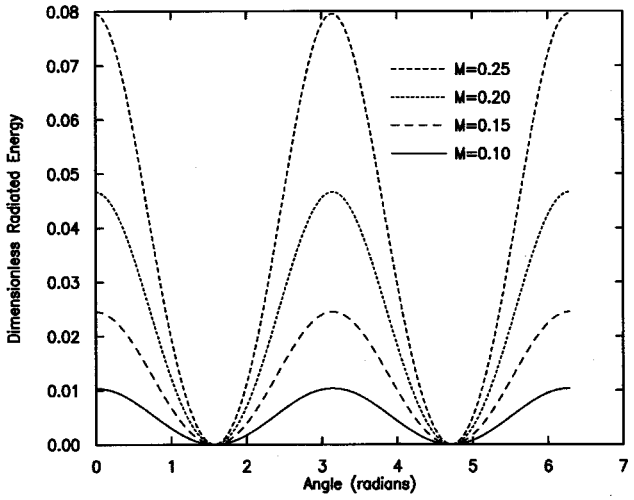


FIG. 2. Angular distribution of the (normalized) energy radiated by the moving monopole, Eq. (22), for Mach numbers in the range $M=0.10$ to $M=0.25$.

$$E(\theta)=\frac{q^2M^2\cos^2\theta}{32\pi\rho_0 c\tau}\int_{-\pi/2}^{\pi/2}\frac{\cos^2\sigma\,d\sigma}{[1+M\cos\theta\sin\sigma]^5}$$

$$=\frac{q^2M^2\cos^2\theta[1+(M^2/4)\cos^2\theta]}{64\rho_0 c\tau[1-M^2cos^2\theta]^{7/2}}, \tag{22}$$

where

$$E=2\pi\int_0^{\pi}E(\theta)\sin\theta\,d\theta$$

$$=\frac{q^2M^2\pi[1-(M^2/4)]}{48\rho_0 c\tau[1-M^2]^{5/2}}$$

$$=\frac{q^2M^2\pi}{48\rho_0 c\tau}[1+9M^2/4+\cdots],\quad M\ll 1, \tag{23}$$

specifies the total radiated energy.

It may be noted that the result in Eq. (22) fully agrees with the outcome of another determination based on the power formula, Eq. (11), *viz.*,

$$E(\theta)=\int_{-\infty}^{\infty}P(\theta,t)\,dt$$

$$=\frac{q^2}{16\pi^2\rho_0 c}\int_{-\infty}^{\infty}\frac{\cos^2\theta(dM/dt)^2}{[1-M(t)\cos\theta]^5}dt, \tag{24}$$

where

$$M(t)=\begin{cases}-M\sin(\pi t/2\tau) & \text{for } |t|<\tau,\\ \text{constant} & \text{otherwise.}\end{cases} \tag{25}$$

In the above, $M=v_0/c$. Therefore, the radiated energy varies inversely with the time span 2τ of accelerated source motion.

A plot of the dimensionless radiated energy $64\rho_0 c\tau E(\theta)/q^2$ *vs* θ from Eq. (22) is shown in Fig. 2 for four different values of M ranging from $M=0.10$ to $M=0.25$. The plot is obviously symmetric with respect to the direction of motion, and its two lobes are seen to grow in size with increasing Mach numbers.

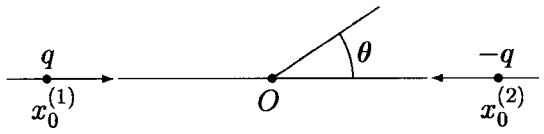


FIG. 3. Rectilinear path of a pair of monopole sound sources on a collision course. Each source reverses direction after collision.

It is clear that for very small values of τ—which would occur for truly instantaneous total reversals in course direction—there would be a very large energy output from the source, as could be verified from Eq. (22).

There are current aircraft that can incredibly undergo a 180-degree change in direction practically on a dime. Interactive videos of such performances were available to the general public, among other places, at the website http://www.usnews.com/whispers during the week of 26 March 2001. The videos show the eye-popping hyperacrobatics of the new Su-37 Russian super Flanker fighter jet. A simplified model for the sound field radiated by such aircraft could be a point source reversing course as we have just described it here. Two such aircraft approaching each other and then quickly completely reversing courses, in a standard maneuver of air-shows, can also be modeled in an elementary fashion, by the present approach, as will be illustrated next.

B. ''Colliding'' monopoles

Consider two monopoles of strengths $\pm q$ on the negative/positive ranges of the x-axis, which initially approach and subsequently recede from the origin (or ''impact'' point) in a symmetrical fashion, as shown in Fig. 3. If their respective coordinates and speeds are denoted by

$$x_s^{(1)}(t)<0,\quad x_s^{(2)}(t)=-x_s^{(1)}(t)>0,$$
$$v_s^{(1)}(t)=-v_s^{(2)}(t), \tag{26}$$

and $x_s^{(1)}$ and $v_s^{(1)}$ are defined as in Eqs. (12) and (13), then the substitution of the composite source function

$$Q(\mathbf{r},t)=q\,\delta(x-x_s^{(1)}(t))-q\,\delta(x+x_s^{(2)}(t)) \tag{27}$$

into the integral of Eq. (7) yields, with a rearrangement in the manner employed earlier,

$$\omega\int_{-\infty}^{\infty}Q(\mathbf{r},t)e^{i\omega(t-\mathbf{n}\cdot\mathbf{r}/c)}\,d\mathbf{r}\,dt$$

$$=-Mqi\cos\theta\int_{-\pi/2}^{\pi/2}\cos\sigma$$
$$\times\left[\frac{e^{(2i\omega\tau/\pi)M\cos\theta+(2i\omega\tau/\pi)(\sigma-M\cos\theta\cos\sigma)}}{(1+M\cos\theta\sin\sigma)^2}\right.$$
$$\left.+\frac{e^{-(2i\omega\tau/\pi)M\cos\theta+(2i\omega\tau/\pi)(\sigma+M\cos\theta\cos\sigma)}}{(1-M\cos\theta\sin\sigma)^2}\right]d\sigma, \tag{28}$$

where the second term is obtained from the first by the substitution $M\to -M$ and $q\to -q$. Substitution into Eq. (7) yields $E(\mathbf{n},\omega)$, and integration over all frequencies gives the directional distribution of the total radiated energy in the form

$$E^*(\theta)=2E(\theta)+\frac{q^2M^2\cos^2\theta}{16\pi\rho_0 c\tau}\Re\int_{-\pi/2}^{\pi/2}\frac{\delta[\sigma-\mu+2M\cos\theta-M\cos\theta(\cos\sigma+\cos\mu)]\cos\sigma\cos\mu}{(1+M\cos\theta\sin\sigma)^2(1-M\cos\theta\sin\mu)^2}d\sigma\,d\mu, \tag{29}$$

where $\Re$ denotes real part, $E(\theta)$ is the angular pattern in Eq. (22) characteristic of each source in isolation, which is shown in Fig. 2, and the remainder of Eq. (29) constitutes an interference term arising from the interaction between the pair of sources. To reduce Eq. (29) requires the analysis of the implicit equation,

$$\sigma-\mu+2M\cos\theta-M\cos\theta(\cos\sigma+\cos\mu)=0, \tag{30}$$

that defines σ, say, as a function of μ. We now let

$$\epsilon=M\cos\theta,\quad \sigma=\mu-\epsilon f(\sigma,\mu),$$
$$f(\sigma,\mu)=2-\cos\sigma-\cos\mu, \tag{31}$$

and then Lagrange's expansion formula with $\epsilon<1$,

$$\sigma=\mu+\sum_{n=1}^{\infty}(-1)^n\frac{\epsilon^n}{n!}\frac{d^{n-1}}{d\sigma^{n-1}}[f(\sigma,\mu)]^n\Big|_{\sigma=\mu}, \tag{32}$$

implies the development

$$\sigma=\mu-2\epsilon(1-\cos\mu)+2\epsilon^2\sin\mu(1-\cos\mu)+\cdots. \tag{33}$$

Similarly,

$$\sin\sigma=\sin\mu+\sum_{n=1}^{\infty}(-1)^n\frac{\epsilon^n}{n!}\frac{d^{n-1}}{d\sigma^{n-1}}[\cos\sigma(f(\sigma,\mu))^n]\Big|_{\sigma=\mu}$$
$$=\sin\mu-2\epsilon\cos\mu(1-\cos\mu)$$
$$+2\epsilon^2\sin\mu(-1+3\cos\mu+2\cos^2\mu)+\cdots, \tag{34}$$

and an analogous expression for $\cos\sigma$ also holds. On the basis of these and additional developments in powers of ϵ or M, it is found that

$$E_T=2\pi\int_0^{\pi}E^*(\theta)\sin\theta\,d\theta$$
$$=\frac{q^2M^2\pi}{12\rho_0 c\tau}+\frac{q^2M^4}{\rho_0 c\tau}\left(\frac{8}{15}-\frac{\pi}{16}\right)+\cdots, \tag{35}$$

which for $M\ll 1$ describes the total radiated energy for interacting monopole sources of opposite strengths. If the monopole pair has equal strengths, then a lesser energy level is found, *viz.*,

$$E_T=\frac{q^2M^4}{\rho_0 c\tau}(\pi/4-8/15)+\cdots(M\ll 1). \tag{36}$$

The leading term in Eq. (35) can be derived from a dipole source whose moment $p(t)=2qx_s(t)$ is suggested by a pair of oppositely signed sources with a variable separation $2x_s(t)$, as illustrated in Fig. 4. Thus, applying the representation Eq. (9) when $f=0$ and $g(t)=p(t)$, and replacing the Doppler factor $1-M(t)\cos\theta$ by unity, an approximation valid for slow motions, one obtains

$$P(\mathbf{n},t)\approx\frac{\cos^2\theta}{\rho_0 c^3(4\pi)^2}\left(\frac{d^2p}{dt^2}\right)^2$$
$$=\frac{q^2M^2}{16\rho_0 c\tau^2}\cos^2\theta\cos^2(\pi t/2\tau). \tag{37}$$

Consequently the energy output is

$$E=2\pi\int_0^{\pi}\sin\theta\,d\theta\int_{-\tau}^{\tau}P(\mathbf{n},t)\,dt=\frac{q^2M^2\pi}{12\rho_0 c\tau}(M\ll 1), \tag{38}$$

in full agreement with Eq. (35).

C. Further analysis of equal/opposite strength colliding monopoles

To appreciate why the source pairs of equal/opposite strengths generate different orders of acoustical energy output, consider the retarded potential representation

$$\phi(\mathbf{r},t)=\frac{1}{4\pi\rho_0}\int\frac{Q(\mathbf{r}',t-|\mathbf{r}-\mathbf{r}'|/c)}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}', \tag{39}$$

and the circumstances which permit a multipole expansion of the potential. Let the spatial extent of a source domain be small enough for Q to change only slightly in the time required for a sound wave to traverse it. In other words, the largest frequencies $\nu=\omega/2\pi$ present in the Fourier decomposition of the source density function Q must be such that $\nu L/c\ll 1$, where L is a characteristic dimension of the source domain. Equivalently, the corresponding wavelength $\lambda=c/\nu$ must be such that $\lambda\gg L$. Then, the difference in retarded times, $t-|\mathbf{r}-\mathbf{r}'|/c$, between various parts of the source distribution is small, and to a first approximation, all retarded times can be identified with that of a fixed point, say the origin of coordinates. More precisely, the difference between $t-|\mathbf{r}-\mathbf{r}'|/c$ and $t-r/c$ can be taken into account by expansion of the source function, thus

$$Q(\mathbf{r},t-|\mathbf{r}-\mathbf{r}'|/c)=Q(\mathbf{r}',t-r/c)+(r-|\mathbf{r}-\mathbf{r}'|)$$
$$\times(1/c)\frac{\partial}{\partial t}Q(\mathbf{r}',t-r/c)+\cdots. \tag{40}$$

A more direct procedure is possible if we are primarily concerned with the far field (i.e., the field at large distances from the center of the source region). Now, $|\mathbf{r}'|$ is regarded as small compared to r whenever it occurs in the retarded potential expressions, though now a separate approximation

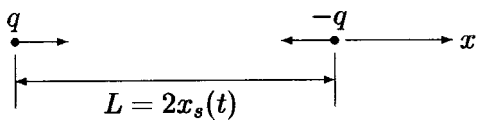


FIG. 4. Rectilinear path of a colliding pair of oppositely signed sources $\pm q$ showing their dipole moment $\boldsymbol{\mu}(t)=qL$.

for the wavelength applies. Hence, we introduce a Taylor series expansion

$$\frac{Q(\mathbf{r}',t-|\mathbf{r}-\mathbf{r}'|/c)}{|\mathbf{r}-\mathbf{r}'|}=[1-\mathbf{r}'\cdot\nabla+(1/2)(\mathbf{r}'\cdot\nabla)^2-\cdots]\times\frac{Q(\mathbf{r}',t-r/c)}{r} \tag{41}$$

into Eq. (39) and obtain

$$\phi(\mathbf{r},t)=\frac{q(t-r/c)}{4\pi\rho_0 r}-\frac{1}{4\pi\rho_0}\nabla\cdot\frac{\boldsymbol{\mu}(t-r/c)}{r}+\frac{1}{8\pi\rho_0}\nabla\cdot\nabla:\frac{\boldsymbol{\Gamma}(t-r/c)}{r}+\cdots, \tag{42}$$

where the total source strength, the dipole moment vector, and the quadrupole moment dyadic are, respectively,

$$q(t)=\int Q(\mathbf{r},t)\,d\mathbf{r},\quad \boldsymbol{\mu}(t)=\int \mathbf{r}Q(\mathbf{r},t)\,d\mathbf{r},\quad \boldsymbol{\Gamma}(t)=\int \mathbf{rr}Q(\mathbf{r},t)\,d\mathbf{r}, \tag{43}$$

where the dyadic notation $\mathbf{A}:\mathbf{B}$ is employed to signify an inner product, *viz.*,

$$\mathbf{A}:\mathbf{B}=\sum_i\sum_j A_{ij}B_{ji}. \tag{44}$$

For a pair of colliding monopoles with the characterization

$$Q(\mathbf{r},t)=q\,\delta[\mathbf{r}-\mathbf{r}_s(t)]\pm q\,\delta[\mathbf{r}+\mathbf{r}_s(t)], \tag{45}$$

the quantities in Eq. (43) are specified by

$$q(t)=\begin{cases}2q, & \oplus,\\ 0, & \ominus,\end{cases} \tag{46}$$

$$\boldsymbol{\mu}(t)=\begin{cases}0, & \oplus,\\ 2q\mathbf{r}_s(t), & \ominus,\end{cases} \tag{47}$$

$$\boldsymbol{\Gamma}(t)=\begin{cases}2q\mathbf{r}_s(t)\mathbf{r}_s(t), & \oplus,\\ 0, & \ominus,\end{cases} \tag{48}$$

For an asymmetrical (i.e., oppositely signed) source pair,

$$\phi(\mathbf{r},t)\approx\frac{-2q}{4\pi\rho_0}\frac{\partial}{\partial x}\left[\frac{x_s(t-r/c)}{r}\right]\approx\frac{-qM}{2\pi\rho_0 r}\sin\frac{\pi(t-r/c)}{2\tau}\cos\theta, \tag{49}$$

and the radial far-field energy flux is

$$S_r=-\rho_0\frac{\partial\phi}{\partial t}\frac{\partial\phi}{\partial r}\approx\frac{q^2M^2}{16\rho_0 c\tau^2}\frac{\cos^2\theta}{r^2}\cos^2\frac{\pi(t-r/c)}{2\tau}. \tag{50}$$

Hence, the total energy radiated during the time span 2τ is

$$E=\frac{q^2M^2}{16\rho_0 c\tau^2}2\pi\int_0^\pi\cos^2\theta\sin\theta\,d\theta\int_{-\tau}^{\tau}\cos\frac{\pi(t-r/c)}{2\tau}dt=\frac{q^2M^2\pi}{12\rho_0 c\tau}, \tag{51}$$

which agrees with the leading term of Eq. (35) for opposite strength monopoles. For a pair of equal strength sources, the dipole moment vanishes and the unsteady source representation is quadrupolar, *viz.*,

$$\phi(\mathbf{r},t)\approx\frac{q}{4\pi\rho_0 r}\frac{\partial^2}{\partial x^2}[x_s^{(2)}(t-r/c)]. \tag{52}$$

Proceeding like before, this far-field estimate yields a radial energy flux

$$S_r=\frac{q^2M^4\cos^4\theta}{4\rho_0 c\tau^2 r^2}\left[\sin\frac{\pi(t-r/c)}{\tau}-\frac{1}{2}\sin\frac{\pi(t-r/c)}{2\tau}\right]^2, \tag{53}$$

and the amount of energy radiated during a time-span 2τ is

$$E=\frac{q^2M^4\pi}{4\rho_0 c\tau}, \tag{54}$$

which shows a small discrepancy with the previous result in Eq. (36). However, a direct approach using the quadrupole strength $\Gamma_{xx}=2qx_s^2(t)$ jointly with the representation

$$P(\mathbf{n},t)=\frac{\cos^4\theta}{16\pi^2\rho_0 c^5}\left[\frac{d^3}{dt^3}\frac{1}{2}\Gamma_{xx}(t)\right]^2 \tag{55}$$

and the $x_s(t)$ in Eq. (12) gives

$$P(\mathbf{n},t)=\frac{\cos^4\theta}{16\pi^2\rho_0 c^5}\left[\frac{v_0\pi q}{\tau}\left(2\sin\frac{\pi t}{\tau}-\sin\frac{\pi t}{2\tau}\right)\right]^2, \tag{56}$$

and its integral over the time-span 2τ is

$$\int_{-\tau}^{\tau}P(\mathbf{n},t)\,dt=\frac{q^2M^4\cos^4\theta}{16\rho_0 c\tau}\left(5-\frac{32}{3\pi}\right). \tag{57}$$

Then the energy radiated in this symmetrical case is

$$E=2\pi\int_0^\pi\sin\theta\,d\theta\int_{-\tau}^{\tau}P(\mathbf{n},t)\,dt=\frac{q^2M^4\pi}{\rho_0 c\tau}\left(\frac{1}{4}-\frac{8}{15\pi}\right), \tag{58}$$

which fully agrees with Eq. (36) for a pair of monopoles of equal strength in a collision course.

III. CLOSURE

This article has dealt with some cases of radiation from point sources which undergo acceleration only during a finite time interval, and it presents a manner of analyzing paired source combinations at small Mach numbers. In all cases considered, the total energy radiated and some of its angular or spectral distributions are determined and graphically displayed.

ACKNOWLEDGMENTS

Author GCG gratefully acknowledges partial support from the In-house Laboratory Independent Research (ILIR) program of the Naval Surface Warfare Center, Carderock Division (NSWCCD), and from the Army Research Laboratory.

[1] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).
[2] A. P. Dowling and J. E. Ffowcs Williams, *Sound and Sources of Sound* (Horwood, Wiley, New York, 1983).
[3] H. Hönl, ''On the sound field of a point sound source in uniform translatory motion,'' Technical Report 1362, National Advisory Committee for Aeronautics, Washington, DC, 1954 [translation from Ann. Phys. (Leipzig) **43(5)**, 437–464 (1943)].
[4] M. J. Lighthill, ''On sound generated aerodynamically,'' Proc. R. Soc. London, Ser. A **260**, 147–182 (1962).
[5] H. Levine and S. M. Candel, ''Acoustic power output from moving point singularities,'' J. Acoust. Soc. Am. **81**, 1695–1702 (1987).
[6] H. Levine, ''The output of acoustic sources,'' J. Acoust. Soc. Am. **67**, 1935–1946 (1980).
[7] G. C. Gaunaurd, ''Contributions to the Theory of Acoustic Radiation by Moving Multipole Sources,'' Ph.D. thesis, Catholic University of America, Washington, DC, 1971. (Also available with more detail as Naval Ordnance Laboratory Report NOLTR-72-75, 1972, and Technical Note NOLTN-9578, May 1972.)
[8] J. Schwinger, ''On the classical radiation of accelerated electrons,'' Phys. Rev. **75**, 1912–1925 (1949).
[9] G. C. Gaunaurd and T. J. Eisler, ''Classical electrodynamics and acoustics: Sound radiation by moving multipoles I,'' J. Vibr. Acoust., **119**, 271–282 (1997).
[10] G. C. Gaunaurd and T. J. Eisler, ''Classical electrodynamics and acoustics: Sound radiation by moving multipoles II,'' J. Vibr. Acoust., **121**, 126–130 (1999).

Optimal focusing by spatio-temporal inverse filter. I. Basic principles

M. Tanter, J.-F. Aubry, J. Gerber, J.-L. Thomas, and M. Fink
Laboratoire Ondes et Acoustique, ESPCI, Université Paris VII, U.R.A. C.N.R.S. 1503, 10 rue Vauquelin, 75005 Paris, France

(Received 16 January 2001; accepted for publication 9 April 2001)

A focusing technique based on the inversion of the propagation operator relating an array of transducers to a set of control points inside a medium was proposed in previous work [Tanter *et al.*, J. Acoust. Soc. Am. **108**, 223–234 (2000)] and is extended here to the time domain. As the inversion of the propagation operator is achieved both in space and time, this technique allows calculation of the set of temporal signals to be emitted by each element of the array in order to optimally focus on a chosen control point. This broadband inversion process takes advantage of the singular-value decomposition of the propagation operator in the Fourier domain. The physical meaning of this decomposition is explained in a homogeneous medium. In particular, a definition of the number of degrees of freedom necessary to define the acoustic field generated by an array of limited aperture in a focal plane of limited extent is given. This number corresponds to the number of independent signals that can be created in the focal area both in space and time. In this paper, this broadband inverse-focusing technique is compared in homogeneous media with the classical focusing achieved by simple geometrical considerations but also with time-reversal focusing. It is shown that, even in a simple medium, slight differences appear between these three focusing strategies. In the companion paper [Aubry *et al.*, J. Acoust. Soc. Am. **110**, 48–58 (2001)] the three focusing techniques are compared in heterogeneous, absorbing, or complex media where classical focusing is strongly degraded. The strong improvement achieved by the spatio-temporal inverse-filter technique emphasizes the great potential of multiple-channel systems having the ability to apply completely different signal waveforms on each transducer of the array. The application of this focusing technique could be of great interest in various ultrasonic fields such as medical imaging, nondestructive testing, and underwater acoustics.
[DOI: 10.1121/1.1377051]

PACS numbers: 43.20.−f, 43.20.Bi, 43.20.El, 43.20.Fn [ANN]

I. INTRODUCTION

A wide range of ultrasonic applications, such as medical imaging or therapy, but also nondestructive testing and underwater acoustics, requires the focusing of an ultrasonic beam deep inside the medium under investigation. Focusing is classically obtained by the assumption of constant sound speed in the medium, and hence the application of a cylindrical time-delay law on the elements of a multiple-transducer array in order to compensate for the time-of-flight differences between elements. This classical focusing is unfortunately strongly degraded for a heterogeneous medium. However, most ultrasonic medical systems do not take into account the inhomogeneities of human tissue. The same assumption is often made in nondestructive testing and underwater acoustics.

In medical applications, the degradations induced by tissue heterogeneities on the ultrasonic focused beam, known as wavefront aberrations, have been widely studied during the past three decades.[1,2] However, most investigations modeled the aberrations as a very thin layer located close to the array which only introduced varying time delays along the array aperture. Thus, scientists working in the medical ultrasound domain restricted their means of compensation to time shift in the transmitting and receiving aperture. By studying wavefront distortion due to breast tissue, Zhu and Steinberg[3,4] have shown that this is not a complete solution of the aberration correction problem. The aberrations cannot be described only by a phase distortion; rather, an amplitude distortion also has to be taken into account.

The inability of time-shift compensation in the receiving aperture to correct perfectly for the effect of a phase and amplitude screen that is not located close to the receiving aperture was shown experimentally and explained by Wu *et al.*[5,6] In 1994, Liu and Waag[7] proposed to model the aberrating medium as a phase screen located at some distance from the receiving aperture. In that case, amplitude and phase distortions can be removed by numerically backpropagating the wavefront to the aberrating layer before applying the time-shift compensation.

However, in many configurations, the aberrations cannot be modeled as a thin phase and amplitude screen located at a given distance from the array. Heterogeneities are dispersed in the whole insonified area. For a pulsed echo, each spectral component of the signal emitted by each element of the array is subject to different phase and amplitude distortions: therefore, the optimal focus can only be achieved by applying a different waveform on each transducer of the array. It straightforwardly emphasizes the utility of array systems made of independent and completely programmable emitting transducers. Time reversal[8,9] was shown to represent an

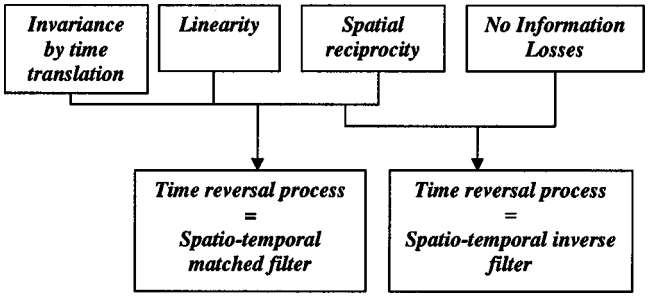


FIG. 1. Properties of the time-reversal process.

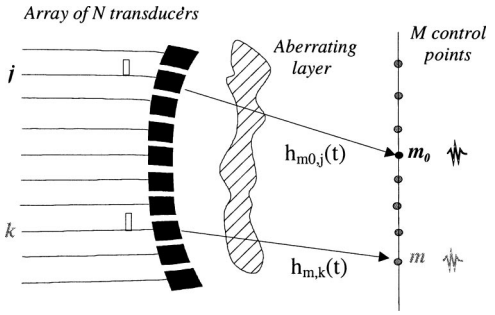


FIG. 2. The propagation operator $h(\mathbf{r},\mathbf{r}',t)$ relating the emitting plane to the reception control plane is discretized by a set of $[M\times J]$ temporal signals, $h(m,j,t)$, called the source–transducer impulse responses.

original and elegant way to compensate for distortions due to such varied heterogeneous media. Indeed, if linearity and spatial reciprocity assumptions are valid in the medium, the time-reversal focusing process corresponds to a spatial and temporal matched filter[10] of the propagation: In other words, the time-reversal process maximizes the output amplitude received at a given location and a given time (the signal amplitude received at focus at a given time) for a given input energy (corresponding to the signals applied on the emitting array). Note that time-reversal processing only optimizes the pressure amplitude received at focus but does not impose any constraints on the field around the focus. However, this simple property of linearity combined with spatial reciprocity explains the strong robustness and efficiency of the time-reversal process in such varied domains of application as medical ultrasound,[11] nondestructive testing,[12] and underwater acoustics.[13] Moreover, as long as information losses do not appear, the time-reversal invariance of the wave equation holds and thus the spatial and temporal matched filter corresponds to an inverse filter of the propagation. Therefore, time reversal applied on a closed surface surrounding a lossless medium ensures the optimal focusing.

When information losses appear during the time-reversal process, as was shown in Ref. 10, time reversal can no longer be considered as an inverse filter of the propagation (Fig. 1). If we want to construct such an inverse filter, we have to acquire the whole propagation operator relating the emitting surface (the transducer array) to a set of control points surrounding the desired focus (in our case a control plane that will be referred to as the focal plane). In fact, this propagation operator acts as a spatio-temporal filter on the emitted signals and its inversion allows us to recover the optimal focusing on a desired point of the control plane. Contrary to the time-reversal process, this focusing technique optimizes not only the pressure wave field received at the focus but also imposes constraints on the signal received on each point of the control plane (here, the focal plane). The problem of imposing constraints on control points located inside the medium can be related to the active noise control approach encountered in room acoustics,[14,15] but also to the radiation pattern control developed for ultrasonic hyperthermia by Seip *et al.*[25] and by Ebbini and Cain.[26]

The restriction of this inverse filter process to a monochromatic approach was described with a matrix formalism in Ref. 10. The aim of this paper and its companion is to extend this formalism to the time domain. Section II introduces the spatio-temporal propagation operator. In Sec. III, the physical meaning of the singular-value decomposition of the propagation operator is illustrated and explained for a homogeneous medium. In Sec. IV, this formalism and its implications are extended to the time domain. In particular, an expression for the number of independent waveforms that can be created during a given time T in the focal plane is given.

The theoretical explanations are illustrated by simulation and experimental results conducted in a homogeneous medium (water). For our purposes, the classical and time-reversal focusing techniques are compared in a homogeneous medium to the spatio-temporal inverse-filter technique. Surprisingly, we observe that even in a such simple case the three techniques differ slightly. Indeed, due to their typical technical characteristics (finite size and impulse response), the transducers induce limited spatial and frequency bandwidth both on transmit and receive signals. We will see that these limited bandwidths are responsible for the differences between the three focusing techniques in homogeneous media as they can be considered as partial losses of information.

As the aim is, of course, optimal focusing in more complex media such as those with heterogeneities in sound speed, density, or absorption, but also reverberating, chaotic, or multiply scattering media, a companion paper, hereafter referred to as Paper II, will demonstrate some potential applications of this new focusing technique achieved by spatio-temporal inverse filter and its impressive efficiency in comparison with classical or time reversal techniques.

II. THE PROPAGATION OPERATOR $\{h_{mj}\}(t)$

We define the linear operator relating the elements of the transducer array to a set of control points embedded in the medium. The signals applied to the transducer elements are the input of the linear operator, and the output consists of the signals measured on the control points.

For simplicity, we restrict ourselves to a one-dimensional array of transducers and a spatial sampling of the field, the set of control points, along a line parallel to the array aperture, Fig. 1. This choice is strictly for convenience and the process can be applied to 2D arrays and a set of control points distributed as desired in the medium. We define an impulse response, $h_{mj}(t)$, for each pair (m,j) comprising a control point and a transducer element such that $h_{mj}(t)$ is the signal received on the mth control point after a tem-

poral delta function is applied on the jth transducer of the array, Fig. 2. This response includes all the propagation effects through the medium under investigation as well as the acousto-electric responses of the array element and the receiver. As the transformations are assumed to be linear and invariant under a time shift, the $M\times J$ temporal functions $h_{mj}(t)$ for $1\leqslant m\leqslant M$ and $1\leqslant j\leqslant J$ can describe any transmit–receive operation for the same arrangement. This set of $M\times J$ temporal functions characterizes the *propagation operator*.

Indeed, let $e_j(t)$, $1\leqslant j\leqslant J$, be the J input signals on the transducer array. The output signals $f_m(t)$, $1\leqslant m\leqslant M$, received in the control domain are given by

$$f_m(t)=\sum_{j=1}^{J} h_{mj}(t)\overset{*}{_t} e_j(t), \quad 1\leqslant m\leqslant M, \tag{1}$$

where $*$ is a temporal convolution operator.

A temporal Fourier transform leads to the following relation:

$$F_m(\omega)=\sum_{j=1}^{J} \mathbf{H}_{mj}(\omega)E_j(\omega), \quad 1\leqslant m\leqslant M. \tag{2}$$

Equation (2) can be written in matrix form

$$F(\omega)=\mathbf{H}(\omega)E(\omega), \tag{3}$$

where $E(\omega)=(E_j(\omega))_{1\leqslant j\leqslant J}$ is the column vector of the Fourier transform of the transmitted signals and $F(\omega)=(F_m(\omega))_{1\leqslant m\leqslant M}$ is the column vector of the Fourier transform of the received signals. The transfer matrix $\mathbf{H}(\omega)=\{H_{mj}(\omega)\}_{1\leqslant m\leqslant M,1\leqslant j\leqslant J}$ describes the propagation in the medium from the array to the set of control points for a chosen frequency component and thus is called the *monochromatic propagation matrix.*

III. THE MONOCHROMATIC APPROACH: THE SPATIAL INVERSE FILTER

A. The monochromatic inverse propagation matrix $\mathbf{H^{-1}}$

In previous work,[10] we explained how to deal with an inverse diffraction experiment in the monochromatic case. In that case, the aim was to determine the desired field distribution E on a boundary surface (the transducer array) that would give rise after propagation to the field distribution F on another surface (the focal plane or control plane). As we wish to focus in the focal plane, we typically take as the target pattern F either a spatial delta function or the optimal focusing obtained in homogeneous medium. To solve this problem, we first acquire the direct propagation matrix $\mathbf{H}$ between the transducer array and the focal plane and then invert it. Thus, E and F are related by

$$F=\mathbf{H}E\Leftrightarrow E=\mathbf{H}^{-1}F. \tag{4}$$

As is the case in all inverse problems, the inversion of the propagation matrix $\mathbf{H}$ is *ill-conditioned*.[16] Numerically, the solution to the inversion does not depend continuously on the data, but rather, small errors in these data produce very large errors in the reconstructed results. Thus, a regularization of the problem is introduced by computing the singular-value decomposition of the matrix $\mathbf{H}$ before inversion

$$\mathbf{H}=\mathbf{U}\mathbf{D}\tilde{\mathbf{V}}. \tag{5}$$

Here, D is a diagonal matrix of singular values (arranged in decreasing order) and U and V are unitary matrices. The matrix inversion is only applied to the main (i.e., physically relevant) singular vectors of the singular-value decomposition of $\mathbf{H}$. It gives rise to a noise-filtered approximation $\hat{\mathbf{H}}^{-1}$ of the inverse matrix $\mathbf{H}^{-1}$

$$\hat{\mathbf{H}}=\mathbf{V}\hat{\mathbf{D}}^{-1}\tilde{\mathbf{U}}=\mathbf{V}\begin{bmatrix} \lambda_1^{-1} & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & 0 & & & \vdots \\ \vdots & 0 & \lambda_N^{-1} & & & \vdots \\ \vdots & & & 0 & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{bmatrix}\tilde{\mathbf{U}}, \tag{6}$$

where N is the number of physically relevant singular values. In 1956, Gabor showed that this number N is associated with the number of eigensolutions of the homogeneous Helmholtz equation.[17]

As soon as $\mathbf{H}$ has been experimentally acquired, $\hat{\mathbf{H}}^{-1}$ can be easily calculated. Then, we need only to choose as an objective the field distribution F_0 we would like to generate on the control points. Given F_0, the inverse propagation matrix $\hat{\mathbf{H}}^{-1}$ provides for a straightforward calculation of the monochromatic vector E to be emitted on the array. This emitted field will then give rise, after propagation to the control plane, to a directivity pattern F_{pr}

$$F_{pr}=\mathbf{H}\hat{\mathbf{H}}^{-1}F_0=\mathbf{U}\begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & & & \vdots \\ \vdots & 0 & 1 & & & \vdots \\ \vdots & & & 0 & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{bmatrix}\tilde{\mathbf{U}}F. \tag{7}$$

This field F_{pr} is the best fit of F in terms of the least-mean-squared optimization.[18] Equation (7) shows that the focusing obtained by propagating the emission vector calculated by monochromatic inverse filter process can be easily predicted: it corresponds to the projection of the objective F on the nonzero singular vectors of the propagation matrix. Thus, as soon as we acquire the propagation matrix we can numerically predict the optimal focusing F_{pr} that will be obtained by the system in the control plane.

B. The singular value decomposition of the monochromatic propagation matrix

As mentioned above, the singular-value decomposition allows us to separate the rank N of the monochromatic propagation matrix from its physically irrelevant kernel. How can we predict the number N of degrees of freedom of $\mathbf{H}$?

This number N of relevant singular values has been predicted and explained in previous work:[10] the limited aperture

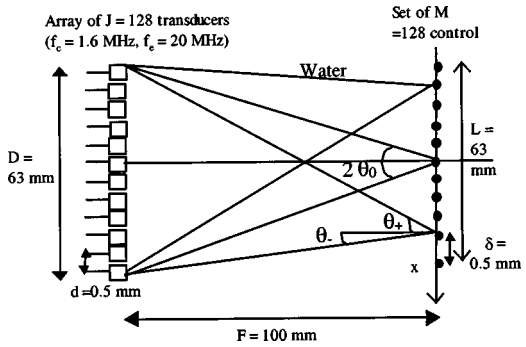


FIG. 3. Simulation and experimental setup.

of the emitting array limits the maximal information content that can be conveyed by the field. Indeed, the limited aperture of the array introduces a cutoff in the angular frequencies of the wave field received at each location in the medium. In our configuration, the number N of spatial degrees of freedom necessary to define the monochromatic field received in the control plane is given by:

$$N=\frac{2L}{\lambda}\sin\left(\tan^{-1}\frac{D}{2F}\right), \tag{8}$$

where D is the array aperture, F is the distance between the array and the control plane, and L is the aperture synthesized by the set of control points, Fig. 3. This definition of N is valid until $L \gg D$. Thus, this definition remains valid in the near field of the transducers. Finally, in the Fresnel assumption, this formula can be simplified in

$$N=\frac{LD}{\lambda F}. \tag{9}$$

In this final case, we easily recognize the number of independent focal spots (lateral resolution $\lambda F/D$) that the array is able to generate in the aperture L of the image plane.

As an example, assume that we use a 127-transducer array working at a central frequency of 1.6 MHz ($\lambda = 0.93$ mm). The signal sampling frequency is chosen to be equal to 20 MHz. The array aperture is equal to 63 mm. The array pitch is 0.5 mm, nearly equal to $\lambda/2$. A set of 127 control points is located along a line parallel to the array aperture in the plane $z=F=100$ mm. The spatial step between two control points is equal to 0.5 mm.

We define the acousto-electric impulse response $r(t)$ of the transducers; see Fig. 4. This signal corresponds to the ultrasonic signal generated by a transducer when a delta function is applied on this element. We also introduce a transducer directivity $\mathcal{D}(\theta)=\mathrm{sinc}(d\theta/\lambda)$, taking into account the finite size d of each element (for sake of simplicity, we assume the directivity to be independent of frequency). The propagation operator $h_{mj}(t)$ is then defined by the transducer–receiver impulse responses

$$\begin{aligned}&\forall\{m,j\}\in[1,M]\times[1,J],\\ &h_{mj}(t)=\mathcal{D}\left(\cos^{-1}\frac{F}{d_{mj}}\right)\frac{r(t-d_{mj}/c)}{d_{mj}},\end{aligned} \tag{10}$$

where d_{mj} is the distance between the mth control point and the jth transducer. In the Fourier domain at a chosen frequency $\omega=\omega_0$, the propagation operator is described by the monochromatic propagation matrix $\mathbf{H}(\omega_0)$

$$\begin{aligned}&\forall\{m,j\}\in[1,M]\times[1,J],\\ &\mathbf{H}_{mj}(\omega_0)=\mathcal{D}\left(\cos^{-1}\frac{F}{d_{mj}}\right)\frac{\mathcal{R}(\omega_0)e^{-j\omega_0\cdot d_{mj}/c}}{d_{mj}},\end{aligned} \tag{11}$$

where $\mathcal{R}(\omega_0)$ is the Fourier component of $r(t)$ at frequency $\omega_0/2\pi$.

This matrix has been computed at $\omega_0/2\pi=1.6$ MHz. The singular-value decomposition of this matrix leads to the singular-value distribution shown in Fig. 5 (vertical dashed line.) We can see that the theoretical number N predicted by Eq. (8) provides a good estimate ($N=43$) of the boundary between the physically relevant singular values and the noise.

What is the physical meaning of these main singular vectors of the monochromatic propagation matrix? First of all, we observe that each singular vector V_i of the propagation matrix satisfies

$${}^{t}\mathbf{H}^{*}\mathbf{H}V_i=(\mathbf{V}\mathbf{D}^{t}\mathbf{U}^{*}\mathbf{U}\mathbf{D}^{t}\mathbf{V}^{*})V_i=(\mathbf{V}\mathbf{D}^{2}\,{}^{t}\mathbf{V}^{*})V_i=\lambda_1^2V_i. \tag{12}$$

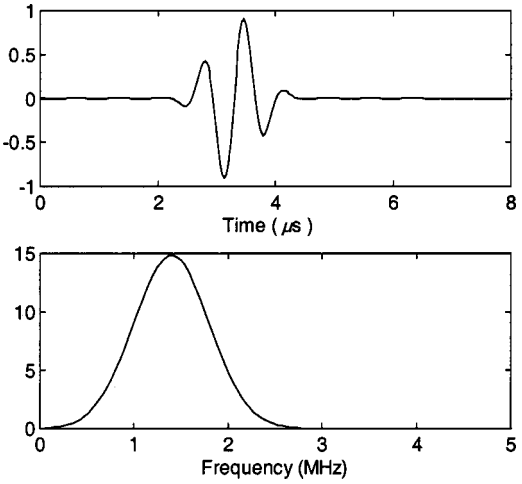


FIG. 4. Computed acousto-electric impulse response $g(t)$ of the tranducers and the absolute value of its Fourier transform.

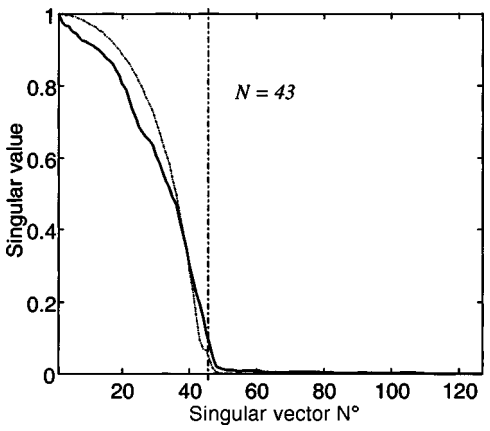


FIG. 5. Computed singular-value distribution for the described configuration dashed line. Comparison with the experimental distribution in the same configuration solid line. Equation (8) predicts $N=43$ freedom degrees.

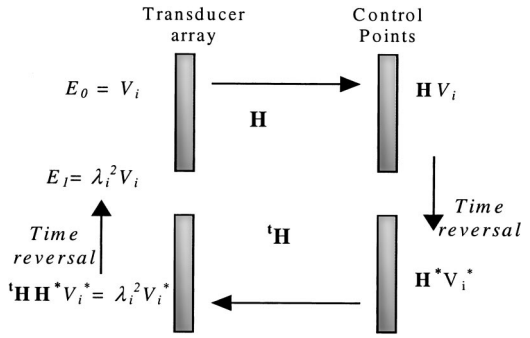


FIG. 6. Physical meaning of **H** singular vectors during a thought experiment: an iteration of the time-reversal operation between the array of transducers and the set of control points. If we emit the vector $E_0 = V_i$ on the array at the beginning of the first iteration, then we emit $E_1 = \lambda_i^2 V_i$ at the beginning of the second iteration: V_i is an invariant of such a process.

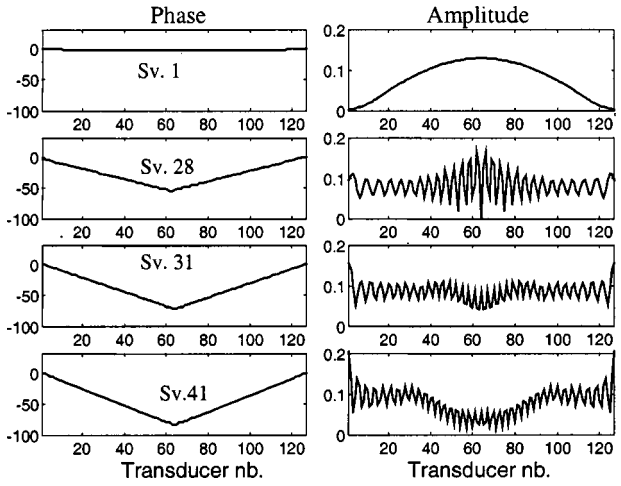


FIG. 7. Amplitude and phase representation of four main singular vectors of the propagation matrix $\mathbf{H}(r, \omega = \omega_0)$ at a chosen frequency $f_0 = 1.6$ MHz.

Thus, the singular vectors of the propagation matrix correspond to the eigenvectors of $^tH^*H$ known in the literature as the time-reversal operator.[19] This matrix has been introduced by Prada in the D.O.R.T. method for target detection in nondestructive testing but also underwater acoustics. In our case we can give another physical meaning to this $^t\mathbf{H}^*\mathbf{H}$ matrix by introducing the following thought experiment: emission of a singular vector V_i on the transducer array. After propagation in the medium, the wave-field distribution received in the control plane is $\mathbf{H}V_i$. This wave field is then time-reversed (which simply consists in a phase conjugation operation for this monochromatic case). Then, we obtain after propagation in the medium the vector $^t\mathbf{H}\mathbf{H}^*V_i^* = \lambda_i^2 V_i^*$ on the array of transducers. Here, we only assume the spatial reciprocity of the medium (in other words, the propagation from the control plane to the array is described by $^t\mathbf{H}$). This vector received on the array is finally time-reversed to achieve a first iteration of a time-reversal experiment between two limited apertures. Thus, we finally obtain the vector $\lambda_i^2 V_i$ ready to be emitted for a second iteration. In conclusion, a singular vector of **H** remains invariant (to within the factor λ_i^2) during a double time-reversal experiment between the two apertures (Fig. 6).

Moreover, we are here in a particular configuration: the array and control points apertures are parallel, have the same symmetry axis, and the same width (Fig. 3). In this particular case, the singular vectors can be to some extent related to the plane-wave decomposition. Figure 7 represents the amplitude and phase of four main singular vectors. As one can notice, the first singular vector corresponds to the emission of a plane wave in the direction perpendicular to the array aperture. The next ones correspond to the emission of two symmetric plane waves on the array aperture (Fig. 7, singular values numbers 28, 31, and 41). The angular direction of these plane waves progressively increases as the singular-value weight decreases.

Due to the spatial symmetries of the configuration, **H** satisfies the relation $\mathbf{H} = {}^t\mathbf{H}$ and thus, $\mathbf{U} = \mathbf{V}^*$. Each singular vector V_i then satisfies

$$\mathbf{H}V_i = \lambda_i U_i = \lambda_i V_i^* . \tag{13}$$

Thus, each singular vector is transformed into its phase conjugate after propagation from the array to the control plane. We can confirm this with the phase representation of the singular vectors. For each singular vector, the two semiplane waves cross each other during propagation and finally recover the phase conjugate form of the initially emitted vector (Fig. 8). Analogous behavior was recently observed by Lin *et al.* in the time-domain eigenfunctions of a scattering operator in the case of 2D arrays.[20]

Furthermore, we can analyze the spatial frequency distribution of these singular vectors. Figure 9 represents the spatial Fourier transform of four main singular vectors. As previously stated, the first singular vector clearly corresponds to the zero-angle plane-wave emission. The subsequent singular vectors correspond to progressively increasing spatial frequencies.

Finally, the entire set of **H** singular vectors can be represented in the spatial Fourier domain (Fig. 10). We clearly recognize the increasing spatial frequencies of the 43 main singular vectors. The following singular vectors correspond to very weak singular values and so can be considered as representing the kernel of the propagation matrix.

This matrix $\mathbf{H}(\omega = \omega_0)$ has been experimentally acquired in the same configuration (Fig. 3). As previously no-

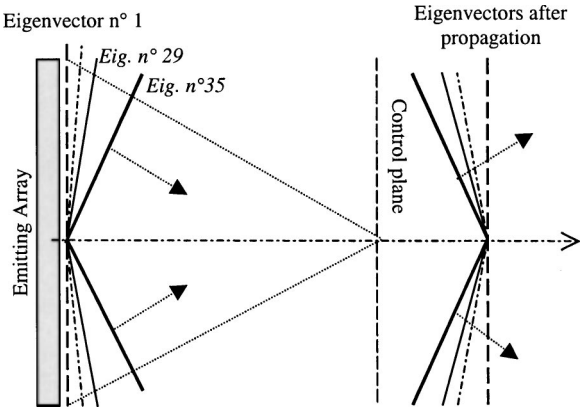


FIG. 8. In the configuration corresponding to Fig. 4, the main singular vectors represented in Fig. 7 correspond to ''semiplane'' waves that are changed into their conjuguate after propagation from the array to the control plane.

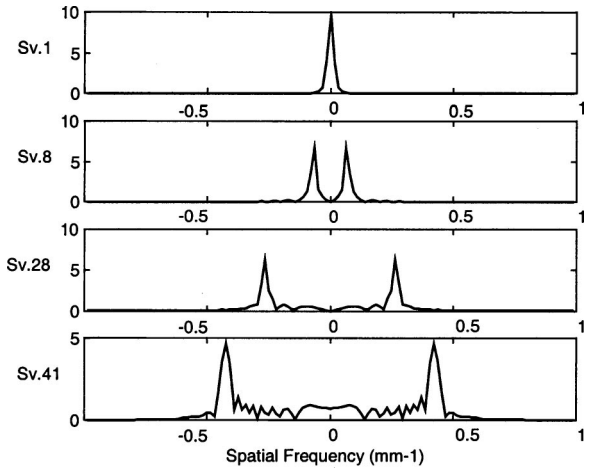


FIG. 9. Absolute value of the spatial Fourier transform of the four previous singular vectors of the propagation matrix $\mathbf{H}(r,\omega=\omega_0)$ at a chosen frequency $f_0=1.6$ MHz.

ticed in Fig. 5 (solid line), the singular-value distribution of this experimental matrix clearly agrees with the computed result. The spatial Fourier transform of the singular vectors of the experimental matrix has also been performed and can be compared in Fig. 11 with the computed one in Fig. 10. We observe again in this experimental acquisition the spatial frequency components represented by each main singular vector. Figure 12 presents the spatial Fourier transform of the singular vectors of the experimental matrix at a higher frequency $f=2.5$ MHz. As the frequency is higher, the number of degrees of freedom and also the exploited spatial frequency range is bigger.

C. Spatial inverse filter focusing: Simulation and experiment

As explained in Sec. III A, we choose as an objective $F=\{0,0,...,0,1,0,...,0\}$. In Fig. 13, we present the focusing pattern F_{pr} obtained for this monochromatic case on the control points. This focusing pattern is first presented for the computed matrix corresponding to Eq. (11). F_{pr} is of course different from the spatial delta vector F. Due to the array limited aperture, the spatial inverse filter is limited by the

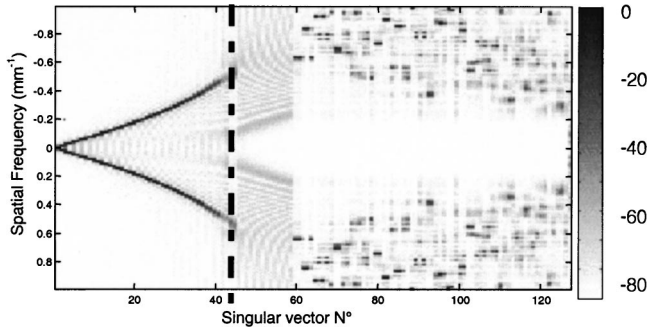


FIG. 10. Spatial Fourier transform in the emitting plane of all singular vectors of the propagation matrix $\mathbf{H}(\omega=\omega_0)$ at a chosen frequency $f_0=1.6$ MHz. One can notice that each singular vector bound to a nonzero singular value represents a single spatial frequency. The dashed line at $N=43$ [corresponding to Eq. (7)] separates the physically relevant singular vectors from the noise. The amplitude is plotted on a dB gray scale.

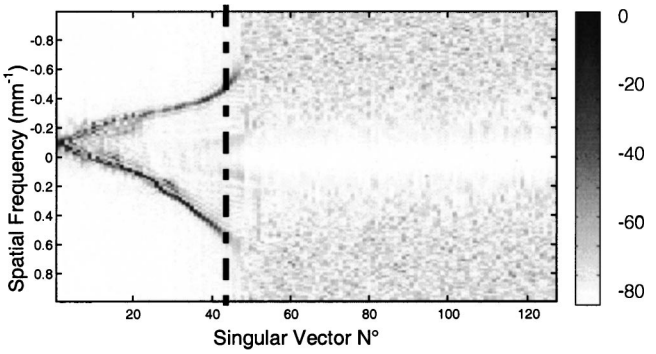


FIG. 11. Experimental acquisition of the propagation matrix $\mathbf{H}(\omega=\omega_0)$ at a chosen frequency $f_0=1.6$ MHz. Here, we represent the spatial Fourier transform in the emitting plane of the whole singular vectors, in others words the spatial Fourier transform of $\mathbf{V}$. The number of freedom degrees agrees with the predicted one ($N=43$, Fig. 5, solid line). The amplitude is plotted on a dB gray scale.

classical diffraction limits (the well-known sinc directivity pattern) and cannot recover a spatial delta function. This illustrates the physical meaning of the main singular vectors as the optimal attainable focusing (here, the sinc function of classical diffraction theory,) corresponding to the projection of the objective F on the main singular vectors.[21]

The experimental acquisition of $\mathbf{H}$ allowed us also to calculate the emission vector E. This vector has been experimentally reemitted by the array and the wave field generated in the control plane has been scanned by a small acoustic receiver. This experimental monochromatic focusing pattern F_{exp} obtained by inverse-filter focusing is also presented in Fig. 13 and shows the same result as the prediction F_{pr} derived from Eq. (11).

IV. THE BROADBAND APPROACH: THE SPATIO-TEMPORAL INVERSE FILTER

The previous results concern only one monochromatic component of the propagation operator. However, these results can be extended to broadband signals. In that case, the propagation operator $\{h_{ij}\}(t)$, defined on temporal domain $0\leqslant t\leqslant \Delta T$, can be described in the Fourier domain by a

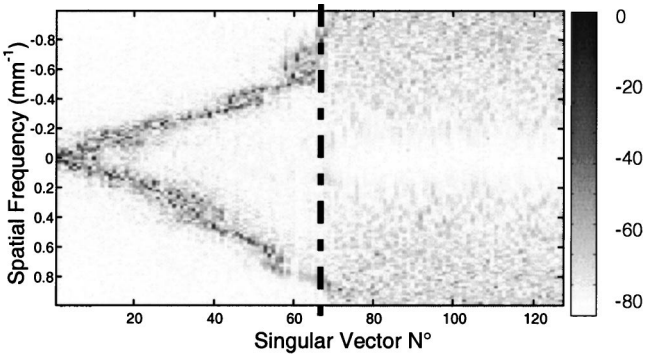


FIG. 12. Experimental acquisition of the propagation matrix $\mathbf{H}(\omega=\omega_0)$ at another frequency $f_0=2.4$ MHz. Here, we represent the spatial Fourier transform in the emitting plane of the whole singular vectors, in others words the spatial Fourier transform of $\mathbf{V}$. The number of freedom degrees agrees with the predicted one $N=68$. The amplitude is plotted on a dB gray scale.

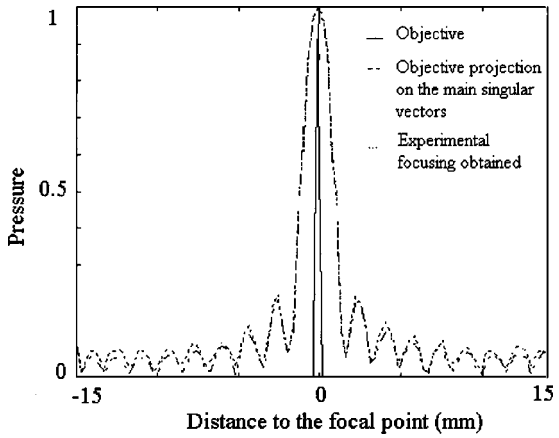


FIG. 13. Directivity patterns in the control plane: the objective F_0 (solid line), the focusing F obtained by monochromatic (spatial) inverse-filter process for the computed matrix (dashed-dotted line), and in the experimental case (dashed line).

whole set of monochromatic propagation matrices $\mathbf{H}(\omega)$. The singular-value decomposition is then applied independently to each monochromatic matrix $\mathbf{H}(\omega)$.

A. The spatio-temporal singular-value distribution

By applying the singular-value decomposition of the propagation operator for each frequency ω_k, the entire singular-value distribution $\lambda_i(\omega)$ can be represented as a 2D surface. This singular-value distribution corresponding to the computed propagator defined by Eq. (10) is presented in Fig. 14. According to Eq. (8), the number of degrees of freedom of each monochromatic matrix depends linearly on the frequency

$$\mathcal{N}(f)=\frac{2L}{c_0}f\sin\left(\tan^{-1}\frac{D}{2F}\right). \tag{14}$$

This straight line $\mathcal{N}(f)$ has been plotted (white line) in Fig. 14(a). As one can readily observe, this definition of $\mathcal{N}(f)$ clearly defines the space corresponding to the relevant singular values in the computed singular-value distribution presented in Fig. 14(a). Figure 14(b) corresponds to the broadband singular-value distribution of the propagation operator acquired experimentally under the same conditions (Fig. 3.) We clearly obtain the same results on these experimental data. Again, the straight line $\mathcal{N}(f)$ given by Eq. (14) fits the area defined by the negligible singular values.

For broadband signals, as space and time are independent variables, the global number $\mathcal{N}$ of spatio-temporal degrees of freedom induced by the emitting array of transducers in the control plane corresponds to the summation over frequency of the number of monochromatic degrees of freedom multiplied by the duration of the signal, ΔT

$$\mathcal{N}=\Delta T\int_{f_{\min}}^{f_{\max}}N(f)df. \tag{15}$$

Integration over frequency gives

$$\mathcal{N}=\frac{L\Delta T}{c_0}\sin\left(\tan^{-1}\frac{D}{2F}\right)(f_{\max}-f_{\min})(f_{\max}+f_{\min}), \tag{16}$$

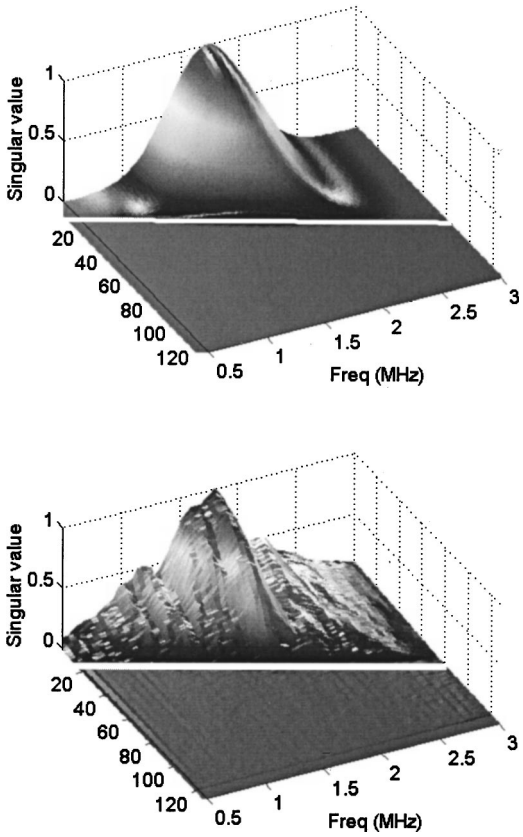


FIG. 14. Spatio-temporal singular-values distribution of the simulated (a) and experimental (b) propagation operator $\{h_{ij}(r,t)\}$.

where $\{f_{\min},f_{\max}\}$ defines the frequency bandwidth of the array. If we introduce B_f, the frequency bandwidth of the array ($B_f=f_{\max}-f_{\min}$) and its central frequency $f_c=(f_{\max}+f_{\min})/2$, we obtain

$$\mathcal{N}=\frac{2Lf_c}{c_0}B_f\Delta T\sin\left(\tan^{-1}\frac{D}{2F}\right). \tag{17}$$

We recognize here the product of the number of degrees of freedom for the temporal dimension $B_f^*\Delta T$ with the spatial number of degrees of freedom for the central frequency. The result comes from the mean value theorem combined with the fact that the spatial number of degrees of freedom increases linearly with the frequency. This number $\mathcal{N}$ corresponds to the number of spatially and temporally independent signals that can be induced by the emitting array in the control plane in a finite time domain ΔT. In the Fresnel assumption, one can verify that this number $\mathcal{N}$ is easily reduced to the simple product

$$\mathcal{N}=\frac{L\Delta TB_f}{\delta}. \tag{18}$$

This result is the extension in both time and space dimensions of Eq. (9) for which $\mathcal{N}$ was found equal to the number of independent focals spots (whose width was equal to $\delta=\lambda F/D$) that the array is able to generate in the aperture L. Here, we take also the number of impulse responses (whose length is equal to $1/B_f$) that can be independently introduced in the time window ΔT.

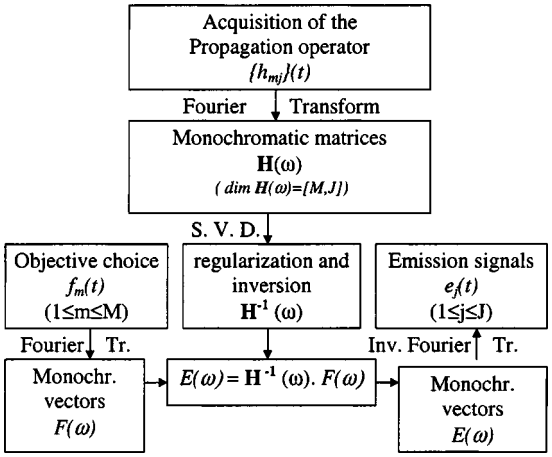


FIG. 15. The spatio-temporal inverse-filter focusing process.

B. The spatio-temporal inverse-filter process

The regularization process can be applied to the whole set of monochromatic matrices by only keeping the relevant singular vectors and the inverse matrices $\hat{\mathbf{H}}^{-1}(\omega)$ can be computed at each frequency. We can choose a spatio-temporal objective $\{f_m(t)\}$ in the control plane, corresponding to the m temporal signals we would like to generate on the set of m control points in a given temporal domain ($t \in [0,\Delta T]$). This set of temporal signals can be expressed in the Fourier domain as a set of monochromatic vectors $F(\omega)$. Then, as explained in Sec. III A, the optimal emission vectors $E(\omega)$ can be computed at each frequency ω according to Eq. (4). Finally, the optimal set of broadband emission signals $\{e_j(t)\}$ is simply deduced by an inverse Fourier transform

$$\forall j \in [1,J], \quad e_j(t) = \mathcal{F}^{-1}(E_j(\omega)) = \int_\omega E_j(\omega) e^{-i\omega t}\, d\omega, \tag{19}$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform. According to Eq. (4) and (19), each optimal emission signal can also be expressed

$$e_j(t) = \int_\omega \sum_{m=1,\ldots,M} \hat{\mathbf{H}}_{jm}^{-1} F_m(\omega) e^{-i\omega t}\, d\omega$$

$$= \sum_{m=1,\ldots,M} \mathcal{F}^{-1}(\hat{H}_{jm}^{-1} F_m(\omega)). \tag{20}$$

Thus, we can define an inverse propagation operator $\{h_{jm}^{-1}\}(t)$ as a set of $[J \times M]$ signals defined by

$$\forall \{j,m\} \in [I,J]x[1,M], \quad \hat{h}_{jm}^{-1}(t) = \mathcal{F}^{-1}(\hat{\mathbf{H}}_{jm}^{-1}(\omega)). \tag{21}$$

This inverse propagation operator allows us to express the J broadband signals $e_j(t)$ to be emitted on the array transducers which will give rise, after propagation in the medium, to the broadband signals $f_m(t)$ received on each control point

$$\forall j \quad 1 \leqslant j \leqslant J, \quad e_j(t) = \sum_{m=1}^{M} \hat{h}_{jm}^{-1}(t) {}_t^{*} f_m(t). \tag{22}$$

The whole space-time inverse-filter focusing process is described by Fig. 15. Assume that we acquired in a first step the whole diffraction operator $h_{mj}(t)$. In a second step, we can easily invert each monochromatic matrix and deduce the inverse propagation operator $\{\hat{h}_{jm}^{-1}\}(t)$. The regularization process that consists of keeping only the main singular vectors is applied for each monochromatic matrix, in other words in both space and time. This regularization process is achieved by keeping only the nonzero singular values of the propagation matrices studied in Sec. IV A.

In order to achieve an optimal focusing in the control plane, we have to choose in a third step an objective $\{f_m\}(t)$ corresponding to the signals we would like to be received in the control plane. The simplest one corresponds to a spatio-temporal delta function creating a temporal delta function at time $t = t_0$ and at a single position $m = m_0$

$$\forall m \quad 1 \leqslant m \leqslant M, \quad f_m(t) = \delta_m(t - t_0) = \delta(m - m_0)\,\delta(t - t_0),$$

$$0 \leqslant t \leqslant \Delta T. \tag{23}$$

Of course, the spatio-temporal objective can be completely different and chosen to meet particular needs depending on the encountered applications. The objective chosen will be addressed in the next subsection.

C. Comparison with classical and time-reversal focusing

In this section, the inverse-filter process is applied in a homogeneous medium (water) and its focusing ability is compared with the classical cylindrical focusing implemented on commercial echographic devices and with the time-reversal focusing technique.

For broadband signals, the choice of the spatio-temporal focusing objective $\{f_m(t)\}$ is a crucial component of the process: we must define the temporal signals we would like to be received during a time window of duration ΔT on each control point. The most straightforward objective corresponds to the spatio-temporal delta function defined in Eq. (23). This objective has been used in the spatio-temporal inverse-filter process of the computed propagation operator [defined by Eq. (10)]. Given this objective, a set of emission signals was calculated and the resultant focusing obtained is presented in Fig. 16(a). This focusing is compared to the one obtained with a classical focusing technique consisting to apply a temporal delta function with a cylindrical time-delay law on each array element, Fig. 16(b).

The inverse filter attempts at each frequency to minimize the difference between the focusing pattern and the focusing objective in terms of mean least squares. So, the use of a spatial delta function in this process enhances lateral resolution rather than lateral contrast whereas the classical focusing technique favors lateral contrast over lateral resolution. Thus, the inverse-filter technique introduces a weak ''antiapodization'' on the emitting aperture to partially compensate the edge elements directivity. Consequently, it increases the lateral resolution of the focal spot to best approach the objective. This result is corroborated here, as the sidelobe level is slightly better (about 2 dB) in Fig. 16(b) than in Fig. 16(a).

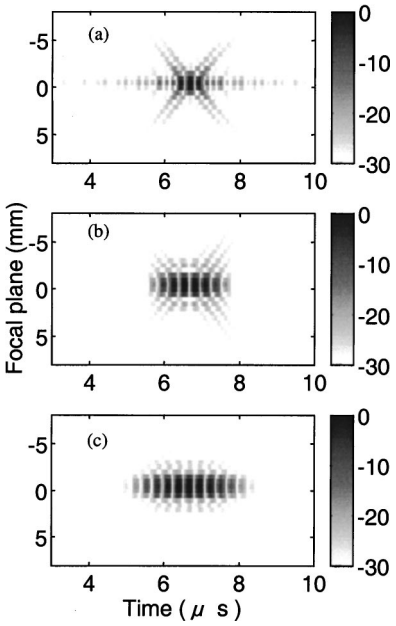


FIG. 16. Simulation results. Focusing patterns in the control plane: (a) spatio-temporal inverse filter using a delta function as the objective: (b) classical focusing, and (c) time-reversal focusing.

Moreover, the same phenomenon appears in the temporal domain: the use of a temporal delta function in the inverse-filter process favors the bandwidth—in others words the axial resolution—to detriment of axial contrast. For our purpose, Fig. 17 presents the Fourier components of the signal applied to an element of the array in the case of inverse-filter processing and classical focusing. For the classical focusing, a temporal delta function is applied to the element, so its Fourier transform is flat. On the contrary, the inverse-filter focusing process proposes a different emission signal: it tries to compensate the element bandwidth. This bandwidth compensation improves the axial resolution of the focal spot. Figure 17(b) represents the Fourier transform of the signal

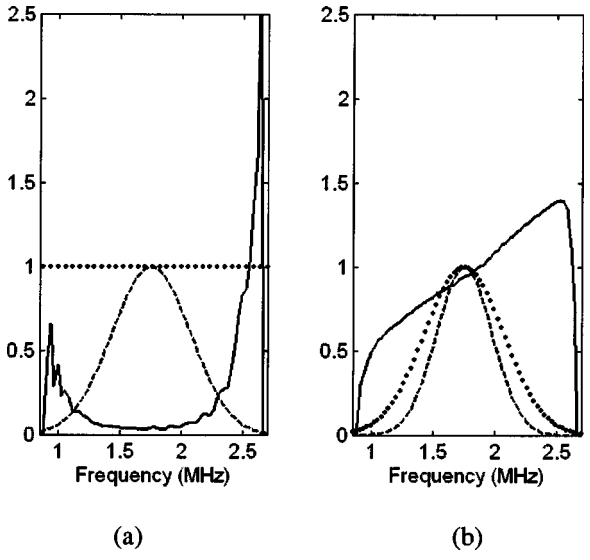


FIG. 17. Fourier spectrum of one transducer emission signals (a) and Fourier spectrum of the corresponding signals received at focus (b) for the three focusing techniques: inverse filter (solid line), classical focusing (dotted line), and time reversal (dashed line).

received at the focus by inverse-filter processing (solid line) and classical focusing (dotted line). As one can observe, the signal received at the focus with inverse-filter processing has a broader bandwidth than the one due to classical focusing. This spectrum of the signal received at focus by inverse-filter processing is not flat, but the higher frequencies are slightly advantaged [Fig. 17(b)]. Indeed, due to our objective choice (a spatio-temporal delta function) the number of singular vectors taken into account during inversion is greater for high frequencies than for low frequencies. So, the signal received at focus contains more high-frequency components. Of course, the axial resolution improvement induced by the inverse filter comes at the expense of the maximal pressure amplitude received at the focus: as the transducer is used in a forced regime, the ratio between the energy received at focus and the energy applied on the emitting array decreases.

The inverse and classical focusing techniques are finally compared to time-reversal focusing. Figure 16(c) presents the spatio-temporal signals received in the focal plane by using this last technique. The axial and lateral resolutions are slightly degraded in comparison with the other techniques. Indeed, the emission signals correspond in that case to the time-reversed version of the focus Green's function recorded on the array. Thus, the Fourier transform of the emission signals applied to the array elements will correspond to the bandwidth of the transducer, as presented in Fig. 17(a) (dashed line). During reemission, the signals experience for a second time the bandwidth of the elements before propagating in the medium and so, the Fourier spectrum of the signal received in the focal plane is slightly narrowed [Fig. 17(b) (dashed line)]. Thus, the axial resolution is slightly poorer than for classical and inverse filter. The same phenomenon appears for the spatial contribution. Due to the directivity of the transducers, the wavefront coming from the focal spot location is apodized on the array aperture. After time reversal and reemission, this apodization of the wavefront induces a slight decrease of the lateral resolution. One must, however, keep in mind that these changes in lateral and axial resolution are very small. What is the counterpart to this small decrease? As shown in previous papers, we know that the time-reversal process is a spatial[10] and temporal[22] matched filter, provided that the spatial reciprocity of the medium is valid: it maximizes the ratio between the energy received at focus and the energy applied on the emitting array. So, time-reversal focusing optimizes the sensitivity of the focusing.

These three focusing patterns are finally compared with the experimental results, Figs. 18(a), (b), (c). We find again the same results. Time-reversal focusing has a better sidelobe level (due to the apodization effect) than the classical focusing and a slightly poorer axial resolution (as signals experience two times the transducers' bandwidth). In that experimental case, the inverse-filter focusing pattern is obtained by applying the inverse process to the experimentally acquired propagation operator. The inverse-filter focusing is here significantly better in comparison with the two other techniques (more than in the simulation case in Fig. 16). Indeed, in the simulation case, all transducers were assumed to be identical. Now, in the experimental configuration, array elements have of course nonuniform characteristics in terms of sensitivity,

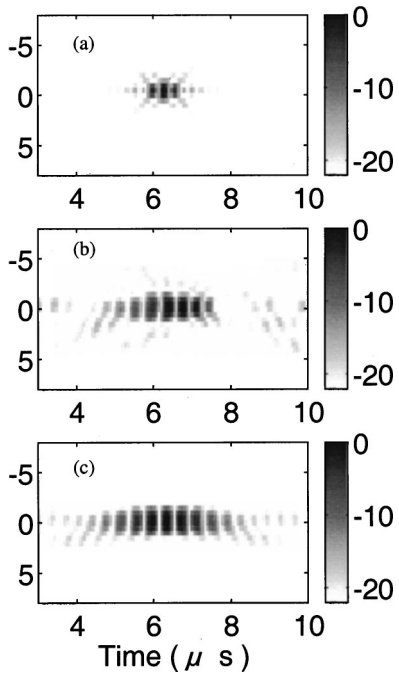


FIG. 18. Experimental results in water. Focusing patterns in the control plane: (a) spatio-temporal inverse filter using a delta function as the objective: (b) classical focusing, and (c) time-reversal focusing.

impulse response, and element position and directivity.[23] These temporal as well as spatial discrepancies are taken into account and optimally corrected by the spatio-temporal inverse filter. However, one must not forget that the pressure amplitude received at the focus is much smaller during inverse filter focusing in comparison with time reversal (the most ''sensitive'' technique): in our water experiment, the ratio between focal-pressure amplitudes between these two techniques was 13 dB.

D. Inverse filter focusing in heterogeneous and absorbing medium: Choice of the objective

As just seen, due to limited bandwidth and size of the transducers, there can be some small differences even in water between spatio-temporal inverse filter, time reversal, and classical focusing. Moreover, one can notice that the use of a spatio-temporal delta function in the inverse process is not necessarily judicious. The improvement of axial and lateral resolutions comes at the expense of a non-negligible loss of sensitivity. However, if we do not want to strongly compensate the transducer's bandwidth, we only have to change the objective to a more realistic one. For example, the spatio-temporal focusing induced by the array aperture in a homogeneous medium can be numerically computed and chosen as the objective.

In this paper, the study has only been carried out for a homogeneous and nonabsorbing medium. However, as soon as the medium is heterogeneous and absorbing, the previous weak differences between the three focusing techniques become very important. Indeed, whereas classical focusing is strongly degraded by medium heterogeneities, time-reversal focusing remains very good as long as information losses are not induced during the time-reversal process.[13] In absorbing media, for instance, its focusing is unfortunately also degraded. In that final extreme case, only inverse filter focusing should remain optimal as it compensates dissipation effects.

In such heterogeneous and absorbing media, the choice of the objective in the inverse process is of great importance. Two configurations can be encountered. If we have no idea about the spatio-temporal focusing pattern (resolution and contrast) that we can reasonably reach (for example, in reverberating or multiply scattering or wave guided media), a spatio-temporal delta function may be chosen as the objective. If the sensitivity decreasing of the inverse-filter focusing due to this particular objective choice is too important, we can fix another objective: the spatio-temporal focusing reached by the same system in homogeneous nonabsorbing medium (obtained either experimentally or by simulation). An experimental illustration of the importance of the choice of objective during the inverse-filter process will be given in a companion paper for an experiment attempting to focus through titanium plates.

V. CONCLUSION

As it is based on the knowledge of the entire propagation operator relating an array of transducers to a set of control points in the medium, the spatio-temporal inverse-filter focusing technique allows one to attain the optimal focusing at each control point. As seen in the simple case of a homogeneous medium, the choice of the focusing objective in the inverse-filter process allows one to define precisely the desired axial and lateral resolutions. Thus, with an extreme choice of a spatio-temporal delta function as the objective, the axial and lateral resolutions can be strongly improved by this broadband inverse technique subject to a non-negligible loss of sensitivity. As we will see in Paper II,[24] the slight differences between classical, time-reversal, and the spatio-temporal inverse-filter focusing techniques in homogeneous medium become very important as soon as the medium is heterogeneous and absorbing. It straightforwardly emphasizes the great potential of multiple-channel systems having the ability to apply completely different signal waveforms on each transducer of the array. Thus, the application of this broadband inverse-filter focusing technique could be of great interest in various ultrasonic fields such as medical imaging, nondestructive testing, and underwater acoustics.

[1] D. N. White, J. M. Clark, J. N. Chesebrough, M. N. White, and J. K. Campbell, ''Effect of skull in degrading the display of echoencephalographic B and C scans,'' J. Acoust. Soc. Am. **44**, 1339–1345 (1968).
[2] M. O'Donnell and S. W. Flax, ''Phase aberration measurements in medical ultrasound: Human studies,'' Ultrason. Imaging **10**, 1–11 (1988).
[3] Q. Zhu, B. D. Steinberg, and R. Arenson, ''Wavefront amplitude distortion and sidelobes levels. I. Theory and computer simulations,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **40**, No. 6, 747–753 (1993).
[4] Q. Zhu, B. D. Steinberg, and R. Arenson, ''Wavefront amplitude distortion and sidelobes levels. II. *In vivo* experiments,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **40**, No. 6, 754–762 (1993).
[5] F. Wu, M. Fink, R. Mallart, J.-L. Thomas, N. Chakroun, D. Cassereau, and C. Prada, ''Optimal focusing through aberrating media: A comparison between time reversal mirror and time delay correction technique,'' in *Proc. IEEE 1991 Ultrason. Symp.*, pp. 1195–1199.
[6] F. Wu, J.-L. Thomas, and M. Fink, ''Time reversal of ultrasonic fields. II. Experimental results,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control, **39**, No. 5, 567–578 (1992).

[7] D.-L. Liu and R. C. Waag, ''Correction of ultrasonic wavefront distortion using backpropagation and a reference waveform method for time-shift compensation,'' J. Acoust. Soc. Am. **96**, 649–660 (1994).
[8] M. Fink, ''Time reversal of ultrasonic fields. I. Basic principles,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, No. 5, 555–566 (1992).
[9] M. Fink, ''Time-reversed acoustics,'' Phys. Today **20**, 34–40 (1997).
[10] M. Tanter, J.-L. Thomas, and M. Fink, ''Time reversal and the inverse filter,'' J. Acoust. Soc. Am. **108**, 223–234 (2000).
[11] J. L. Thomas, F. Wu, and M. Fink, ''Time reversal mirror applied to lithotripsy,'' Ultrason. Imaging **18**, 106–121 (1996).
[12] C. Prada, F. Wu, and M. Fink, ''The iterative time reversal mirror: A solution to self-focusing in pulse-echo mode,'' J. Acoust. Soc. Am. **90**, 1119–1129 (1991).
[13] H. C. Song, W. A. Kuperman, W. S. Hodgkiss, T. Akal, C. Ferla, and D. R. Jackson, ''Iterative time reversal in the ocean,'' J. Acoust. Soc. Am. **105**, 3176–3184 (1999).
[14] V. Martin, ''Active control of moving sound source radiation—numerical modeling in the space-frequency and space-time domains,'' J. Sound Vib. **228**(3), 477–509 (1999).
[15] N. Kahana and H. Kirkeby, ''A multiple microphone recording technique for the generation of virtual acoustic images,'' J. Acoust. Soc. Am. **105**, 1503–1516 (1999).
[16] M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Wiley, New York, 1991).
[17] D. Gabor, in *Proceedings of the Symposium on Astronomical Optics*, edited by Z. Kopal (North-Holland, Amsterdam, 1956), p. 17.
[18] W. H. Press, *Numerical Recipes in C* (Cambridge University Press, Cambridge, 1988).
[19] C. Prada and M. Fink, ''Eigenmodes of the time reversal operator: A solution to selective focusing in multiple target media,'' Wave Motion **20**, 151–163 (1994).
[20] F. Lin, A. I. Nachman, and R. C. Waag, ''Quantitative imaging using a time domain eigenfunction method,'' J. Acoust. Soc. Am. **108**, 899–912 (2000).
[21] M. Tanter, ''Application du retournement temporel à l'hyperthermie ultrasonore du cerveau,'' Thesis manuscript, University of Paris VII, May 1999.
[22] C. Dorme and M. Fink, ''Focusing in transmit–receive mode through inhomogeneous media: The time reversal matched filter approach,'' J. Acoust. Soc. Am. **98**, 1155–1162 (1995).
[23] D.-L. Liu, R. C. Waag, and C. Oakley, ''Compensation for array nonidealities by electronics,'' J. Acoust. Soc. Am. **101**, 3115 (1997).
[24] J. F. Aubry, M. Tanter, J. Gerber, J. L. Thomas, and M. Fink, ''Optimal focusing by spatio-temporal inverse filter. II. Experiments. Application to focusing through absorbing and reverberating media,'' J. Acoust. Soc. Am. **110**, 48–58 (2001).
[25] R. Seip, P. VanBaren, and E. Ebbini, ''Dynamic focusing in ultrasound hyperthermia using implantable hydrophone arrays,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **41**, 706713 (1994).
[26] E. S. Ebbini and C. Cain, ''Multiple-focus ultrasound phased-array pattern synthesis: Optimal driving signal distribution for hyperthermia,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **36**, 540548 (1989).

Optimal focusing by spatio-temporal inverse filter. II. Experiments. Application to focusing through absorbing and reverberating media

J.-F. Aubry, M. Tanter, J. Gerber, J.-L. Thomas, and M. Fink
Laboratoire Ondes et Acoustique, ESPCI, Université Paris VII, U.R.A. C.N.R.S. 1503, 10 rue Vauquelin, 75005 Paris, France

(Received 16 January 2001; accepted for publication 9 April 2001)

To focus ultrasonic waves in an unknown heterogeneous medium using a phased array, one has to calculate the optimal set of signals to be applied on the transducers of the array. (In most applications of ultrasound, medical imaging, medical therapy, nondestructive testing, the first step consists of focusing a broadband ultrasound beam deeply inside the medium to be investigated.) Focusing in a homogeneous medium simply requires to compensate for the varying focus–array elements geometrical distances. Nevertheless, heterogeneities in the medium, in terms of speed of sound, density, or absorption, may strongly degrade the focusing. Different techniques have been developed in order to correct such aberrations induced by heterogeneous media (time reversal, speckle brightness, for example). In the companion to this paper, a new broadband focusing technique was investigated: the spatio-temporal inverse filter. Experimental results obtained in various media, such as reverberating and absorbing media, are presented here. In particular, intraplate echoes suppression and high-quality focusing through a human skull, as well as hyper-resolution in a reverberating medium, will be shown. It is important to notice that all these experiments were performed with fully programmable multichannel electronics whose use is required to fully exploit the spatio-temporal technique.
[DOI: 10.1121/1.1377052]

PACS numbers: 43.20.−f, 43.20.Bi, 43.20.El, 43.20.Fn [ANN]

I. INTRODUCTION

In most applications of ultrasound, such as medical imaging, medical therapy, or nondestructive testing, the first step consists in focusing a broadband ultrasonic beam deeply inside the medium under investigation. Focusing in a homogeneous medium simply requires compensation for the varying geometrical distances between the array elements and the focus. Nevertheless, heterogeneities in the medium, in terms of speed of sound, density, or absorption, may strongly degrade the focusing.[1–3] Different techniques have been developed in order to correct such aberrations induced by heterogeneous media (for example, time reversal[4] and speckle brightness[5]). In the companion to this paper, we investigated a new broadband focusing technique: the spatio-temporal inverse filter. This technique is based on the inversion of the propagation operator relating the elements of a transducer array to a set of control points embedded in the medium and allows one to calculate the optimal set of signals to be emitted by each independent element of the array. Indeed, it will be shown in this article that if we want to attain optimal focusing through a complex medium, distinct signal waveforms have to be emitted on each transducer of the array. A clear implication of this work is the great potential of multichannel systems made up of independently controlled transducers

The domain of potential applications of the spatio-temporal inverse filter is vast. In this paper, we present results of experimental application of this technique in various media which demonstrate its usefulness in nondestructive testing, medical imaging, and therapy. In each case, inverse filter focusing will be compared to time reversal or classical cylindrical focusing. In Sec. II, we study the focusing through a solid plate. In this case, the focusing is perturbed by an anterior wavefront due to reflections.[6] These intraplate echoes or reverberations degrade not only the lateral focusing but also the axial resolution. We will see that inverse filtering completely suppresses the intraplate echoes and optimizes both axial and lateral resolution. This is an interesting result for the nondestructive testing community, as it could enable the detection of defects behind plates, tubes, spars, ribs, or more complicated structures.

In Sec. III, we study an absorbing sample located between the array of transducers and the focal plane. In that case, absorption strongly degrades the classical cylindrical focusing. Time-reversal focusing,[7] while able to correct for sound-speed heterogeneities, is also degraded in such an absorbing medium. On the contrary, the inverse filter compensates for both kind of heterogeneities, namely sound speed and absorption, and achieves optimal focusing.

As the first two parts show that the inverse filter is a very efficient way to focus through aberrating media, it could lead to interesting medical applications, especially in imaging. One of the most problematic media for ultrasound in the human body is certainly the skull.[8] Indeed, the skull induces both large phase aberrations and strong absorption. This is why ultrasonic brain imaging remains beyond current technological limitations. We show, in Sec. IV, that it is possible to calculate by broadband inverse filter processing a set of signals to be emitted by each element of a transducer array

 0001-4966/2001/110(1)/48/11/$18.00

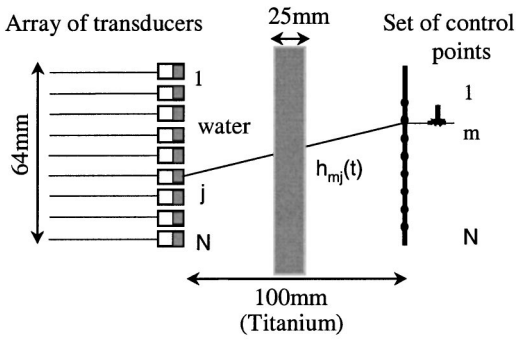


FIG. 1. A titanium plate is located between the 127 control points and the array of 127 transducers.

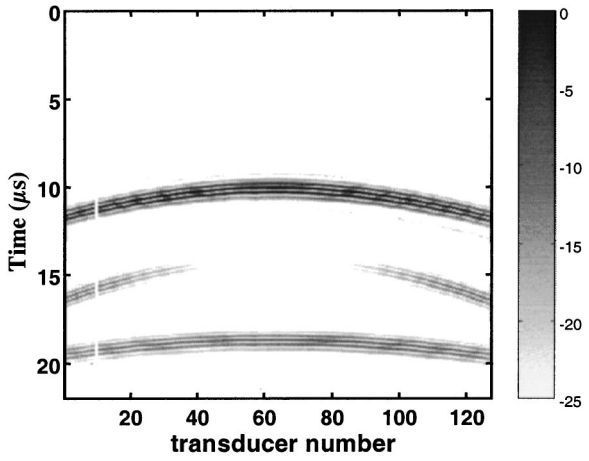


FIG. 2. Waveform received through the titanium plate located 40 mm from the array. Signal amplitude is shown on a dB gray scale; the horizontal axis is the transducer number and the vertical axis is the evolution in time.

that optimally focus through the skull; moreover, we demonstrate that the spatial and temporal focusing obtained in water and through the skull have the same quality. This constitutes the first step toward ultrasonic brain imaging and could have important diagnostic applications, as it could be possible to practice real-time flow mapping in the brain at a very low cost, compared to magnetic resonance imaging.

Finally, in Sec. V, regarding the results we obtained in different ultrasonic applications, we study the advantages that this technique could have for the propagation in very complex and reverberating structures, as it can be involved in nondestructive testing but also in sound propagation through urban geometries. Indeed, the inverse filter is particularly suited to precisely focus the information in space and time, taking advantage of the complexity of the medium. In particular, we observe an important hyper-resolution[9,13] by using the inverse filter in the reverberating medium presented in the last part.

II. SUPPRESSION OF INTRAPLATE ECHOES

When focusing with a cylindrical law through a metallic plate, the main wavefront is generally followed by secondary wavefronts due to reflections internal to the plate. In this section, we present results of an experiment in which we use the spatio-temporal inverse-filter technique to ensure an optimal focusing pattern while at the same time suppressing the intraplate echoes.

A. Experimental setup

All the experiments presented in this article were carried out with an apparatus which has been discussed in previous work.[7] The salient features of the setup are shown in Fig. 1. The linear transducer array is made of 128 piezoelectric elements working at 1.5 MHz. The aperture of each transducer element is 0.5 mm wide and 10 mm high. The array pitch is 0.5 mm. The total aperture of the array of transducers is thus equal to 64 mm. The 128 elements are connected to a fully programmable electronic apparatus which includes 128 32-Kbyte buffer memories. This enables us to record and emit any temporal signal on each transducer of the array. A single transducer element 1 mm wide and 10 mm high is placed in the focal plane. This monoelement is translated at each location of the control points and used in receive mode as a hydrophone to acquire the signals coming from the array. This monoelement is also used to scan the pressure field in order to check the focusing. Each transducer is connected to a fully programmable channel that enables us to record and emit any temporal shape.

For the study of plate echoes, a uniform 25-mm-thick titanium plate is placed between the monoelement and the transducer array ($V_L = 6.1\,\text{mm}/\mu\text{s}$, $V_T = 3.1\,\text{mm}/\mu\text{s}$). The transducer elements and the titanium plate are immersed in water. We used 127 control points with a regular 0.3-mm pitch.

One can see the influence of the titanium plate by emitting a pulse at the central control point location (i.e., at the $N/2$nd position) and by recording the signals received on all the transducers of the array. On a B-mode representation, the signals recorded on each transducer are then plotted in a gray scale as a function of time in Fig. 2.

A first principal wavefront arrives at time 11 μs corresponding to the direct transmission of the longitudinal wave, or L wave, that arose on the first interface and was transmitted through the second interface. There can be seen on the plot two symmetric arrivals at 15 μs centered at ± 6 mm. These correspond to transverse waves, T waves, in the plate. Indeed, at the first interface a part of the wavefront was converted into a T wave propagating inside the plate. After propagation through the plate, this wavefront is transmitted through the second interface and radiates an L wave in water that propagates to the array. As the transverse wave speed is 2 times smaller than the longitudinal one, it arrives 4 μs later. Note that no transverse wave arrives in the center due to vanishing longitudinal–transverse wave coupling at direct incidence. A third wavefront arrives at 19 μs (8 μs after the direct one), due to the first intraplate reflection of the L wavefront.

B. Experimental results

The propagation operator $\{h_{mj}(t)\}$ introduced in the companion article, (please see p. 19) hereafter referred to as Paper I,[14] is acquired as follows: each transducer j emits successively a chirp that is recorded at each control point m. A chirp rather than a pulse is emitted in order to increase the

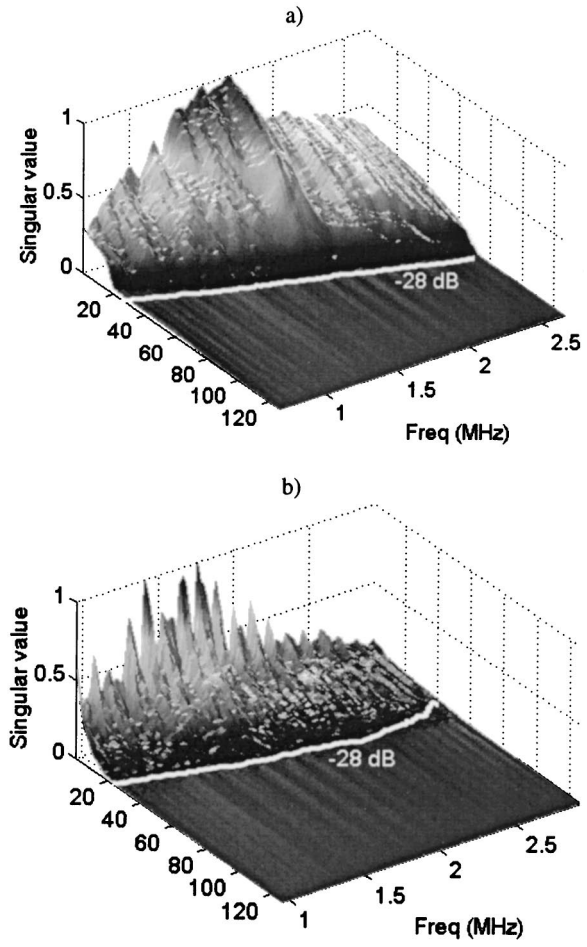


FIG. 3. Spatio-temporal singular-values distribution of the propagation operator $\{h_{mj}(t)\}$ acquired (a) in water and (b) through a titanium plate.

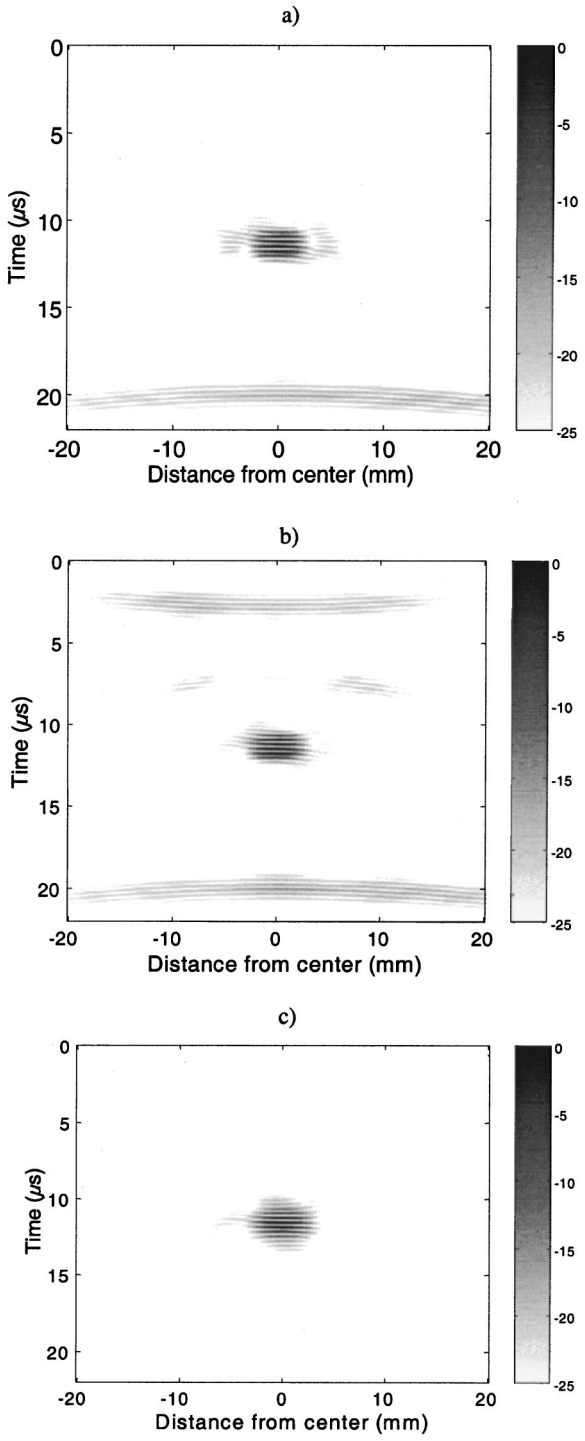


FIG. 4. (a) B scan in the focal plane obtained when focusing with the cylindrical law. The amplitude of the signals recorded on each of the 128 transducers is represented in dB on a gray scale as a function of time. (b) B scan in the focal plane obtained when using time reversal. (c) B scan in the focal plane obtained with the inverse filter.

sensitivity on frequencies far from the central one. Deconvolution of the received signal by the signal applied on the transducer gives the impulse response $h_{mj}(t)$. We denote the ensemble of signals $\{h_{mj}(t)\}$ as the $M\times J$ matrix $H(t)$. A Fourier transform of each element of the matrix $H(t)$ gives $H(\omega)$.

The inversion of the matrix $H(\omega)$ was carried out using the singular-value decomposition technique, giving the noise-filtered inverse propagation matrix $\hat{\mathbf{H}}^{-1}(\omega)$. As explained in Paper I, an inversion is carried out at each frequency ω. The singular-value distribution and corresponding cutoff is given in Fig. 3(b), the singular-value distribution obtained in a homogeneous medium given as a reference in Fig. 3(a). Through the plate, strong peaks appear at regularly spaced frequencies corresponding to resonant modes inside the plate. Singular values that are below the -28-dB threshold are considered as noise and thus are not taken into account in the inversion.

Having calculated $\hat{\mathbf{H}}^{-1}(\omega)$, it is straightforward to obtain the inverse-filter-derived emission vector $\{e^{1\mathrm{F}}(t)\}$ using Eq. (19) given in the companion to this article and choosing a particular objective in the focal plane. We emit this set of signals with the array of transducers and compare the signals received in the focal plane to time-reversal and cylindrical-law focusing. The B scans (spatio-temporal representation of the temporal signals received on a whole set of transducers) obtained in the focal plane by using cylindrical-law focusing, time-reversal, and the inverse filter are presented in Figs. 4(a), (b), and (c), respectively.

As expected, the cylindrical-law focusing wavefront shows a focal spot (11 μs) followed by a broad reflected wavefront (19 μs) arriving 8 μs later and corresponding to the first intraplate reflection of the longitudinal wave [Fig.

4(a)]. The part of the signals that was converted to T waves while passing through the plate does not result in a significantly perturbing wavefront.

Figure 4(b) shows the result of time-reversal focusing obtained by time reversing the signals previously presented in Fig. 2 and re-emitting them on the array of transducers. As time reversal corresponds at focus to the autocorrelation function of the received signal, which is made of two wavefronts, one can see three wavefronts. Basically, two main wavefronts are successively emitted corresponding to the direct transmission of the L wave and its first reflection. First, the reflected wavefront is emitted and will arrive at time 3 μs on the B scan. The corresponding reflected wavefront exists but has a small amplitude and will interfere constructively with the main wavefront transmitted later. Then, the main wavefront is emitted and focuses straight at the focal point; but, a part of this wavefront is reflected by the plate and gives rise to a second wavefront at 21 μs that is similar to the first one received at 3 μs.

This experiment also illustrates that even for lossless media, the result of the time reversal may not be optimal as soon as information losses occur during the process. Here, the loss of information obviously comes from the nonrecorded backscattered wave: a second time-reversal mirror would be needed on the opposite side to perform a complete time-reversal experiment. However, for more complex distribution of heterogeneities, only a time-reversal cavity would ensure time-reversal invariance. When loss of information occurs, time-reversal invariance is broken; however, time-reversal process achieves a spatio-temporal matched filter as long as linearity and reciprocity remain valid.[12]

Figure 4(c) shows the B scan obtained by emitting the signal $\{e^{1F}(t)\}$ calculated with the inverse filter. For the result presented here, we take as our objective the main focus of the time-reversal experiment. That is to say, our objective consists only of the main focus of the time-reversal experiment [at time 13 μs in Fig. 4(b)]: before and after this focus we force the objective signals to be equal to zero, at the exact places where plate echoes induced undesired wavefronts. It is evident on the experimental focusing pattern obtained by inverse filter [Fig. 4(c)] that the intraplate reflections have been largely suppressed as well as parasite signals due to mode conversion. In fact, the obtained focusing pattern corresponds almost perfectly to the chosen objective.

Visual inspection of $\{e^{IF}(t)\}$ (Fig. 5) confirms how this is done. A main wavefront is emitted at time 8 μs. The longitudinal part of this wavefront generates the focal spot at $t=12$ μs. A second wavefront (16 μs in Fig. 5) is emitted later and will destructively interfere with the intraplate echoes of the main wavefront.

In this case, the intraplate echoes are well suppressed but, due to the plate refraction angle, the effective aperture is quite small and the resolution is quite low. Indeed, as the sound velocity in the plate is higher than in water, the effective aperture is reduced according to the Snell–Descartes laws. We wanted to check if the system could reach a higher resolution by choosing as an objective a spatio-temporal Dirac delta function.

Figure 6 shows the experimental B-scan obtained by

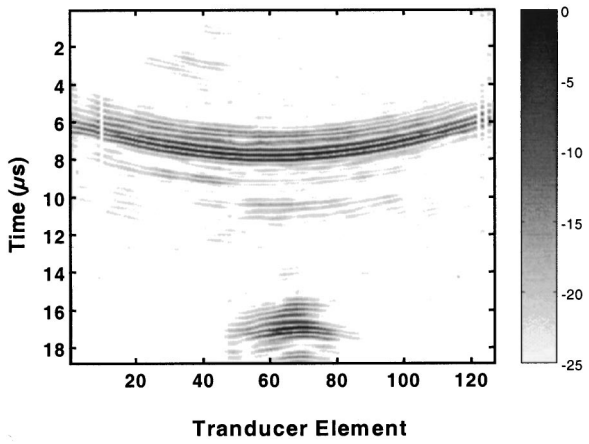


FIG. 5. Emission vector $\{e^{IF}(t)\}$ calculated by the inverse filter. The plate echoes of main front emitted at 7 μs will destructively interfere with the second wavefront emitted at 16 μs.

emitting the signal $\{e^{IF}(t)\}$ calculated with the inverse filter using a spatio-temporal delta function objective. We see that the intraplate reflections are still largely suppressed as well as parasite signals due to mode conversion. Moreover, the focal point is now much finer compared to the B-scan presented in Fig. 4(c). One can readily observe the improvement in the focusing quality compared to the classical focusing obtained in Fig. 4(a).

It is of great interest to look at the set of emission signals $\{e^{IF}(t)\}$ calculated with the inverse-filter focusing technique (Fig. 7). Basically, as in the previous experiment (Fig. 5) the main wavefront can be seen passing through 7 μs on the first array element, followed by a second wavefront emitted at 15 μs, which will suppress the intraplate reflections due to the ''direct'' wavefront. In order to have a broader focal spot, the main wavefront is antiapodized: this is the simplest way to enhance the aperture of the array. But, there exists another way to increase the effective aperture of the transducer array, which is limited by the plate refraction in this geometry. Because transverse waves travel more slowly than longitudinal waves, they undergo less refraction. Con-

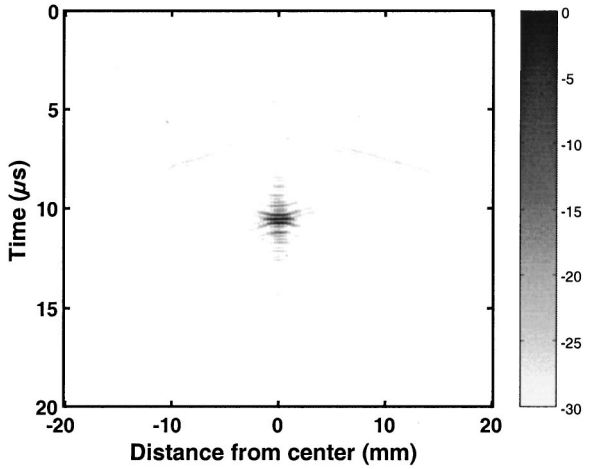


FIG. 6. B scan in the focal plane obtained with the inverse filter. The amplitude of the signals recorded on each of the 128 transducers is represented in dB on a gray scale as a function of time.

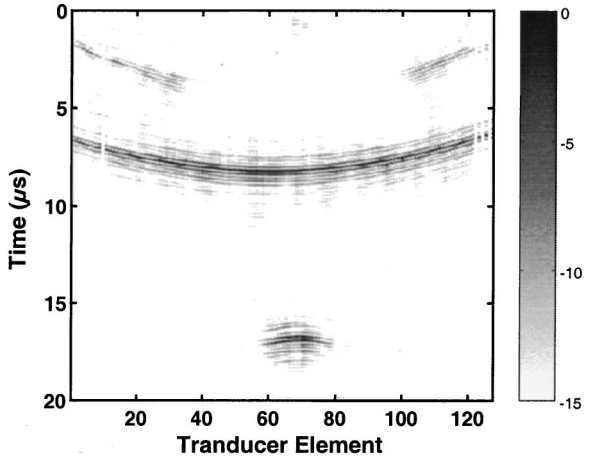


FIG. 7. Emission vector $\{e^{IF}(t)\}$ calculated by the inverse filter. The main front at 12 μs will be modulated at the focus by the two preceding fronts (at 3 and 4 μs) by means of a transverse wave and an intraplate echo, respectively.

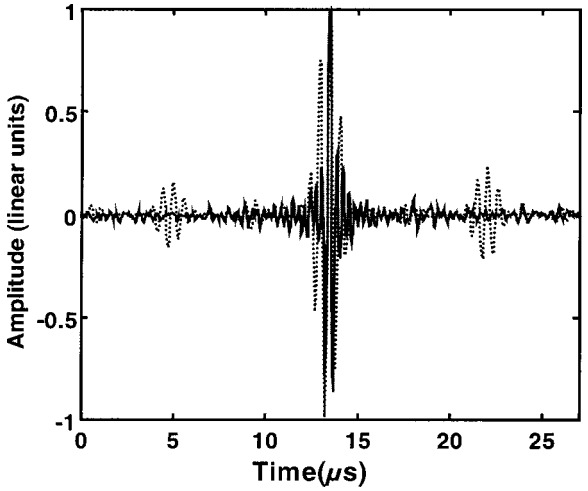


FIG. 8. Signal received at the focal point as a function of time by using time-reversal (dashed line) and the inverse-filter technique (solid line).

sequently, it is possible to favor the use of transverse waves in order to increase the effective aperture. High-incidence angles are thus emitted at time 3 μs on the first element that will be transmitted in the plate by longitudinal–transverse wave conversion. Those signals are emitted 4 μs before the main front, so that the transverse wave generated at the first interface of the plate will give rise on the second interface to a longitudinal wave in water that will reach the focus point on time with the longitudinal waves of the wavefront. Of course, the longitudinal part of the waves sent at time 4 μs will arrive sooner and degrade the focusing. Thus, a slight wavefront can be seen arriving between 5 and 10 μs in Fig. 6, but as the transverse part of the signal is set to focus on the spatio-temporal Dirac function, the longitudinal part is differently refracted by the plate and does not focus: it is diluted with a small amplitude over a large distance. It is interesting to notice that even if the solution mathematically given by the inverse filter may in some cases seem very odd, it always has a physical meaning. Thus, in this case, signals are counterintuitively emitted before the main wavefront, but as we just demonstrated, it actually enhances the focusing.

Finally, to confirm that the plate echoes are well suppressed, it is interesting to look at the signal received at the focal point as a function of time. In Fig. 8, we plot the signal received at the focal point after a time-reversal operation (dashed line) as a function of time. This signal is compared to the one obtained at the focal point by using the inverse filter (solid line).

This signal confirms the suppression of the plate echoes. Moreover, one can also notice that the time compression obtained with the inverse filter is better than the one obtained with time reversal: as predicted in the first theoretical part of the article, the inverse filter enlarges the actual bandwidth of the system. This improves the axial resolution. One can also notice that the frequency component of the inverse filter is higher than that of time reversal, which is completely relevant with the choice of the objective: a spatio-temporal Dirac function. But, the lateral resolution is also improved as the inverse filter compensates for the angle-dependent transmission coefficient by using bulk-shear mode conversion. This is confirmed by the directivity pattern itself given in Fig. 9.

In summary, we have demonstrated that when focusing through a reflective medium, the inverse filter optimizes the set of emission signals on each transducer, in order to fit the objective. Thus, temporal resolution is enhanced not only by suppressing the plate echoes but also by enlarging the effective bandwidth of the transducers. The spatial focusing is also enhanced by compensating the angle-dependent transmission coefficient. This property lends this technique great promise for application in complex geometries, for example trying to image defects behind plates or tubes in nondestructive testing.

III. FOCUSING THROUGH ABSORBING MEDIA

When focusing through a medium with a varying absorption, parts of the wavefront that encounter different absorption can no longer optimally interfere. This leads to an increase in the sidelobes. As absorption breaks the time-reversal invariance of the wave equation, the time-reversal

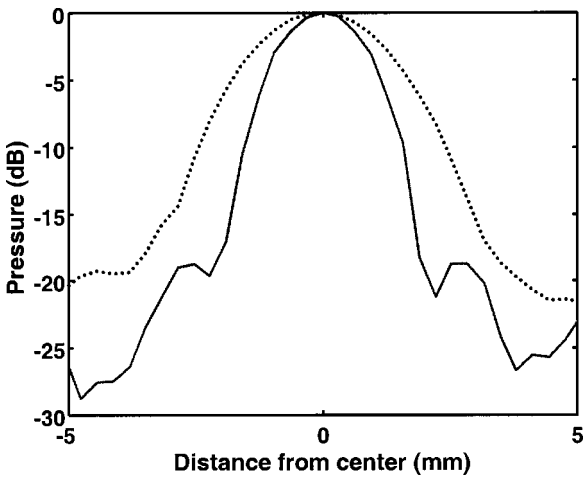


FIG. 9. Directivity patterns (maximum in the temporal domain of the signals recorded at each location) obtained by time reversal (dotted line) and inverse filter through the titanium plate (solid line).

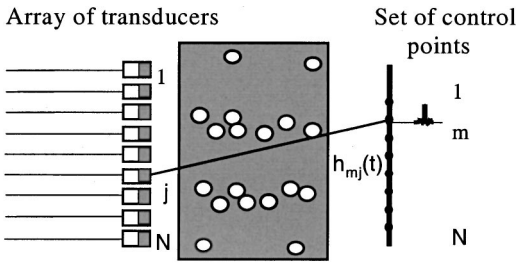


FIG. 10. An UREOL® sample is located between the image plane and the array of 127 transducers.

focusing is also degraded. But, the inverse filter compensates absorption if necessary : if the propagation operator H corresponds to an attenuator, then the inverse operator H^{-1} acts as an amplifier. We demonstrated this by focusing through an UREOL® plate with heterogeneously distributed absorption zones.

A. Experimental setup

A strongly absorbing UREOL® plate (2.5 dB cm^{-1} at 1.5 MHz) is placed between the array of transducers and the array of control points, as presented in Fig. 10. The sample was semirandomly pierced in order to have a varying absorption. As usual, the experiment is immersed in water, so that each hole is not absorbing. Globally, the plate is strongly absorbing in the middle and at both ends.

As UREOL® impedance is very close to that of water, there are no reflections. Moreover, there is no phase distortion because the sound velocities in both media are very close : the experiment models a purely absorbing medium.

B. Experimental results

As previously described, the propagation operator H was experimentally acquired, inverted, and the focusing vector was calculated. Its spatio-temporal singular-value distribution is compared to the one acquired in water in Fig. 11. Here, singular values below the -28-dB threshold are considered as noise and thus are not taken into account in the inversion. The influence of absorption is straightforward. As one can observe, the weight of the singular values of the propagation operator acquired through the UREOL® sample is weaker than those of the propagation operator acquired in water. Only the two or three first singular values have a significant weight in UREOL®, whereas all the physical singular values above the noise have approximately the same weight in water.

Once the focusing vector was calculated, it was emitted on the array of transducers and the pressure field in the focal plane was measured in order to obtain the directivity pattern. This directivity is compared to the one obtained by using a cylindrical law (Fig. 12).

The amplitude modulation of the focusing wavefront induced by the UREOL® sample leads to a high sidelobe level in the directivity pattern obtained with the cylindrical law. But, the inverse filter corrects the aberrations of the absorbing medium, as the sidelobes level decreases significantly. This suggests that the inverse filter is an efficient technique

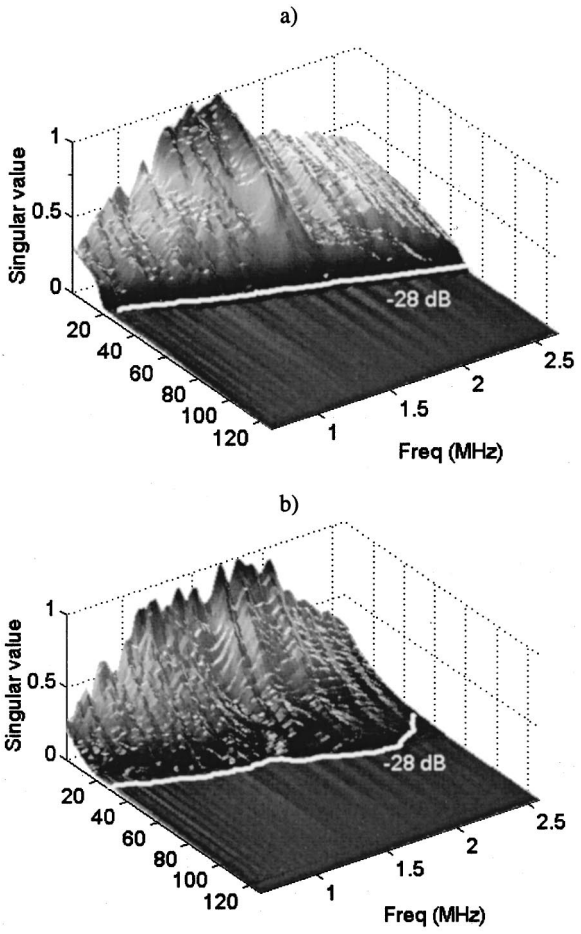


FIG. 11. Spatio-temporal singular-values distribution of the propagation operator $\{h_{mj}(r,t)\}$ acquired (a) in water and (b) through the UREOL® sample.

to focus through complex absorbing media. At the frequencies used in echographic imaging, human tissues are significantly absorbing. So, inverse filtering could improve the resolution and the contrast of images. However, as it is time-consuming to acquire and invert the propagation operator H, at this moment we restricted this technique to a medium whose aberrations don't move. Thus, correcting the aberra-

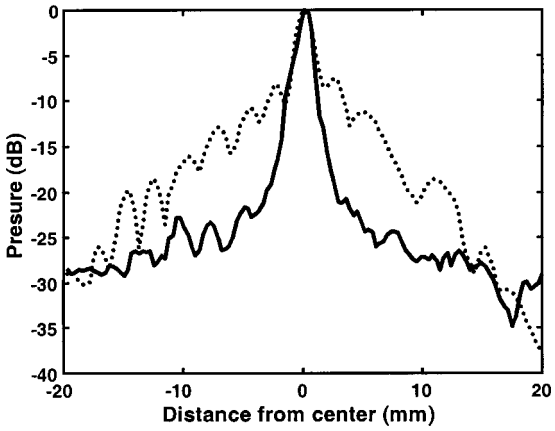


FIG. 12. Directivity patterns obtained with a cylindrical law through the UREOL® sample (dotted line), compared to the one obtained with the inverse filter (solid line).

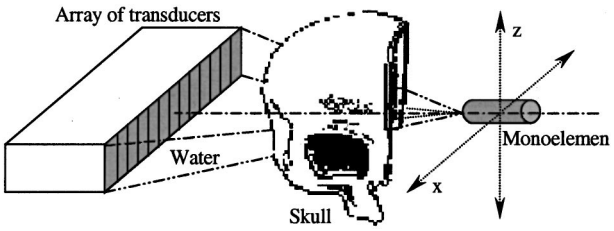


FIG. 13. One half-skull is located between the image plane and the array of 127 transducers.

tions induced by the human skull was an adequate challenge that we will now develop in Sec. IV.

IV. FOCUSING THROUGH AN ABERRATING MEDIUM: THE HUMAN SKULL

The human skull is an interesting medium to test the efficiency of the inverse filter technique as it induces phase aberration and absorption that are both frequency dependent. The sound velocity in human skull[10] varies from 1500 ms^{-1} to 3000 ms^{-1}. The sound absorption equals 1 to 4 $dB\,mm^{-1}\,MHz^{-1}$. The huge aberrations induced by the skull have a strong defocusing effect on ultrasonic beams. That is why brain echography currently remains limited. While it is possible to perform newborn brain imaging by focusing through the fontanel, for adults, brain imaging can only be performed by focusing through the temple. However, as the temple affords a small aperture, the resolution is very low. This is why it would be interesting to correct the aberrations induced by the skull over a large aperture. Solutions were proposed in previous work by combining time-reversal processing, amplitude compensation, and numerical backpropagation.[11] Spatial inverse filtering through the skull was also performed.[12] Although these methods improved the focusing quality in terms of resolution, the sidelobe level remained worse than the one obtained in a homogeneous medium. The goal here was to test if it was physically possible to find a set of signals that would optimally focus through a large area of the skull.

A. Experimental setup

In our experiments, a half-skull is located between the array of transducers and the array of control points, close to the array, as shown in Fig. 13. The whole experiment is immersed in a water tank.

B. Experimental results

The matrix propagator is experimentally acquired and inverted both in water and through the skull. Figure 14 compares the spatio-temporal singular-values distribution of the propagation operator $\{h_{mj}(r,t)\}$ acquired through the skull and in water in the same geometry. We clearly see in Fig. 14(a) by comparison with Fig. 14(b) the frequency dependence of the absorption coefficient in the skull: high frequencies are much more absorbed than the low ones. This will be naturally compensated in the inversion process.

As this experiment consists of the first step toward a possible ultrasonic transcranial imaging, the objective implemented in our codes was not a spatio-temporal Dirac function, but the focusing obtained in water (i.e., the best quality of focusing naturally achievable with our system without excessive degradation of the focal sensitivity). Once this objective was set, we could calculate the set of emitted signals $\{e^{IF}(t)\}$. Those signals are represented in Fig. 15(b) and compared to the set used with a cylindrical focusing [Fig. 15(a)] law. We can see that the set of signals calculated with the inverse filter is a lot more complicated than the cylindrical ones. In order to emit this optimal set of signals, one needs a fully programmable electronic apparatus. In order to check if the focusing is correct, we emitted those signals and scanned the pressure field in the focal plane.

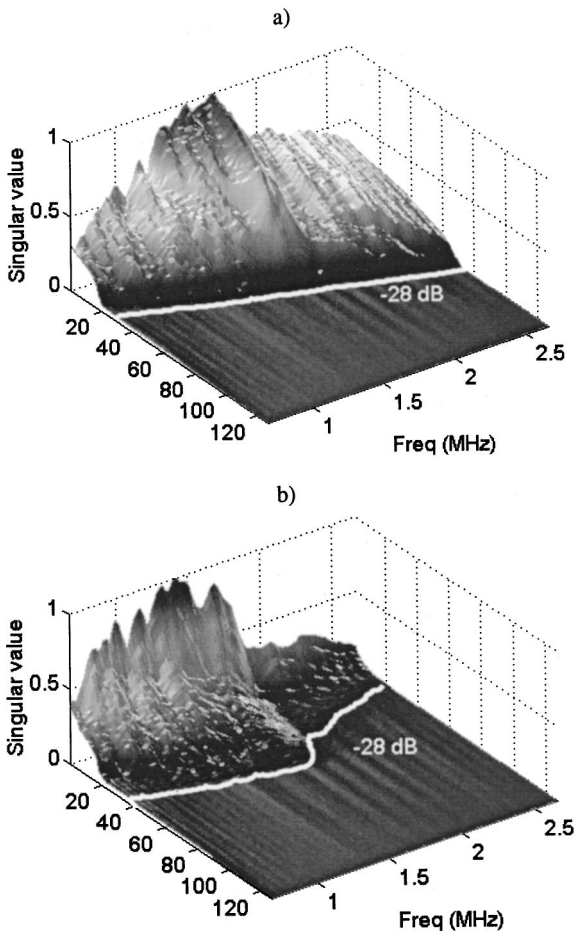


FIG. 14. Spatio-temporal singular-values distribution of the propagation operator $\{h_{mj}(r,t)\}$ acquired (a) in water and (b) through the skull.

Figure 16 compares the spatial focusing obtained through the skull by using a cylindrical law (gray line) to the one obtained in water (solid line), without the skull. The focusing obtained in water is plotted as a reference as it is the optimum focusing we can achieve with our system. We can easily see the dramatic defocusing effects of the skull: the sidelobes are very high (-8 dB), there is a spreading in the main lobe, and the location of the maximum of the pressure amplitude is far from the objective. By using time-reversal (dotted line), phase aberrations induced by the skull are naturally corrected but not attenuation. One can notice that the maximum of pressure amplitude is located at the desired location but the main lobe is rather large and the sidelobes remain high. By comparison we see that the inverse filter

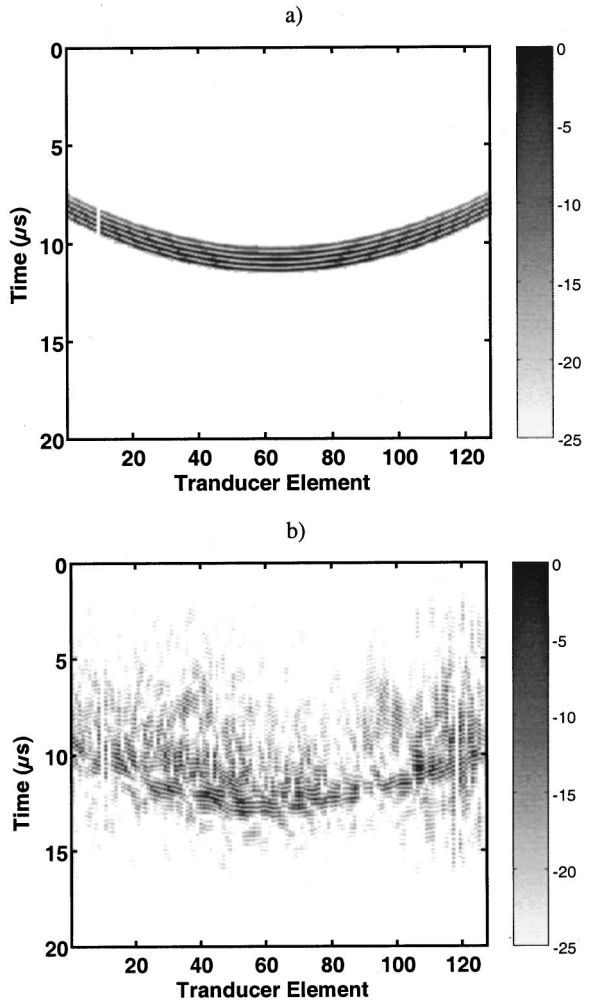


FIG. 15. Emission vector $\{e^{IF}(t)\}$ calculated by the inverse filter (b) compared to the set of emitted signals used with a cylindrical law (a).

(dash-dotted line) perfectly corrects the aberrations induced by the skull: it fits the directivity pattern in water almost perfectly.

Moreover, the temporal compression obtained with this inverse filter (dotted line) is also as good as in water (solid line), as we can see in Fig. 17, when plotting the normalized

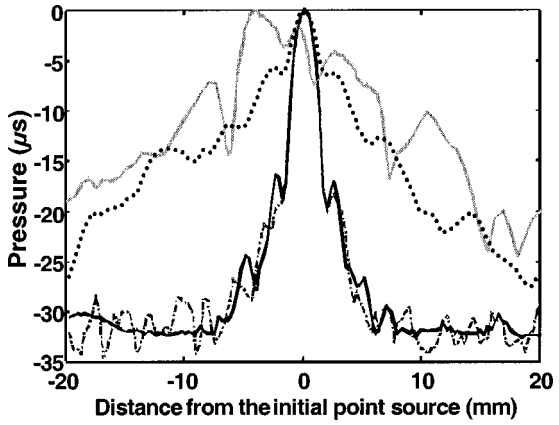


FIG. 16. Directivity patterns obtained by time reversal through pure water (solid line), time reversal (dotted line), inverse filter (dashdotted line), and cylindrical law (gray line).

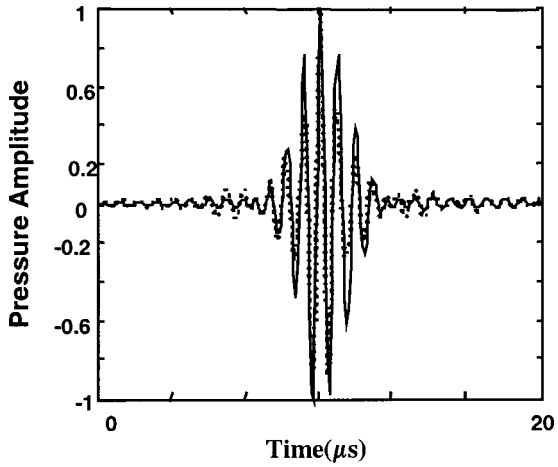


FIG. 17. Normalized pressure amplitude recorded in time at the focal point in water (solid line) and through the skull with inverse-filter focusing (dotted line).

signal recorded at the focal point as a function of time. Again, one can hardly tell the difference between the focusing in water and the inverse filter focusing through the skull. Of course, the pressure amplitude recorded at the focus is smaller when using the inverse filter than when using the cylindrical law because the inverse filter naturally applies a gain on signals that suffered a big loss when passing through the skull. An important part of the emitting energy is dissipated inside the skull. But, as the energy is largely spread in the focal plane when using the cylindrical law, the inverse filter energy is only 5% smaller than the cylindrical one. In terms of energy, time reversal will always reach the optimum as it acts as a spatial and temporal matched filter.

This means that due to the inverse-filter method, it is possible to find a set of signals that optimally focuses through the skull over a large aperture. Thus, it should be possible to perform brain imaging. Of course, this has only been performed through a half-skull. At present, we are developing techniques using two arrays of transducers located at the opposite sides of the entire skull in order to achieve real-time brain imaging, and flow mapping.

V. APPLICATION TO REVERBERANT MEDIA

Trying to achieve nondestructive imaging through complex reverberant structures is very difficult. Indeed, reverberant media strongly diffuse the ultrasonic beams, degrading the focusing. The same problem occurs when studying the sound attenuation in urban geometries.[13] Nevertheless, the inverse filter could be of great interest as it can take advantage of all the reflections to improve the focusing.

A. Experimental setup

We modeled a reverberating environment by placing highly reflective titanium blocks between the array of transducers and the set of control points. The exact model geometry is given in Fig. 18.

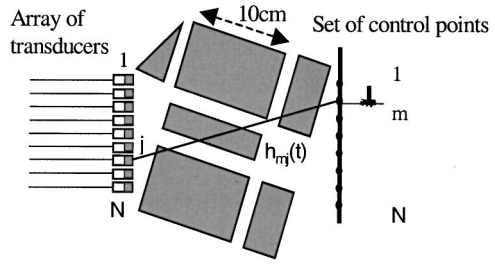


FIG. 18. Reflective metal blocks model a reverberating environment.

B. Experimental results

The matrix propagator is experimentally acquired and inverted both in water and through the model. Figure 19 compares the spatio-temporal singular-values distribution of the propagation operator $\{h_{mj}(r,t)\}$ acquired through the model and in water in the same geometry.

We observe that at each frequency, in a highly reverberant medium [Fig. 19(b)], the singular values strongly differ from the one obtained in a homogeneous medium [Fig. 19(a)]. Thus, whereas a mean number of 20 significant singular values emerges in water, about 40 physical singular values can be seen in the reverberating model. Indeed, thanks to all the reflections, higher spatial frequencies can be generated in the medium. Thus, the complexity of the medium

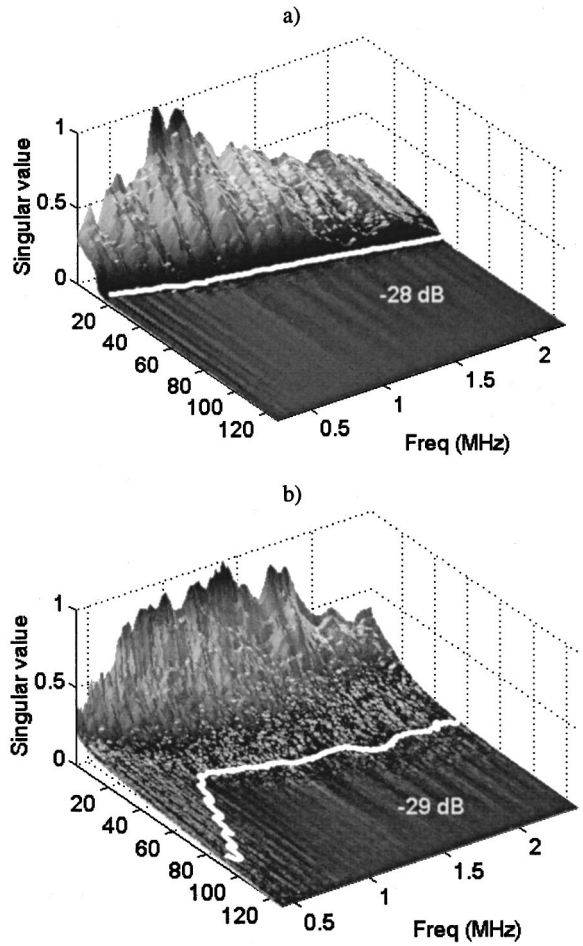


FIG. 19. Spatio-temporal singular-values distribution of the propagation operator $\{h_{mj}(r,t)\}$ acquired (a) in water and (b) through the model.

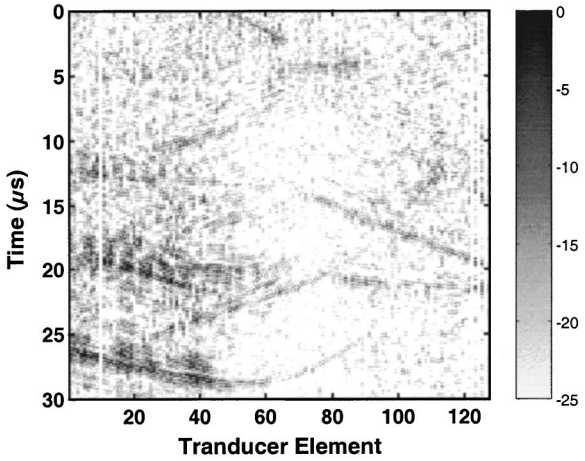


FIG. 20. Emission vector $\{e^{IF}(t)\}$ calculated by the inverse filter. The inverse filter fully takes advantage of all the reflections.

strongly increases the global number of degrees of freedom of the propagation operator. The same phenomenon occurs in waveguides. The singular values corresponding to those waves have a very low amplitude but cannot be considered as noise. The threshold here is thus harder to adjust than in the previous experiments. Once the threshold is adjusted, one can choose an objective [we first tried to focus through the reverberating environment on a spatio-temporal Dirac located at the central position of the control points (N/2nd position)] and determine the set of emitted signals $\{e^{IF}(t)\}$ represented in Fig. 20.

We can see in Fig. 20 that the inverse filter fully takes advantage of all the reflections, so that many wavefronts are successively emitted that will, after propagation and reflections in the medium, constructively interfere at the focal point. We emitted this emission vector $\{e^{IF}(t)\}$ with the array of transducers and tried to compare it to other focusing techniques. The results are presented in Fig. 21. The directivity pattern obtained in water without the model is plotted as a reference (dotted line).

Using a cylindrical law (dash-dotted line) to focus through a reverberating environment is definitely not a solu-

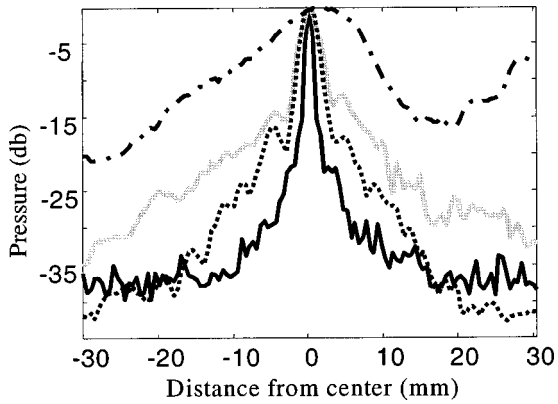


FIG. 21. Directivity patterns obtained through the reverberating model by using a classical cylindrical law (dashed-dotted line), the spatio-temporal inverse filter (solid line), and time reversal (gray line), compared to the one obtained in pure water (dotted line).

tion: the blocks induce a huge spreading in the main lobe. On the contrary, we see that the inverse filter focusing (solid line) is very accurate, as the directivity pattern is even sharper than the one obtained in water: we obtain an interesting hyper-resolution. Such hyper-resolution has been first experimentally shown and explained by Derode *et al.*[9] when using time reversal in a strongly multiple scattering medium. It has been recently theoretically studied, also in the case of a time reversal focusing technique, by Blomgren *et al.*[15] We can also notice that in our geometry, time reversal (Fig. 21, gray line) leads to good focusing, comparable to the one obtained in water. However, in this configuration, time reversal does not involve hyper-resolution, in contrast to inverse filter. On the one hand time reversal does not correct absorption involved in the studied reverberating medium. On the other hand, the chosen configuration is not optimally reverberating: after propagation through this medium, the spatial frequencies components of the time-reversed wavefront do not have the same weight. On the contrary, the spatio-temporal inverse-filter technique corrects the absorption and gives back the same weight to all spatial frequencies components at focus. So, in such a configuration the spatio-temporal inverse filter is the most accurate focusing method.

Moreover, we saw that we could set any objective geometry. For example, to check if the spatio-temporal inverse filter could enable hyper-resolution, we took as the objective a spatio-temporal Dirac function. Then, we wanted to check if it was possible to focus simultaneously on several points in the focal plane. This could allow a very fast scanning of the focal plane, as a given number of focal spots could be examined in the same time. This would greatly enhance the amount of information we can get from the focal plane within one shot. In the companion to this article we gave a definition of the maximum amount of information that can be induced in homogeneous medium by using a limited aperture. To illustrate this, we tried to focus on different locations: one spatio-temporal Dirac function that would focus at time 2 μs at the center, followed by two, four, or eight regularly spaced Dirac functions set to focus at time 7, 12, and 17 μs, respectively. The whole set of emitting signals was calculated and emitted together. The result is given in Fig. 22.

We see that all the spots are well separated. This gives an idea of the minimum size of the cells it might be possible to use in a given geometry. The 15 spots are all illuminated at the same time. Figure 23 shows the set of emitted signals: it is impossible to separate the wavefronts corresponding to each spot.

The inverse-filter technique enables one to send a signal on any of the control points at the same time, as soon as those points are separated by more than the size of a focal spot. This is very interesting, as we saw that the inverse filter takes full advantage of the complex geometry to reduce to the minimum the size of the focal spots. Complex reverberant geometries are thus particularly suited to the inverse-filter technique. Further investigations will be very promising in cavities, waveguides, or multiscattering media.

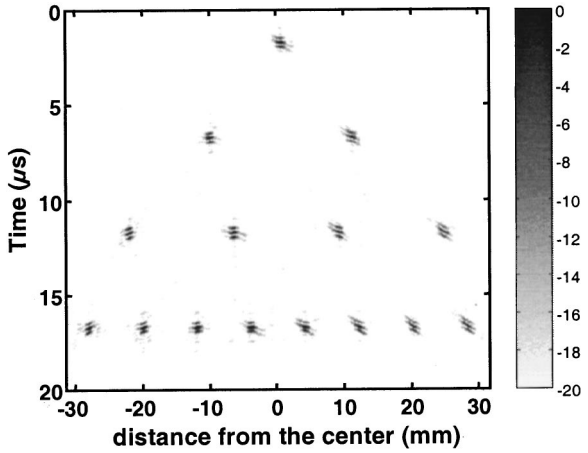


FIG. 22. B scan in the focal plane obtained with the inverse filter. The amplitude of the signals recorded in the focal plane at different distances from the center is represented in dB on a gray scale as a function of time.

VI. CONCLUSION

We have demonstrated experimentally, in various aberrating media, that the spatio-temporal inverse filter presented in Paper I, the companion to this article, provides a method for achieving optimally focused beams through complex heterogeneous media. Comparison was made to time-reversal and cylindrical time-delay-law focusing methods. An experiment where focusing was achieved through a titanium plate proved that the inverse-filter technique was able to craft a compound wavefront capable of suppressing—by destructive interference—echoes due to intraplate reflections. It was also shown that the inverse-filter process can compensate for complicated attenuating media by focusing a beam through an inhomogeneous piece of UREOL®. A successful focusing experiment through a human skull, which combines the complications of attenuation and internal reflections and heterogeneity, demonstrated the applicability of this new technique to extremely complex media. In addition, it has been shown that the spatio-temporal filter can take advantage of complex reverberant media to achieve focal characteristics superior to the homogeneous case. A focusing experiment through a re-

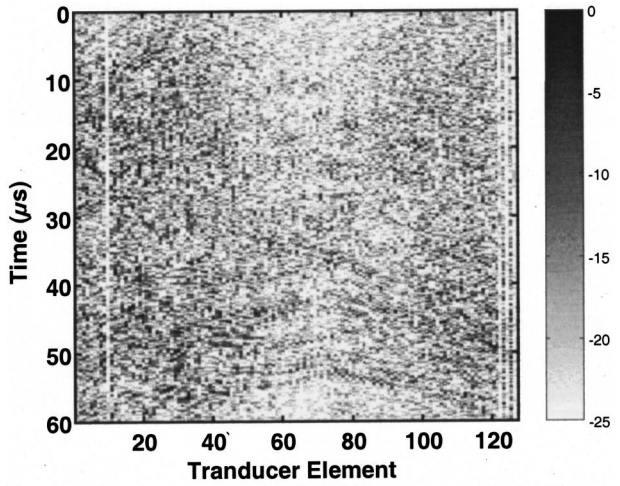


FIG. 23. Emission vector $\{e^{IF}(t)\}$ calculated by the inverse filter. The complex mixing of wavefronts corresponds to the simultaneous focusing on all spots.

verberant environment comprised of metal blocks in water led to ''hyper-resolution,'' in which the focal length is smaller than that achievable by the transducer array in a homogeneous environment. This is possible because the inverse filter takes advantage of the complicated geometry to extend the effective aperture of the array. It was also demonstrated in the context of this experiment that this technique provides a means of focusing on a complicated spatio-temporal pattern.

This technique shows extraordinary promise in several domains. The ability to suppress internal echoes suggests the possibility of precise nondestructive imaging behind complex geometries such as plates, tubes, spars, or ribs. In that case, the inverse-filter technique can be seen as a calibration process allowing to correct aberrations that does not move with regard to the imaging array. Imaging systems which would take advantage of the hyper-resolution obtained in reverberant media are easily envisioned. The demonstration that there exists a set of signals that optimally focuses through a strongly aberrating medium like the skull is an important first step to the development of transcranial ultrasonic imaging. However, in that application, we cannot directly use a secondary set of sources inside the brain. We are currently working on extending the process by creating a set of virtual sources behind the entire skull, in order to achieve noninvasive echography of the brain.

This broadband inverse-filter focusing technique extends well beyond the experiments presented in this paper and we believe it will find myriad applications in cavities, waveguides, multiscattering media, and more generally in various ultrasonic fields including medical imaging, nondestructive testing, and underwater acoustics.

[1] M. O'Donnell and S. W. Flax, ''Phase aberration measurements in medical ultrasound: Human studies,'' Ultrason. Imaging **10**, 1–11 (1988).

[2] M. E. Anderson, M. S. McKeag, and G. E. Trahey, ''The impact of sound speed errors on medical ultrasound imaging,'' J. Acoust. Soc. Am. **107**, 3540–3548 (2000).

[3] M. Moshfeghi and R. C. Waag, ''*In vivo* and *in vitro* ultrasound beam distortion measurements of a large aperture and a conventional aperture focused transducer,'' Ultrasound Med. Biol. **14**, 415–428 (1988).

[4] M. Fink, ''Time reversal of ultrasonic fields. I. Basic principles,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, No. 5, 555–566 (1992).

[5] L. Nock, G. E. Trahey, and S. W. Smith, ''Phase aberration correction in medical ultrasound using speckle brightness as a quality factor,'' J. Acoust. Soc. Am. **85**, 1819–1833 (1989).

[6] W. R. Scott and P. F. Gordon, ''Ultrasonic spectrum analysis for nondestructive testing of layered composite materials,'' J. Acoust. Soc. Am. **62**, 108–116 (1977).

[7] J.-L. Thomas and M. Fink, ''Ultrasonic beam focusing through tissue inhomogeneities with a time reversal mirror: Application to transskull therapy,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, No. 6, 1122–1129 (1996).

[8] D. N. White, J. M. Clark, J. N. Chesebrough, M. N. White, and J. K. Campbell, ''Effect of skull in degrading the display of echoencephalographic B and C scans,'' J. Acoust. Soc. Am. **44**, 1339–1345 (1968).

[9] A. Derode, A. Tourin, and M. Fink, ''Time reversal in multiply scattering media,'' Ultrasonics **36**, 443–447 (1998).

[10] F. J. Fry and J. E. Barger, ''Acoustical properties of the human skull,'' J. Acoust. Soc. Am. **65**, 1576–1590 (1978).

[11] M. Tanter, J.-L. Thomas, and M. Fink, ''Focusing and steering through absorbing and aberating layers: Application to ultrasonic propagation through the skull,'' J. Acoust. Soc. Am. **103**, 2403–2410 (1998).

[12] M. Tanter, J.-L. Thomas, and M. Fink, ''Time reversal and the inverse filter,'' J. Acoust. Soc. Am. **108**, 223–234 (2000).

[13] H. Kuttruff, ''A mathematical model for noise propagation between buildings,'' J. Sound Vib. **85**, 125–128 (1982).

[14] J. F. Aubry, M. Tanter, J. Gerber, J.-L. Thomas, and M. Fink, ''Optimal focusing by spatio-temporal inverse filter. I,'' J. Acoust. Soc. Am. **110**, 37–47 (2001).

[15] P. Blomgren, G. Papanicolaou, and H. Zhao, ''Super-resolution in time-reversal acoustics,'' J. Acoust. Soc. Am. (to be published).

Rayleigh waves on a viscoelastic solid half-space

Maurizio Romeo[a)]
D.I.B.E. Università, via Opera Pia 11/a, 16145 Genova, Italy

(Received 31 July 2000; accepted for publication 17 April 2001)

It is shown that the secular equation for Rayleigh waves propagating on a viscoelastic half-space always admits only one complex root corresponding to a surface wave. This result is proved in those cases in which the displacement field can be obtained analytically, including the isotropic case. The root is obtained in terms of complex integrals extending to the viscoelastic case, a result by Nkemzi [Wave Motion **26**, 199–205 (1997)]. The wave solution is shown to represent an admissible surface wave for any viscoelastic relaxation kernel compatible with thermodynamics. A correspondence is then established between elastic and viscoelastic Rayleigh waves and their propagation properties are pointed out by an approximated analysis of the solution. Illustrative results are given by a numerical evaluation of the complex root for anisotropic viscoelastic half-spaces. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1378347]

PACS numbers: 43.20.Bi, 43.35.Mr, 43.35.Pt [ANN]

I. INTRODUCTION

The theoretical investigation of acoustic surface waves propagating along the boundary of a solid half-space has been developed in order to account for many different physical settings such as layered media, piezoelectric materials, interfaces between different media, and so on. These studies are motivated by a wide variety of applications which run from the detection and analysis of seismic surface waves to the design of electronic devices for processing electroacoustic signals. A well-established theory exists which is mainly based on linear elasticity in continuum mechanics,[1–3] and a number of relevant questions have been successfully answered in connection with anisotropic media about the existence of surface waves and the evaluation of their speed (cf. Ref. 4 and, more recently, Refs. 5 and 6). Recent results also concern porous media,[7,8] prestressed materials,[9] and generalized continua.[10]

Although it is not difficult, in principle, the extension of the theory to viscoelastic solids has been limited to few results.[11–16] This is essentially due to the occurrence of complex valued moduli in the governing equations of the problem, which, in the usual approach, need to be solved numerically to ascertain the existence of Rayleigh waves and to evaluate the wave speed, even in the isotropic half-space.[17] In this case, in fact, it seems that the result of uniqueness for the admissible Rayleigh wave in elasticity has no counterparts in viscoelasticity. Specifically, the analysis of viscoelastic Rayleigh waves performed in Ref. 12 leads to a pair of admissible surface waves and one of them is the viscoelastic counterpart of the classical elastic result. From a general point of view, the additional viscoelastic surface wave is unexpected since a correspondence principle may be applied to the Rayleigh problem,[18] which ensures that the viscoelastic solution can be obtained from the elastic one by simply substituting the elastic moduli with their viscoelastic, complex valued, counterparts in the secular equation. In particular, one would expect that nonadmissible solutions to the elastic secular equation, if any, lead to nonadmissible solutions to the corresponding viscoelastic equation. In the customary approach the determination of admissible Rayleigh waves relies on the solution of an algebraic cubic equation which follows by squaring the secular equation. Spurious solutions are also introduced which must be selected by the requirement of exponential decay of the wave amplitude with the depth in the half-space. In the elastic case it has been shown that only one of the three possible roots represents an effective surface wave,[19] but the same result for viscoelasticity has never been proved.

In a recent paper, Nkemzi[20] has shown that the solution of the secular equation in the elastic isotropic Rayleigh problem is unique and real. As a consequence, only one root of the cubic equation satisfies the original secular equation and the question of the admissibility of the other two solutions simply does not apply. In the present paper Nkemzi's approach is extended to the viscoelastic half-space and it is shown that the secular equation admits only one complex root. This root turns out to represent an effective surface wave for any complex valued viscoelastic moduli. In particular, the proof is concerned with isotropic and anisotropic solids for which the secular equation can be obtained in a closed form.

The viscoelastic model is outlined in Sec. II, where the symmetry properties of the solid half-space and the governing equations for surface wave propagation are given. The secular equation is discussed in detail in Sec. III with the proof of the uniqueness of the viscoelastic surface wave in a anisotropic half-space which possesses at least orthorhombic symmetry. The same problem for isotropic solids is briefly analyzed in Sec. IV and the reduction to Nkemzi's result is performed. In Sec. V it is remarked that the present result agrees with the correspondence principle, which is often invoked in the search of solutions to the boundary value problems in linear viscoelasticity. Singular properties of surface wave propagation are derived by an approximated analysis of the solution to the secular equation in the isotropic case. Viscoelastic surface waves are shown to propagate with

a)Electronic mail: romeo@dibe.unige.it

speed greater than the speed of elastic surface waves and smaller than that of the slowest bulk waves propagating in the same direction. In addition, the penetration depth of viscoelastic surface waves turns out to be greater than that of the corresponding elastic waves. Such results are confirmed by a numerical computation of the more general anisotropic problem, performed in Sec. VI.

II. VISCOELASTIC SURFACE WAVES POLARIZED IN THE SAGITTAL PLANE

A solid half-space is here modeled by means of the well-established linear theory of viscoelasticity. Denoting by $\mathbf{e}(\mathbf{x},t)$ the infinitesimal strain tensor, the following constitutive equation holds for the Cauchy stress $\mathbf{T}$:

$$\mathbf{T}(\mathbf{x},t)=\int_0^\infty \mathbf{G}(\tau)\,\dot{\mathbf{e}}(\mathbf{x},t-\tau)\,\mathrm{d}\tau, \tag{1}$$

where $\mathbf{G}:\mathbb{R}^+\rightarrow \mathrm{Lin(Sym)}$ is the relaxation kernel and where a superimposed dot denotes the material time derivative. Harmonic solutions to the mechanical problem are considered, such that the displacement $\mathbf{u}$ can be written as

$$\mathbf{u}(\mathbf{x},t)=\hat{\mathbf{u}}(\mathbf{x})\exp(-i\omega t), \tag{2}$$

where $\omega\in\mathbb{R}^{++}$ is the angular frequency. From Eq. (1) it follows that

$$\mathbf{T}(\mathbf{x},t;\omega)=\boldsymbol{\mathcal{G}}(\omega)\,\hat{\mathbf{e}}(\mathbf{x})\exp(-i\omega t), \tag{3}$$

where $\hat{\mathbf{e}}(\mathbf{x})=\frac{1}{2}[\nabla\hat{\mathbf{u}}(\mathbf{x})+(\nabla\hat{\mathbf{u}})^T(\mathbf{x})]$ and

$$\boldsymbol{\mathcal{G}}(\omega)=\mathbf{G}(0)+\mathbf{G}_c'(\omega)+i\mathbf{G}_s'(\omega)=:\boldsymbol{\mathcal{G}}^a(\omega)+i\boldsymbol{\mathcal{G}}^b(\omega).$$

The quantities $\mathbf{G}_c'$ and $\mathbf{G}_s'$ represent, respectively, the half-range Fourier cosine and sine transforms of the derivative of $\mathbf{G}(t)$, and $\mathbf{G}(0)$ is the instantaneous elastic modulus. From these definitions it follows that

$$\lim_{\omega\rightarrow\infty}\boldsymbol{\mathcal{G}}(\omega)=\mathbf{G}(0),\quad \lim_{\omega\rightarrow 0}\boldsymbol{\mathcal{G}}(\omega)=\mathbf{G}(\infty). \tag{4}$$

It is customary to assume that the equilibrium elastic modulus $\mathbf{G}(\infty)$ be positive definite. The exploitation of the second law of thermodynamics implies that $\mathbf{G}(0)-\mathbf{G}(\infty)$ be positive semidefinite and $\mathbf{G}_s'(\omega)$ be negative semidefinite for any $\omega\in\mathbb{R}^{++}$.[21] As a consequence, since the entries of $\mathbf{G}_c'(\omega)$ are likely to be much smaller than the corresponding entries of $\mathbf{G}(0)$,[15] the following assumption is made:

$$\boldsymbol{\mathcal{G}}^a(\omega)>0,\quad \boldsymbol{\mathcal{G}}^b(\omega)\leqslant 0,\quad \forall\,\omega\in\mathbb{R}^{++}. \tag{5}$$

Solutions to the dynamical problem, in the form of surface waves, are required to satisfy a stress-free condition at the plane boundary $\mathcal{S}$ of the half-space. Denoting by ρ the mass density and by $\mathbf{n}$ the normal to $\mathcal{S}$ the equation of motion, and the boundary condition read

$$\rho\ddot{\mathbf{u}}=\nabla\cdot\mathbf{T},\quad \mathbf{Tn}|_{\mathcal{S}}=0. \tag{6}$$

In order to obtain an explicit analytical solution to the problem of wave propagation, it is necessary to consider some restrictions on the material symmetries of the half-space. To this end, having introduced Cartesian coordinates x_1,x_2,x_3 such that the half-space be placed at $x_3\geqslant 0$, the displacement $\hat{\mathbf{u}}$ is assumed to propagate as a plane wave along the x_1 direction. As usual, the plane π spanned by x_1,x_3 is referred to as the *sagittal plane*.

Owing to the symmetry of the fourth-rank tensor $\mathbf{G}$, the entries of $\boldsymbol{\mathcal{G}}^a$ and $\boldsymbol{\mathcal{G}}^b$ can be expressed in a six-dimensional notation as

$$\mathcal{G}^{a,b}_{\alpha\beta}=\mathcal{G}^{a,b}_{\beta\alpha},\ \ (\alpha,\beta=1,\ldots,6),$$

where the usual indicial correspondence holds between greek indices and the pairs (ij), i.e., $1:=(11)$, $2:=(22)$, $3:=(33)$, $4:=(23)=(32)$, $5:=(13)=(31)$, $6:=(12)=(21)$. It turns out that if x_k lies along a twofold axis of symmetry or along a normal to a plane of symmetry, then[22]

$$\mathcal{G}^{a,b}_{\alpha\beta}=0,\quad \alpha\in\{(ik),i\neq k\},\quad \beta\neq\alpha. \tag{7}$$

In view of conditions (7) surface acoustic waves propagating along x_1 are necessarily polarized in the sagittal plane if either π is a plane of symmetry or the normal to π is a twofold axis of symmetry.[23] To meet such conditions the solid must have at least a monoclinic symmetry. Then a two-component displacement field is allowed. A further simplification is obtained if also x_1 is a direction normal to a plane of symmetry or lies along a twofold axis of symmetry. This assumption implies the additional restrictions

$$\mathcal{G}_{6\alpha}=\mathcal{G}_{5\alpha}=0,\quad \mathcal{G}_{4\beta}=0,\quad \alpha\neq 5,6,\quad \beta\neq 4,6. \tag{8}$$

The same simplification occurs if the previous properties are satisfied by x_3 instead of x_1. As shown in Ref. 24, crystal classes which satisfy these restrictions belong to orthorhombic, tetragonal, hexagonal, or cubic systems. Substitution of Eqs. (2), (3), and (8) into Eq. (6) yields the following solution (the detailed derivation parallels that of Royer and Dieulesaint;[24] see also Ref. 23):

$$\begin{aligned}
\hat{u}_1(x_1,x_3,\omega)&=A\exp[i\chi(\omega)k(\omega)x_1]\\
&\quad\times\{\exp[-\kappa_+(\omega)k(\omega)x_3]\\
&\quad+p(\omega)\exp[-\kappa_-(\omega)k(\omega)x_3]\},\\
\hat{u}_3(x_1,x_3,\omega)&=A\exp[i\chi(\omega)k(\omega)x_1]\\
&\quad\times\{q_+(\omega)\exp[-\kappa_+(\omega)k(\omega)x_3]\\
&\quad+q_-(\omega)p(\omega)\exp[-\kappa_-(\omega)k(\omega)x_3]\},
\end{aligned} \tag{9}$$

where $\kappa_\pm$ are solutions of the biquadratic equation

$$\begin{aligned}
&\alpha_3\kappa^4+[\alpha_3+1+(\alpha_{13}^2-\alpha_1\alpha_3-1)\chi^2]\kappa^2\\
&\quad+(\alpha_1\chi^2-1)(\chi^2-1)=0,
\end{aligned} \tag{10}$$

with

$$\alpha_1(\omega)=\frac{\mathcal{G}_{11}(\omega)}{\mathcal{G}_{55}(\omega)},\quad \alpha_3(\omega)=\frac{\mathcal{G}_{33}(\omega)}{\mathcal{G}_{55}(\omega)},$$

$$\alpha_{13}(\omega)=1+\frac{\mathcal{G}_{13}(\omega)}{\mathcal{G}_{55}(\omega)}.$$

The quantities $q_\pm$, p, k are complex valued functions of ω and are given by

$$q_{\pm}=i\frac{\alpha_1\chi^2-1-\kappa_{\pm}^2}{\alpha_{13}\kappa_{\pm}^2},\quad p=-\frac{\kappa_{+}-i\chi^2 q_{+}}{\kappa_{-}-i\chi^2 q_{-}},\quad k^2=\frac{\rho\omega^2}{\mathcal{G}_{55}}.$$

Finally, the function $\chi(\omega)$ is required to satisfy the following secular equation:

$$\sqrt{\alpha_3}\sqrt{\alpha_1\chi^2-1}\sqrt{\chi^2-1}=\frac{\alpha_1\chi^2-1}{\alpha\chi^2-1},\tag{11}$$

for all $\omega\in\mathbb{R}^{++}$ where

$$\alpha=\alpha_1-\frac{(\alpha_{13}-1)^2}{\alpha_3}.$$

Equation (11) represents the Rayleigh equation for the problem at hand. Owing to inequalities (5), $\Re\alpha_1>0$ and $\Re\alpha_3>0$, while no restrictions on the sign of the imaginary parts of α_1 and α_3 follow from the second law of thermodynamics. In the case of purely elastic half-spaces the quantities α_1, α_3, α_{13} are real valued and the positive definiteness of $\boldsymbol{\mathcal{G}}$ implies

$$\alpha_1>0,\quad \alpha_3>0,\quad \alpha>0.\tag{12}$$

Really, except for very special cases, the material moduli of known materials are such that $\mathcal{G}_{11}>\mathcal{G}_{55}$ and consequently, $\alpha_1>1$. Moreover, according to the hypotheses on $\boldsymbol{\mathcal{G}}^a(\omega)$, it is also assumed that for the viscoelastic material

$$\Re\alpha_1>1.\tag{13}$$

Equations (9) represent a surface wave if $\hat{u}_1$, $\hat{u}_3$ vanish as $x_3\to\infty$. Accordingly, for any solution of the secular Eq. (11), only those roots of Eq. (10), such that

$$\Re(\kappa_{\pm}k)>0,\tag{14}$$

are admitted as effective surface waves. Without loss of generality the signs of χ and k can be chosen in such a way that $\Re(\chi k)>0$. It follows that with a pertinent choice of the signs of $\kappa_{\pm}$, the requirement (14) can be always satisfied for any solution of Eq. (11) such that $\kappa_{\pm}\neq 0$. Equation (11) admits the solution

$$\chi^2=\frac{1}{\alpha_1}.\tag{15}$$

Its substitution into Eq. (10) yields

$$\kappa_{+}^2=0,\quad \kappa_{-}^2=\frac{1-\alpha_1-\alpha_{13}^2}{\alpha_1\alpha_3}.\tag{16}$$

This solution does not represent a surface wave since one of the inequalities (14) is not satisfied. In fact, substitution of Eqs. (15) and (16) into Eq. (9) yields a superposition of a bulk wave and a surface wave-type component propagating along the sagittal plane. In the elastic case, owing to Eq. (12) and the restriction $\alpha_1>1$, also this last component reduces to a bulk wave.

From the above discussion it follows that all the complex roots of Eq. (11), except for Eq. (15), represent surface waves on a viscoelastic half-space. In the following it will be shown that only one such root exists.

Finally, no conditions have been required about the sign of $\Im(\chi k)$. Although it is customary to choose $\Im(\chi k)>0$, the fact that harmonic waves in viscoelastic media can be treated as inhomogeneous waves, makes such a condition unnecessary.[15]

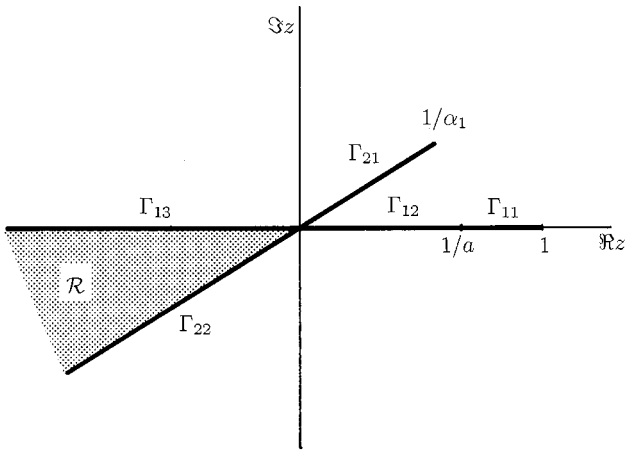


FIG. 1. Arcs Γ_1 and Γ_2 in the complex plane.

III. ANALYSIS OF THE SECULAR EQUATION

The scope of this section is an extension to the viscoelastic anisotropic case, outlined in the previous section, of the analysis of the Rayleigh equation recently proposed by Nkemzi for the isotropic elastic half-space.[20] This approach consists in reducing the solution of Eq. (11) to the search of the complex roots of a suitable polynomial in χ, by the application of the Burniston–Siewert method.[25] Let $z=\chi^2$ and define

$$F_R(z)=\sqrt{\alpha_3}\sqrt{z-1}\sqrt{\alpha_1 z-1}(\alpha z-1)-(\alpha_1 z-1).\tag{17}$$

The zeroes of $F_R(z)$ are the same as the roots of Eq. (11). The function defined in Eq. (17) is analytic in the whole complex plane except for the straight lines

$$\Gamma_1:\ t(\tau)=\tau\ \ \tau\in(-\infty,1],$$

$$\Gamma_2:\ t(\tau)=\tau/\alpha_1\ \ \tau\in(-\infty,1],$$

owing to the discontinuities of the square roots in Eq. (17). Posing $\alpha_1=a+ib$, in view of Eq. (13), it turns out that $1/a<1$. It is convenient to decompose Γ_1 and Γ_2 into the following arcs:

$$\Gamma_{11}=\Gamma_1|_{1/a<\tau\leqslant 1},\quad \Gamma_{12}=\Gamma_1|_{0<\tau\leqslant 1/a},\quad \Gamma_{13}=\Gamma_1|_{\tau\leqslant 0},$$

$$\Gamma_{21}=\Gamma_2|_{0<\tau\leqslant 1},\quad \Gamma_{22}=\Gamma_2|_{\tau\leqslant 0},$$

and denote by $\mathcal{R}$ the plane open sector of $\mathbb{C}$ bounded by Γ_{13} and Γ_{22}. An illustrative picture is given in Fig. 1 where, for definiteness, it has been taken as $b<0$. Now consider the restriction of $F_R(z)$ to $\mathbb{C}\setminus(\mathcal{R}\cup\Gamma_1\cup\Gamma_2)$. Its continuation on $\mathcal{R}$ along Γ_{13} and Γ_{22} is given by

$$F(z)=\begin{cases}F_R(z) & z\in\mathbb{C}\setminus(\mathcal{R}\cup\Gamma_1\cup\Gamma_2)\\ -\sqrt{\alpha_3}\sqrt{z-1}\sqrt{\alpha_1 z-1}(\alpha z-1)-(\alpha_1 z-1) & \\ \qquad z\in\mathcal{R}\cup\Gamma_{13}\cup\Gamma_{22}.\end{cases}\tag{18}$$

On both sides of each arc of $\Gamma=\Gamma_{11}\cup\Gamma_{12}\cup\Gamma_{21}$, the following quantities are defined:

$$F^{\pm}(t)=\lim_{\epsilon\to 0}F(t\pm i\epsilon),\qquad t\in\Gamma, \tag{19}$$

with $\epsilon>0$. Then, denoting by $Q(t)$ the ratio $F^{+}(t)/F^{-}(t)$, for each arc $\Gamma_{11},\Gamma_{12},\Gamma_{21}$, the substitution of Eq. (18) yields, respectively,

$$Q_{11}(t)=\frac{1-iR_1(t)}{1+iR_1(t)},\quad Q_{12}(t)=\frac{1+iR_2(t)}{1-iR_2(t)},$$
$$Q_{21}=\frac{1-iR_2(t)}{1+iR_2(t)}, \tag{20}$$

where

$$R_1(t)=\sqrt{\alpha_3}\frac{\sqrt{1-t}(\alpha t-1)}{\sqrt{\alpha_1 t-1}},$$
$$R_2(t)=i\,\mathrm{sgn}(b)\sqrt{\alpha_3}\frac{\sqrt{1-t}(\alpha t-1)}{\sqrt{1-\alpha_1 t}}. \tag{21}$$

Equations (20) can be also rewritten in the form

$$\ln Q_{11}(t)=-2i\theta_1(t),\quad \ln Q_{12}(t)=-2i\theta_2(t),$$
$$\ln Q_{21}(t)=2i\theta_2(t),$$

where

$$\theta_{1,2}=\arctan(R_{1,2}). \tag{22}$$

The previous positions allow us to show the following result.

Proposition. *$F_R(z)$ always admits only one complex root in addition to $1/\alpha_1$.*

Proof. The proof is organized in two steps. First, it is shown that the analytic continuation of $F(z)$ given by (18) satisfies the statement of the proposition. Second, it is proved that the sought zero does not belong to $\mathcal{R}$.

From the above definitions and results the following facts can be realized. $F(z)$ is analytic on $\mathbb{C}\backslash\Gamma$ and has only one pole (of the second order) at infinity. At the end points of Γ, $t=1/\alpha_1$ and $t=1$, the function $F(z)$ turns out to be analytic with $F(1/\alpha_1)=0, F(1)=1-\alpha_1$. In addition, the limits (19) exist and are different from zero at every internal point of Γ. From Eqs. (20) and (21) it turns out that the limits of $Q(t)$ at the end points of Γ are different from zero. In fact, we have

$$\lim_{\tau\to 1^{-}} Q_{11}(\tau)=1,\qquad \lim_{\tau\to 1^{-}} Q_{21}(\tau/\alpha_1)=-1.$$

In view of the definitions (22), the following function can be introduced:

$$\theta(t)=\begin{cases}\theta_1(t) & t\in\Gamma_{11}\\ \theta_2(t) & t\in\Gamma_{12}\cup\Gamma_{21},\end{cases}$$

which is piecewise continuous and differentiable on $\Gamma\backslash\{1/\alpha_1\}$, and $\mathrm{Lip}_{1/2}$ at $t=1/\alpha_1$.

According to these results, the theory of the Privalov problem can be applied to the function $F(z)$ in $\mathbb{C}\backslash\Gamma$ (see theorem 14.10a in Ref. 26). It follows that a polynomial $P(z)$ exists such that

$$P(z)=F(z)\exp[-g(z)], \tag{23}$$

where

$$g(z)=\frac{1}{\pi}\int_{\Gamma}\frac{\theta(t)}{t-z}\,\mathrm{d}t.$$

The zeroes of $F(z)$ turn out to be the same as those of $P(z)$. Hence $F(z)$ and $\exp[-g(z)]$ are expanded in Laurent series to obtain

$$F(z)=\alpha\sqrt{\alpha_1\alpha_3}\left\{z^2-\left[\frac{1}{2}\left(1+\frac{1}{\alpha_1}\right)+\frac{1}{\alpha}\left(1+\sqrt{\frac{\alpha_1}{\alpha_3}}\right)\right]z\right.$$
$$+\frac{1}{2\alpha}\left(1+\frac{1}{\alpha_1}\right)-\frac{1}{8}\left(1+\frac{1}{\alpha_1^2}\right)$$
$$\left.+\frac{1}{4\alpha_1}\left(1+\frac{4}{\alpha}\sqrt{\frac{\alpha_1}{\alpha_3}}\right)\right\}+O(z^{-1}), \tag{24}$$

$$\exp[-g(z)]=1+\frac{I_0}{z}+\frac{1}{2z^2}(2I_1+I_0^2)+O(z^{-3}),$$

with

$$I_n=\frac{1}{\pi}\int_{\Gamma}t^n\theta(t)\,\mathrm{d}t,\quad n=0,1. \tag{25}$$

Substituting Eqs. (24) into Eq. (23) and accounting for the fact that $P(z)$ is a polynomial, the requirement $P(z)=0$ yields

$$z^2-\left[\frac{1}{2}\left(1+\frac{1}{\alpha_1}\right)+\frac{1}{\alpha}\left(1+\sqrt{\frac{\alpha_1}{\alpha_3}}\right)-I_0\right]z+C=0, \tag{26}$$

where C stands for a complex constant term which depends on α_1,α_3,α. Since one of the roots of Eq. (26) must be equal to $1/\alpha_1$, the remaining root takes the form

$$z_R=\frac{1}{\alpha}\left(1+\sqrt{\frac{\alpha_1}{\alpha_3}}\right)+\frac{1}{2}\left(1-\frac{1}{\alpha_1}\right)-I_0, \tag{27}$$

where, owing to Eqs. (25) and (23),

$$I_0=-\frac{1}{\pi}\int_{1/a}^{1}\theta_1(t)\,\mathrm{d}t-\frac{1}{\pi}\int_{0}^{1/a}\theta_2(t)\,\mathrm{d}t+\frac{1}{\pi a}\int_{0}^{1}\theta_2(t/\alpha_1)\mathrm{d}t. \tag{28}$$

This result completes the first part of the proof.

Now remembering that α_1, α_3, and α are continuous complex functions of ω, Eq. (4) implies

$$\lim_{\omega\to\infty}\alpha_1(\omega)=\alpha_1^0,\quad \lim_{\omega\to\infty}\alpha_3(\omega)=\alpha_3^0,\quad \lim_{\omega\to\infty}\alpha(\omega)=\alpha^0,$$
$$\lim_{\omega\to 0}\alpha_1(\omega)=\alpha_1^\infty,\quad \lim_{\omega\to 0}\alpha_3(\omega)=\alpha_3^\infty,\quad \lim_{\omega\to 0}\alpha(\omega)=\alpha^\infty,$$

where α_1^0, α_3^0, α^0, α_1^∞, α_3^∞, and α^∞ are real quantities. Hence, in both these two limits of ω, the last two integrals in Eq. (28) become equal and opposite so that they cancel out. The corresponding values of z_R are real. For instance,

$$z_{R\infty}=\lim_{\omega\to\infty} z_R$$
$$=\frac{1}{\alpha^0}\left(1+\sqrt{\frac{\alpha_1^0}{\alpha_3^0}}\right)+\frac{1}{2}\left(1-\frac{1}{\alpha_1^0}\right)$$
$$+\frac{1}{\pi}\int_{1/\alpha_1^0}^{1}\arctan\left[\sqrt{\alpha_3^0}\frac{\sqrt{1-t}(\alpha^0 t-1)}{\sqrt{\alpha_1^0 t-1}}\right]\mathrm{d}t. \quad (29)$$

The integral on the right-hand side of Eq. (29) is always greater than $-\pi/2[1-(1/\alpha_1^0)]$. As a consequence, owing to the definition of α,

$$z_R>\frac{1}{\alpha^0}\left(1+\sqrt{\frac{\alpha_1^0}{\alpha_3^0}}\right)>\frac{1}{\alpha^0}>\frac{1}{\alpha_1^0}.$$

Since $F(z)$ does not vanish on $\Gamma\backslash\{1/\alpha_1\}$, the previous inequality implies that, in the limit $\omega\to\infty$,

$$z_{R\infty}>1 \quad (30)$$

A similar result can be worked out in the same way for $\omega\to 0$, i.e., posing $\lim_{\omega\to 0} z_R=z_{R0}$,

$$z_{R0}>1. \quad (31)$$

In view of Eqs. (27) and (28), z_R is a continuous function of the parameters α_1, α_3, α, and, in turn, of ω. Hence z_R moves in the complex plane from z_{R0} to $z_{R\infty}$ continuously as ω ranges from 0 to ∞. Since z_R does not belong to $\Gamma_1\cup\Gamma_2$, owing to Eqs. (30) and (31), it will never reach the region $\mathcal{R}$. Finally, z_R turns out to be always a zero of $F_R(z)$.

In the purely elastic case the parameters α_1,α_3, and α turn out to be real valued and Eq. (27) reduces to

$$z_R=\frac{1}{\alpha}\left(1+\sqrt{\frac{\alpha_1}{\alpha_3}}\right)+\frac{1}{2}\left(1-\frac{1}{\alpha_1}\right)$$
$$+\frac{1}{\pi}\int_{1/\alpha_1}^{1}\arctan\left[\sqrt{\alpha_3}\frac{\sqrt{1-t}(\alpha t-1)}{\sqrt{\alpha_1 t-1}}\right]\mathrm{d}t. \quad (32)$$

This root is real and the result (32) generalizes that of Nkemzi[20] to a anisotropic solid half-space which possesses, at least, orthorhombic symmetry. The following inequality is also satisfied in this case:

$$z_R>1. \quad (33)$$

IV. ISOTROPIC HALF-SPACES

A great part of the existing literature on elastic and viscoelastic Rayleigh waves is concerned with isotropic half-spaces. In this case the secular equation contains only one parameter, σ, which, in the elastic case corresponds to the ratio between the speeds of longitudinal and transverse waves propagating in the bulk of the solid. Denoting by μ and λ the Lamé's moduli of the solid, $\sigma=(2\mu+\lambda)/\mu$ and posing $k^2=\rho\omega^2/\mu$, Rayleigh waves in isotropic media take the form (9) with

$$\kappa_+^2=\chi^2-\frac{1}{\sigma},\quad \kappa_-^2=\chi^2-1,\quad p=-\frac{(\kappa_+\kappa_-)^{1/2}}{\chi},$$
$$q_+=i\frac{\kappa_+}{\chi},\quad q_-=i\frac{\chi}{\kappa_-}. \quad (34)$$

The function χ must satisfy the following secular equation, which is formally the same as the classical one,[27]

$$4\chi^2\sqrt{\sigma\chi^2-1}\sqrt{\chi^2-1}-\sqrt{\sigma}(2\chi^2-1)^2=0. \quad (35)$$

Rayleigh waves are then obtained from Eqs. (34) and (35) provided inequality (14) is satisfied. In particular, there is no isotropic counterpart of the solution (15) to Eq. (11) and the following proposition holds.

Proposition. *Equation (35) always admits only one complex solution.* The proof of this proposition proceeds in a quite similar way as in the anisotropic case and is not given here. As a result, the following root for $\chi^2=:z$ is obtained:

$$z_R=\frac{2+\left(1-\frac{1}{\sigma}\right)^2}{4\left(1-\frac{1}{\sigma}\right)}-I_0, \quad (36)$$

with

$$I_0=-\frac{1}{\pi}\int_{1/a}^{1}\theta_1(t)\,\mathrm{d}t-\frac{1}{\pi}\int_{0}^{1/a}\theta_2(t)\,\mathrm{d}t+\frac{1}{\sigma\pi}\int_{0}^{1}\theta_2(t/\sigma)\,\mathrm{d}t, \quad (37)$$

where $a=\Re(\sigma)$, and

$$\theta_1(t)=\arctan\left[4\frac{t\sqrt{\sigma t-1}\sqrt{1-t}}{\sqrt{\sigma}(2t-1)^2}\right],$$
$$\theta_2(t)=q\arctan\left[4\frac{it\sqrt{1-\sigma t}\sqrt{1-t}}{\sqrt{\sigma}(2t-1)^2}\right],$$

with $q=\mathrm{sign}[\Im(\sigma)]$. In the purely elastic case the last two integrals on the right-hand side of Eq. (37) cancel out and Eq. (36) coincides with the result of Nkemzi,[20]

$$z_R=\frac{2+\left(1-\frac{1}{\sigma}\right)^2}{4\left(1-\frac{1}{\sigma}\right)}+\frac{1}{\pi}\int_{1/\sigma}^{1}\arctan\left[4\frac{t\sqrt{\sigma t-1}\sqrt{1-t}}{\sqrt{\sigma}(2t-1)^2}\right]\mathrm{d}t, \quad (38)$$

which represents the real root of the secular equation for Rayleigh waves on elastic isotropic half-spaces. Inequality (33) also holds in this case.

V. ELASTIC VERSUS VISCOELASTIC SOLUTIONS

The uniqueness of both elastic and viscoelastic solutions to the Rayleigh problem for surface waves is not a surprising result. It can be viewed as a consequence of the principle of correspondence in linear viscoelasticity (see, for example, Refs. 18, 28, and 29). According to this principle, the solution of any boundary value problem with time-independent boundaries in linear viscoelasticity can be obtained from the corresponding problem in linear elasticity. In the present

problem, Eqs. (9)–(11), or their counterparts in the isotropic case, hold for both the elastic and the viscoelastic solution, provided that α_1, α_3, and α_{13} are taken to be the pertinent real or complex valued parameters. Moreover, each solution of the secular Eq. (11) in the viscoelastic case can be put in a one-to-one correspondence with the solution to the elastic problem. However, it is worth remarking that the viscoelastic result (27) cannot be derived from its elastic counterpart (32) by simply substituting the real parameters α_1, α_3, and α with their corresponding viscoelastic complex quantities. This is due to the fact that the derivation of Eq. (27) involves separating the complex parameter α_1 into real and imaginary parts.

The relevant peculiarities of surface wave propagation in viscoelastic solids can be shown directly from the roots of Eq. (11) or (35) under the approximation of small viscous terms. Results based on this approximation are shown in Ref. 12 where the cubic Rayleigh wave has been solved, accounting for first-order imaginary terms added to the real elastic moduli. Solutions of that equation differ from those obtained in the elastic context for the presence of small imaginary terms.

In the present approach, taking advantage of the closed form of the solution to the secular equation, the approximation of small viscosity can be applied directly to the result (27) or (36). In doing so, it can be shown that a first-order correction also arises in the real part of the solution. For the sake of clearness, the isotropic problem is here considered, such that the solution depends on the single parameter σ. It can be written as

$$\sigma = \bar{\sigma}(1 + i\nu\varepsilon), \tag{39}$$

where $\bar{\sigma}$ is a real positive parameter pertaining to the elastic half-space and $0 < \varepsilon \ll 1$. Concerning real materials, in the following it is understood that $\bar{\sigma} \in (2, \infty)$ while no restrictions are imposed on the sign of the real quantity ν. Together with the assumption

$$\mu = \bar{\mu}(1 + i\eta\varepsilon), \quad \eta < 0, \tag{40}$$

position (39) turns out to be equivalent to the hypotheses in Ref. 12 if the value $\bar{\sigma} = 3$ is chosen. Substituting Eq. (39) into Eq. (36) and retaining only terms up to the first order in ε, the following approximation is obtained:

$$z_R = \bar{z}_R + h_2\varepsilon + ih_1\varepsilon, \tag{41}$$

where $\bar{z}_R$ is given by Eq. (36) for $\sigma = \bar{\sigma}$ and

$$h_1 = \frac{\nu}{4\bar{\sigma}}\left[1 - \frac{2}{(1 - 1/\bar{\sigma})^2}\right] + \frac{\nu}{\pi}\int_{1/\bar{\sigma}}^{1} C_1(\bar{\sigma}, t)\, \mathrm{d}t,$$

$$h_2 = -\frac{|\nu|}{\pi}\int_{0}^{1/\bar{\sigma}} C_2(\bar{\sigma}, t)\, \mathrm{d}t.$$

The functions C_1 and C_2 are given by

$$C_1(\bar{\sigma}, t) = \frac{2t\sqrt{t-1}}{\sqrt{\bar{\sigma}}\sqrt{\bar{\sigma}t - 1}(2t-1)^2[1 - A^2(\bar{\sigma}, t)]},$$

$$C_2(\bar{\sigma}, t) = \operatorname{arctanh}[A(\bar{\sigma}, t)] - \frac{2(1-t)(2t-1) + (1 - \bar{\sigma}t)(3 - 2t)}{\sqrt{\bar{\sigma}}(2t-1)^3\sqrt{1-t}\sqrt{1-\bar{\sigma}t}[1 - A^2(\bar{\sigma}, t)]},$$

where

$$A(\bar{\sigma}, t) = 4\frac{t\sqrt{1-\bar{\sigma}t}\sqrt{1-t}}{\sqrt{\bar{\sigma}}(2t-1)^2}.$$

By a numerical integration of C_2, it turns out that h_2 is negative for any admissible value of $\bar{\sigma}$ and this fact bears relevant consequences on the propagation of surface waves. In this respect, owing to Eq. (9), the speed of Rayleigh waves is given by $v_R = \omega/\Re(\sqrt{z_R}k)$. From Eqs. (39), (40), and (41) it follows that

$$\frac{1}{v_R} = \frac{1}{\bar{v}_R}\left(1 + \frac{h_2}{2\bar{z}_R}\varepsilon\right), \tag{42}$$

where $\bar{v}_R = \sqrt{\bar{\mu}/(\rho\bar{z}_R)}$ is the speed of Rayleigh waves in the elastic half-space. Hence, within the approximation of small viscous terms, Rayleigh waves propagate in viscoelastic media faster than in the corresponding elastic ones. Moreover, in view of Eqs. (34),

$$\kappa_+ k = \omega\sqrt{\frac{\rho}{\bar{\mu}}}\bar{\kappa}_+\left\{1 + \frac{1}{2}\frac{\varepsilon}{\bar{\kappa}_+^2}\left[h_2 + i\left(h_1 + \frac{1}{\bar{\sigma}} - \eta\bar{\kappa}_+^2\right)\right]\right\},$$

$$\kappa_- k = \omega\sqrt{\frac{\rho}{\bar{\mu}}}\bar{\kappa}_-\left\{1 + \frac{1}{2}\frac{\varepsilon}{\bar{\kappa}_-^2}[h_2 + i(h_1 - \eta\bar{\kappa}_-^2)]\right\}, \tag{43}$$

where $\bar{\kappa}_+^2 = \bar{z}_R - 1/\bar{\sigma}$, $\bar{\kappa}_-^2 = \bar{z}_R - 1$. As a consequence, the penetration depth of Rayleigh waves in viscoelastic media is greater than that of the corresponding elastic media. Besides, it can be noted that the occurrence of imaginary small terms in Eq. (43) implies that the amplitudes of the displacements u_1 and u_3 slowly oscillate along x_3. Finally, owing to inequality (33) and the result (42), the following relation holds:

$$\bar{v}_R < v_R < \bar{v}_T, \tag{44}$$

where $\bar{v}_T = \sqrt{\bar{\mu}/\rho}$ is the speed of transverse bulk waves propagating along x_1 in the elastic half-space. This last result can be compared with the corresponding one in the numerical analysis of Currie.[14]

Some remarks are in order to compare the present results with those obtained by Currie, Hayes, and O'Leary.[12] In essence, they look for a solution of an approximated cubic Rayleigh equation obtained by a linearization with respect to a small parameter ε, accounting for the imaginary parts of the viscoelastic Lamé's moduli μ and λ. The roots of this equation, evalued up to the first order in ε, are then substituted into the linearized secular equation and selected by the admissibility requirements on the propagation of surface waves. In this way they obtain one ''viscoelastic'' solution, in addition to a ''quasielastic'' one, for special ranges of constitutive parameters. In principle, one can expand the solution of the Rayleigh cubic to the desired order in ε and

look for compatible solutions of the secular equation up to that order in ε. Nevertheless, this does not imply that the expansion converges to a solution of the exact secular equation. Moreover, although the linearization of the equations is performed about the elastic case, where μ and λ are real valued, the viscoelastic solution does not correspond to any admissible surface wave as $\varepsilon \to 0$, nor does it degenerate into a bulk elastic wave in the elastic limit.

Finally, as shown by the present result (41), the small viscosity approximation of the exact solution does not correspond to the solution of the approximated equation in Ref. 12. In particular, a small correction to the real part of the solution appears in Eq. (41), in contrast with the linearized result in Ref. 12, relative to the quasielastic solution. Strictly speaking, although the analysis of Currie, Hayes, and O'Leary is formally correct, their linearized secular equation leads to solutions which, at the same order in ε, are not approximated solutions of the exact equation.

VI. NUMERICAL RESULTS

The main results obtained in the previous section actually hold independently on the approximation of small viscous terms and can be partially extended to the solution of the secular equation in the anisotropic case as given in Sec. III. To show this fact we perform the computation of the root (27) by considering complex valued viscoelastic moduli $\mathcal{G}_{11}$, $\mathcal{G}_{33}$, $\mathcal{G}_{13}$, and $\mathcal{G}_{55}$ in the form

$$\begin{aligned} &\mathcal{G}_{11}=\mathcal{G}^0_{11}(1-i\beta), \quad \mathcal{G}_{33}=\mathcal{G}^0_{33}(1-i\beta), \\ &\mathcal{G}_{13}=\mathcal{G}^0_{13}(1-i\beta), \quad \mathcal{G}_{55}=\mathcal{G}^0_{55}(1-i\gamma). \end{aligned} \tag{45}$$

Here, $\mathcal{G}^0_{11}$, $\mathcal{G}^0_{33}$, $\mathcal{G}^0_{13}$, and $\mathcal{G}^0_{55}$ are real quantities pertaining to a purely elastic anisotropic half-space and β and γ are viscoelastic parameters which, according to the requirements (5), are non-negative. In the present description we avoid any prescription for the underlying dependence of β and γ on the angular frequency ω and assume that $\mathcal{G}^0_{11}$, $\mathcal{G}^0_{33}$, $\mathcal{G}^0_{13}$, and $\mathcal{G}^0_{55}$ are constant. This last hypothesis is justified by the discussion in Sec. II, which precedes inequalities (5), and allows us to express the roots of the secular Eq. (17) in terms of the real parameters β and γ only.

We have numerically computed the root (17) for a half-space of zinc, modeled as in Eq. (45) with

$$\rho = 7135\ \text{kg/m}^3, \quad \mathcal{G}^0_{11}=165\times 10^9\ \text{Pa}, \quad \mathcal{G}^0_{33}=62\times 10^9\ \text{Pa},$$

$$\mathcal{G}^0_{13}=50\times 10^9\ \text{Pa}, \quad \mathcal{G}^0_{55}=39.6\times 10^9\ \text{Pa}$$

(see Ref. 30). The resulting complex z_R has been substituted into the expression of v_R which, in this case, is

$$v_R=\left[\Re\left(\sqrt{\frac{\rho z_R}{\mathcal{G}_{55}}}\right)\right]^{-1}.$$

The Rayleigh wave speed v_R is given in Fig. 2 as a function of the viscoelastic parameter β for three fixed values of γ. The value of v_R corresponding to $\beta=0$ and $\gamma=0$ represents the speed $\overline{v_R}$ of Rayleigh waves in an elastic half-space of zinc. As previously observed in the small viscous approximation, the wave speed is always greater than $\overline{v_R}$, and

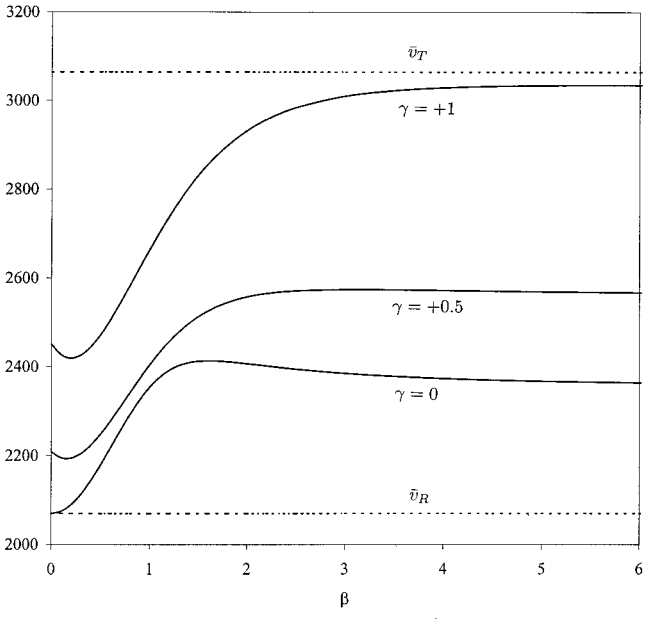


FIG. 2. Speed v_R (m/s) of Rayleigh viscoelastic waves versus the parameter β for three different values of γ.

smaller than the speed $\overline{v_T}$ of transverse bulk waves propagating along x_1. Accordingly, inequality (44) still holds in the present anisotropic case where $\overline{v_T}=\sqrt{\mathcal{G}^0_{66}/\rho}$, with $\mathcal{G}^0_{66}=67\times 10^9$ Pa. Successively, the solution z_R has been substituted into Eq. (10) to evaluate $\kappa_\pm$, accounting for the restriction (14). The results are shown in Figs. 3, 4, 5, and 6 in terms of the nondimensional quantities $\Re(\kappa_\pm k/\overline{k})$ and $\Im(\kappa_\pm k/\overline{k})$, which are plotted versus the viscoelastic parameter β for three fixed values of γ, $(\overline{k}=\omega\sqrt{\rho/\mathcal{G}^0_{55}})$. The values reported in these figures for $\beta=0$ and $\gamma=0$ represent the positive solutions $\overline{\kappa_\pm}$ of the biquadratic Eq. (10) in the elastic case. In particular, Figs. 3 and 4 show that $\Re(\kappa_+ k/\overline{k})$ decreases and

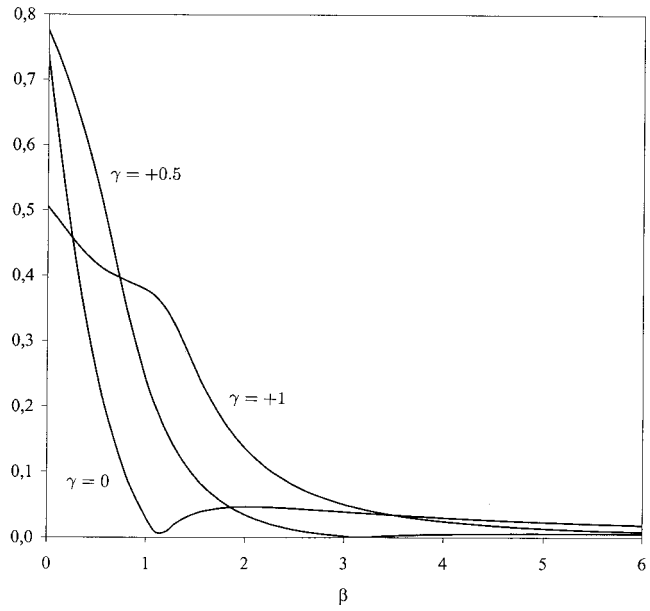


FIG. 3. Real part of $\kappa_+ k/\overline{k}$ vs the parameter β for three different values of γ.

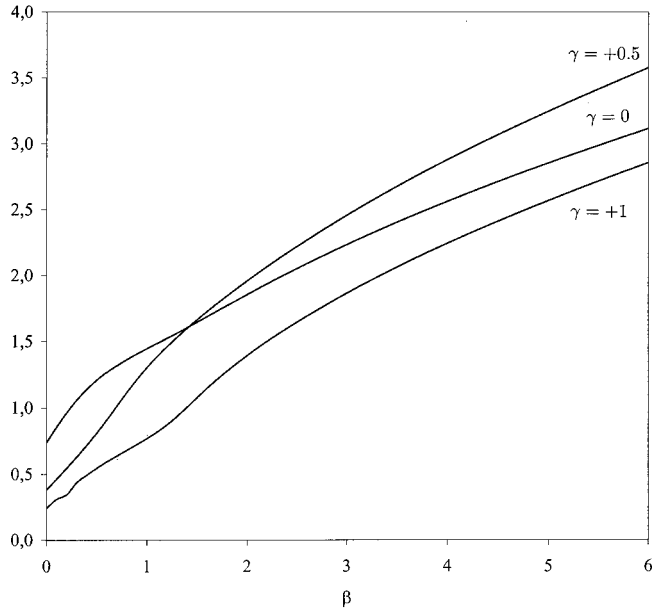


FIG. 4. Real part of $\kappa_- k/\overline{k}$ vs the parameter β for three different values of γ.

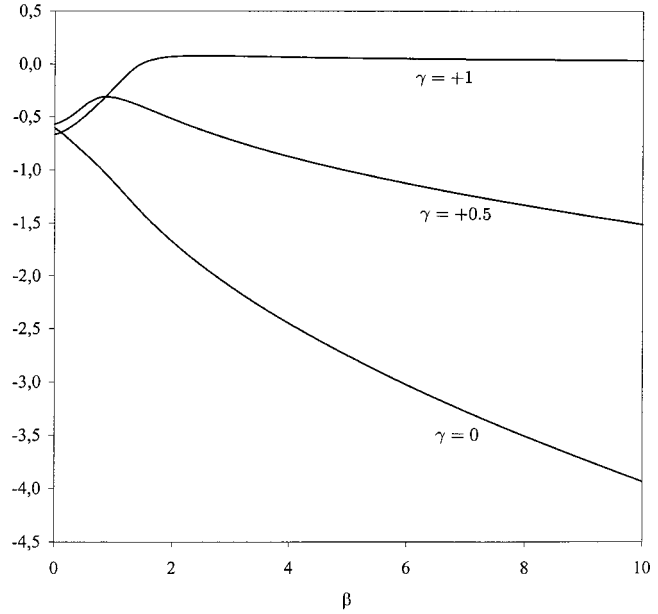


FIG. 6. The imaginary part of $\kappa_- k/\overline{k}$ vs the parameter β for three different values of γ.

$\Re(\kappa_- k/\overline{k})$ increases as β increases. Such a different behavior is peculiar in anisotropic media. In any case, and not only in the small viscosity approximation, the penetration depth, which is proportional to the smallest of the two quantities $[\Re(\kappa_\pm k/k)]^{-1}$, turns out to be enhanced by the presence of viscoelasticity. Finally, nonzero values of $\Im(\kappa_\pm k/\overline{k})$ denote a wave amplitude oscillating along x_3.

VII. CONCLUDING REMARKS

The main theoretical result of the present analysis on surface waves is the proof of uniqueness for the solution to the Rayleigh's wave problem in viscoelastic half-spaces. We have taken advantage of the Nkemzi's result for the same problem in the isotropic elastic case,[20] which independent of the question about its practical utility in determining the Rayleigh wave speed (see Ref. 31 and the results in Ref. 32) represents an effective tool in order to bypass the problem of spurious solutions to the cubic Rayleigh equation. The analogous approach in viscoelasticity presented here rules out some previous results[12] where a viscoelastic surface wave appeared in addition to the so-called quasielastic wave within the small viscosity approximation. On the other hand, the content of the proposition in Sec. III restores the full validity of the correspondence principle in connection with surface wave problems.[18]

It is shown that, within the approximation of small viscous terms in viscoelastic isotropic half-spaces the wave speed and the penetration depth of Rayleigh surface waves are always greater than in the elastic case. Numerical results have pointed out that the same behavior occurs in the more general anisotropic case considered in the present investigation.

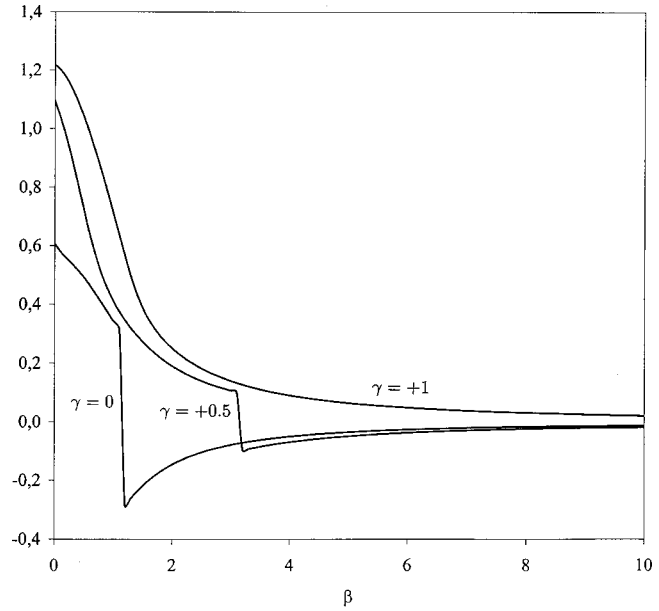


FIG. 5. Imaginary part of $\kappa_+ k/\overline{k}$ vs the parameter β for three different values of γ.

ACKNOWLEDGMENTS

The research leading to this work has been performed under the auspices of GNFM and supported by the MURST–COFIN.98 Project, Modelli Matematici per la Scienza dei Materiali.

[1] J. D. Achenbach, *Wave Propagation in Elastic Solids* (Elsevier, New York, 1984).
[2] I. A. Viktorov, *Rayleigh and Lamb Waves* (Plenum, New York, 1967).
[3] L. Brekhovskikh and V. Goncharov, *Mechanics of Continua and Wave Dynamics* (Springer, Berlin, 1982).
[4] P. Chadwick and G. Smith, ''Foundation of the theory of surface waves in anisotropic elastic media,'' Adv. Appl. Mech. **17**, 303–376 (1977).
[5] D. M. Barnett, P. Chadwick, and J. Lothe, ''The behaviour of elastic surface waves polarized in a plane of material symmetry,'' Proc. R. Soc. London, Ser. A **433**, 699–710 (1991).
[6] A. N. Darinskii, V. I. Alshits, J. Lothe, V. N. Lyubimov, and A. L. Shu-

valov, ''An existence criterion for the branch of two-component surface waves in anisotropic elastic media,'' Wave Motion **28**, 241–257 (1998).
[7] A. Kaptsov and S. Kuznetsov, ''Rayleigh waves in anisotropic porous media,'' J. Acoust. Soc. Am. **106**, 2232–2233 (1999).
[8] W. Lauriks, L. Kelders, and J. F. Allard, ''Surface wave above thin porous layers and periodic structures,'' J. Acoust. Soc. Am. **105**, 1385 (1999).
[9] J. F. Chai and T. T. Wu, ''Propagation of surface waves in prestressed piezoelectric materials,'' J. Acoust. Soc. Am. **100**, 2112–2122 (1996).
[10] I. Vardoulakis and H. G. Georgiadis, ''SH surface waves in a homogeneous gradient-elastic half-space with surface energy,'' J. Elast. **47**, 147–165 (1997).
[11] J. G. Sholte, ''On Rayleigh waves in visco-elastic media,'' Physica (Utrecht) **13**, 245–250 (1947).
[12] P. K. Currie, M. A. Hayes, and P. M. O'Leary, ''Viscoelastic Rayleigh waves,'' Q. Appl. Math. **35**, 35–53 (1977).
[13] P. K. Currie and P. M. O'Leary, ''Viscoelastic Rayleigh waves II,'' Q. Appl. Math. **35**, 445–454 (1978).
[14] P. K. Currie, ''Viscoelastic surface waves on a standard linear solid,'' Q. Appl. Math. **37**, 332–336 (1979).
[15] G. Caviglia and A. Morro, *Inhomogeneous Waves in Solids and Fluids* (World Scientific, Singapore, 1992).
[16] G. Caviglia and A. Morro, ''Inhomogeneous waves in transversely isotropic solids,'' Q. J. Mech. Appl. Math. **49**, 451–469 (1996).
[17] G. Caviglia, A. Morro, and E. Pagani, ''Surface waves on a solid half-space,'' J. Acoust. Soc. Am. **86**, 2456–2459 (1989).
[18] D. R. Bland, *The Theory of Linear Viscoelasticity* (Pergamon, Oxford, 1960).
[19] M. A. Hayes and R. S. Rivlin, ''A note on the secular equation for Rayleigh waves,'' Z. Angew. Math. Phys. **13**, 80–83 (1962).
[20] D. Nkemzi, ''A new formula for the velocity of Rayleigh waves,'' Wave Motion **26**, 199–205 (1997).
[21] M. Fabrizio and A. Morro, ''Viscoelastic relaxation functions compatible with thermodynamics,'' J. Elast. **19**, 63–75 (1988).
[22] M. Romeo, ''Specific directions in viscoelastic crystals,'' Q. J. Mech. Appl. Math. **48**, 77–87 (1995).
[23] A. A. Maradudin and G. J. Stegeman, ''Surface acoustic waves,'' in *Surface Phonons*, edited by W. Kress and F. W. de Wette, Springer Series in Surface Sciences, Vol. 21 (Springer, New York, 1991), pp. 5–35.
[24] D. Royer and E. Dieulesaint, ''Rayleigh wave velocity and displacement in orthorhombic, tetragonal, hexagonal, and cubic crystals,'' J. Acoust. Soc. Am. **76**, 1438–1444 (1984).
[25] E. E. Burniston and C. E. Siewert, ''The use of Riemann problems in solving a class of trascendental equations,'' Proc. Cambridge Philos. Soc. **73**, 111–118 (1973).
[26] P. Henrici, *Computational Complex Analysis* (Wiley, New York, 1986), Vol. III.
[27] Lord Rayleigh, Proc. London Math. Soc. **17**, 4–11 (1885).
[28] M. A. Biot, *Mechanics of Incremental Deformations* (Wiley, New York, 1965).
[29] A. C. Eringen, *Mechanics of Continua* (Krieger, New York, 1980).
[30] R. F. S. Hearmon, *Numerical Data and Functional Relationship in Science and Technology*, Landolt-Börnstein, New Series, Group 3 (Springer, New York, 1969), Vol. I.
[31] P. G. Malischewsky, Comment to ''A new formula for the velocity of Rayleigh waves'' by D. Nkemzi [Wave Motion **26** (1997), 199–205], Wave Motion **31**, 93–96 (2000).
[32] M. Rahman and J. R. Barber, ''Exact expressions for the roots of the secular equation for Rayleigh waves,'' J. Appl. Mech. **62**, 250–252 (1995).

Semi-analytical elastic wave-field modeling applied to arbitrarily oriented orthotropic media

Martin Spies[a)]
Fraunhofer-Institut Zerstörungsfreie Prüfverfahren (IZFP), Universität, Gebäude 37, 66123 Saarbrücken, Germany

(Received 26 July 2000; revised 19 April 2001; accepted 23 April 2001)

Three-dimensional elastic wave-field calculation is addressed for the case of orthotropic materials with arbitrary spatial orientation. Based on a mathematical formulation involving Green's dyadic displacement tensor function, appropriate evaluation yields a representation of the displacement vector of transducer wave fields in anisotropic media which is convenient for effective numerical computation. With respect to bulk wave propagation, the numerical evaluation of Green's dyadic function is circumvented by applying a reciprocity-based approach, which is valid in the (point source) far field. The presented formulation involves characteristic quantities obtained from plane-wave theory and appears as a point-source superposition representation including the respective point-source directivities. Thus, it is in the same form as a corresponding formulation for scattered elastic wave fields presented previously [M. Spies, J. Acoust. Soc. Am. **107**, 2755–2759 (2000)]. Focusing on orthotropic composite materials, numerical evaluation examples are presented for ultrasonic tranducer field patterns and time-dependent radio frequency (rf)-pulse propagation.
 [DOI: 10.1121/1.1380440]

PACS numbers: 43.20.Bi, 43.35.Cg, 43.35.Zc [DEC]

I. INTRODUCTION

Composite materials have gained a considerable importance, being widely applied, e.g., in aerospace industries as unidirectional, layered, or woven structures. Through their complex build-up these materials exhibit anisotropic elastic behavior. Thus, effects like beam splitting, beam distortion, and deviation between wave propagation and energy flow raise considerable difficulties for ultrasonic nondestructive testing techniques. In modeling the interaction of elastic waves with such media, a simple tool of assisting analysis is available. In this respect, simulation and optimization allow for a reduction of experimental work and an increase in reliability of applied testing procedures. High interest has therefore been dedicated to anisotropic wave propagation in past years. Respective studies consider ray tracing, plane-wave theory, or the propagation of pulses in these media (see, e.g., Refs. 1–5). Rose *et al.*[6] have presented a numerical integration model, which includes the consideration of ultrasonic pulse rf signals, using scalar Green's functions as point sources, derived for transversely isotropic media exhibiting mild anisotropy. Newberry and Thompson[7] have introduced a model for anisotropic wave-field calculation employing Gauss–Hermite beams, while a pure Gaussian beam technique for transversely isotropic (TI) media has been presented recently.[8] Also, for such TI media with arbitrary orientation, the modeling of ultrasonic transducer radiation has been addressed in Ref. 9 by using a point-source superposition technique. Recently, Rudolph[10] has also used a Gaussian beam model to treat propagation into a general anisotropic medium. Guo and Achenbach[11] have applied the boundary-element method to calculate the field radiated by a circular transducer into a TI half-space, exploiting simple forms of Green's function for anisotropic media derived by Wang and Achenbach.[12,13] Further, numerical methods have been applied, e.g., by Lord *et al.*,[14] Fellinger *et al.*,[15] and Boström,[16] where emphasis is on problems related to nondestructive testing.

In this article, materials exhibiting orthotropic (OT) elastic symmetry—which is generally the case for composite materials—are addressed. A representation for transducer-field modeling is obtained in Sec. II, which uses fundamental plane wave characteristics and appears as a point-source superposition formulation including the respective point-source directivities. Section III presents the plane-wave relationships, where the orthotropic material's spatial orientation appears as an additional parameter. Since for complex-shaped components the material's natural symmetry planes are in general not identical to the component's surfaces, a respective transformation has been applied to yield a compact elastic tensor representation for such configurations. In Sec. IV, numerical evaluation is presented for a layered composite material, covering point-source directivities and transducer-field patterns for various material orientations. The propagation of rf pulses is also addressed. Generally, two kinds of notations can be used: index notation of tensor analysis and direct notation of vectors and matrices. Throughout the article direct notation is used, since it allows the expression of physical quantities by single letters instead of components, and which reflects their nature of coordinate independence. Thus, the results presented are independent of a special choice of coordinate system.

a)Electronic mail: spies@izfp.fhg.de

 0001-4966/2001/110(1)/68/12/$18.00

II. FORMULATION OF THE TRANSDUCER DISPLACEMENT FIELD

The basis for the elastodynamic beam-field calculation procedure applied in this article is the mathematical formulation of Huygens' principle given by Pao and Varatharajulu.[17] A detailed analysis of such representation theorems can be found, e.g., in Ref. 18. For wave radiation by an isolated vibrating body or a fixed surface enclosing a source, where each point on the surface S of the body vibrates with the same angular frequency ω, the displacement vector outside the surface S can be written as

$$\underline{\mathbf{u}}(\underline{\mathbf{R}},\omega)=\int\int_S\{\underline{\mathbf{u}}(\underline{\mathbf{R}}',\omega)\cdot[\underline{\mathbf{n}}\cdot\underline{\underline{\underline{\Sigma}}}(\underline{\mathbf{R}}-\underline{\mathbf{R}}',\omega)] -[\underline{\mathbf{n}}\cdot\underline{\underline{\mathbf{T}}}(\underline{\mathbf{R}}',\omega)]\cdot\underline{\underline{\mathbf{G}}}(\underline{\mathbf{R}}-\underline{\mathbf{R}}',\omega)\}\mathrm{d}S', \tag{1}$$

where—with $\underline{\underline{\mathbf{T}}}$ being the stress tensor—the traction $\underline{\mathbf{n}}\cdot\underline{\underline{\mathbf{T}}}$ and the displacement $\underline{\mathbf{u}}$ at this surface act as sources of the generated wave field. $\underline{\underline{\mathbf{G}}}$ and $\underline{\underline{\underline{\Sigma}}}$ are Green's dyadic and triadic functions, whose components represent the displacement and stress field, respectively, at position $\underline{\mathbf{R}}$ generated by three mutually perpendicular (point) forces acting at $\underline{\mathbf{R}}'$ on surface S. In Eq. (1), the usually assumed $(\underline{\mathbf{R}},\underline{\mathbf{R}}')$ dependence of the Green's functions has been replaced by the $(\underline{\mathbf{R}}-\underline{\mathbf{R}}')$ dependence, since only translation invariant media are of interest in this article.

For evaluating transducer radiation, S is assumed to lie in the $x-y$ plane of a Cartesian coordinate system, i.e., $\underline{\mathbf{n}}=\underline{\mathbf{e}}_z$. In selecting the Green's tensor functions entering in Eq. (1) one has two options. The first one is to choose the free-space functions as has been done, e.g., by Guo and Achenbach,[11] where, however, $\underline{\mathbf{u}}(\underline{\mathbf{R}}',\omega)$ is an unknown function which has to be determined. Here, the Green's functions for the elastic half-space are chosen, where—considering the surface to be stress-free—the triadic stress-tensor function accordingly fulfills the boundary condition that

$$\underline{\mathbf{e}}_z\cdot\underline{\underline{\underline{\Sigma}}}^{\mathrm{half}}\big|_{z=0}=\underline{\underline{\mathbf{0}}}, \tag{2}$$

so that

$$\underline{\mathbf{u}}(\underline{\mathbf{R}},\omega)=-\int\int_S[\underline{\mathbf{e}}_z\cdot\underline{\underline{\mathbf{T}}}(\underline{\mathbf{R}}',\omega)]\cdot\underline{\underline{\mathbf{G}}}^{\mathrm{half}}(\underline{\mathbf{R}}-\underline{\mathbf{R}}',\omega)\mathrm{d}S' \tag{3}$$

follows, where the dyadic Green's function for the half-space has accordingly been introduced.

A. Far-field representation of Green's dyadic function

From Ref. 19, where the far-field radiation of vibrating point sources in anisotropic media is considered, a far-field expression for $\underline{\underline{\mathbf{G}}}^{\mathrm{half}}$ accounting for the bulk wave contributions can be inferred according to

$$\underline{\underline{\mathbf{G}}}^{\mathrm{half}}_{\mathrm{far}}(\underline{\mathbf{R}}-\underline{\mathbf{R}}',\omega)=\sum_\alpha\ \underline{\mathbf{g}}_\alpha(\hat{\underline{\mathbf{K}}}(\widehat{\underline{\mathbf{R}}-\underline{\mathbf{R}}'}))\hat{\underline{\mathbf{u}}}_\alpha(\hat{\underline{\mathbf{K}}}(\widehat{\underline{\mathbf{R}}-\underline{\mathbf{R}}'})) \times\frac{e^{j\underline{\mathbf{K}}_\alpha(\widehat{\underline{\mathbf{R}}-\underline{\mathbf{R}}'})\cdot(\underline{\mathbf{R}}-\underline{\mathbf{R}}')}}{4\pi|\underline{\mathbf{R}}-\underline{\mathbf{R}}'|}, \tag{4}$$

where $\hat{\underline{\mathbf{u}}}_\alpha$ and $\underline{\mathbf{K}}_\alpha=K_\alpha\hat{\underline{\mathbf{K}}}$ designate the (plane-wave) polarization vector and the wave vector of wave-type α, respectively. The Cartesian components of vector $\underline{\mathbf{g}}_\alpha$ are the directivities for transversely (in x- or y-direction) and normally (in z direction) acting point sources on the stress-free surface. $\underline{\mathbf{g}}_\alpha$ and $\hat{\underline{\mathbf{u}}}_\alpha$—as well as the group velocity vector $\underline{\mathbf{c}}_\alpha$—are functions of the wave propagation direction $\hat{\underline{\mathbf{K}}}$ that produces an energy contribution along the spatial direction $(\widehat{\underline{\mathbf{R}}-\underline{\mathbf{R}}'})$. Generally, the $\hat{\underline{\mathbf{K}}}(\widehat{\underline{\mathbf{R}}-\underline{\mathbf{R}}'})$-relationship has to be evaluated numerically, only in special cases are exact[20] or approximate[21] analytical representations available. An efficient numerical evaluation scheme for this relation is described below. Using Eq. (4) and defining the surface traction

$$\underline{\mathbf{t}}(\underline{\mathbf{R}}',\omega)\equiv\underline{\mathbf{e}}_z\cdot\underline{\underline{\mathbf{T}}}(\underline{\mathbf{R}}',\omega), \tag{5}$$

Eq. (3) finally leads to

$$\underline{\mathbf{u}}(\underline{\mathbf{R}},\omega)\cong-\int\int_S\sum_\alpha\ [\underline{\mathbf{t}}(\underline{\mathbf{R}}',\omega)\cdot\underline{\mathbf{g}}_\alpha(\hat{\underline{\mathbf{K}}}(\widehat{\underline{\mathbf{R}}-\underline{\mathbf{R}}'}))] \times\hat{\underline{\mathbf{u}}}_\alpha(\hat{\underline{\mathbf{K}}}(\widehat{\underline{\mathbf{R}}-\underline{\mathbf{R}}'}))\frac{e^{j\omega|\underline{\mathbf{R}}-\underline{\mathbf{R}}'|/c_\alpha(\widehat{\underline{\mathbf{R}}-\underline{\mathbf{R}}'})}}{4\pi|\underline{\mathbf{R}}-\underline{\mathbf{R}}'|}\mathrm{d}S', \tag{6}$$

which is valid in the far field of the point source. Here, the relationships $\underline{\mathbf{K}}_\alpha\cdot\underline{\mathbf{c}}_\alpha=\omega$ and $\underline{\mathbf{K}}_\alpha\cdot(\underline{\mathbf{R}}-\underline{\mathbf{R}}')=\underline{\mathbf{K}}_\alpha\cdot\underline{\mathbf{c}}_\alpha|\underline{\mathbf{R}}-\underline{\mathbf{R}}'|/c_\alpha=\omega|\underline{\mathbf{R}}-\underline{\mathbf{R}}'|/c_\alpha$ have been exploited, where c_α is the modulus of the group velocity vector. Since the (energy) contribution to a wave field at an observation point $(\widehat{\underline{\mathbf{R}}-\underline{\mathbf{R}}'})$ is characterized by group velocity, it follows for the respective unit vectors that $(\widehat{\underline{\mathbf{R}}-\underline{\mathbf{R}}'})=\hat{\mathbf{c}}$. Equation (6) gives a general formulation for transducer generated bulk wave fields which can be applied to any anisotropic medium. However, the determination of $\underline{\mathbf{g}}_\alpha$ using Green's function is intricate and would require considerable numerical efforts. Therefore, use is made of another way of determining these directivities, which is based on the reciprocity theorem[22] and which has been presented by Wu *et al.*[23] for general anisotropic media. A brief review of this method is given in Appendix A.

B. Computational scheme

Assuming a traction whose magnitude is zero outside the transducer aperture and unity within, the integration in Eq. (6) has to be performed over the transducer aperture. Applying numerical integration on the basis of an equally spaced rectangular grid, where $\underline{\mathbf{R}}_m$ designates the position of the mth grid point, the displacement vector describing the α wave field under concern follows accordingly as

$$\underline{\mathbf{u}}_\alpha(\underline{\mathbf{R}},\omega)\cong\sum_m\ \{[\underline{\mathbf{e}}_i\cdot\underline{\mathbf{g}}_\alpha(\hat{\underline{\mathbf{K}}}(\widehat{\underline{\mathbf{R}}^{\boldsymbol{\Delta}}_{\mathbf{m}}}))]\hat{\underline{\mathbf{u}}}_\alpha(\hat{\underline{\mathbf{K}}}(\widehat{\underline{\mathbf{R}}^{\boldsymbol{\Delta}}_{\mathbf{m}}})) \cdot\exp[-j\omega|\underline{\mathbf{R}}^{\Delta}_m|/c_\alpha(\hat{\underline{\mathbf{K}}}(\widehat{\underline{\mathbf{R}}^{\boldsymbol{\Delta}}_{\mathbf{m}}}))]\cdot|\underline{\mathbf{R}}^{\Delta}_m|^{-1}\}, \tag{7}$$

where the resulting constant factor has been omitted. It is $\underline{\mathbf{R}}^{\Delta}_m=(\underline{\mathbf{R}}-\underline{\mathbf{R}}_m)$, $\widehat{\underline{\mathbf{R}}^{\boldsymbol{\Delta}}_{\mathbf{m}}}$ is the corresponding unit vector, and the unit vector $\underline{\mathbf{e}}_i$ indicates the direction of the applied traction $(i=x,y,\text{or}\,z)$. According to Eq. (7), any (continuous wave)

displacement field can be determined using an equidistant distribution of grid points within the transducing surface in accordance with the sampling theorem.

C. Numerical $\hat{\underline{K}}(\hat{\underline{R}})$ evaluation algorithm

The evaluation of Eq. (7) requires the determination of the wave vector direction $\hat{\underline{K}}$ that produces an energy contribution, i.e., a group velocity direction along the spatial direction $\hat{\underline{R}}$. The group velocity vector $\underline{c}$ is obtained as a function of the $\hat{\underline{K}}$ direction, so that the inversion of this relationship is required. However, due to the complex mathematical structure of these relationships the analytical inversion succeeds only in special cases. This has been performed for transversely isotropic material symmetry for *SH* waves yielding an exact analytical expression,[20] while for *qP* waves an approximate formulation has been obtained for a certain class of TI materials, e.g., unidirectionally grain-structured austenitic steels.[21] In arbitrary anisotropic material the inversion has to be performed numerically, where the shape and possible degeneracy of the slowness surfaces lead to additional difficulties. An efficient $\hat{\underline{K}}(\hat{\underline{R}})$ evaluation scheme has been implemented in the following way. The numerical program steps through the wave-vector positions, parametrized by an azimuthal angle α and a polar angle κ in steps of, e.g., 0.5°, and determines the corresponding group velocity vectors, which are in turn parametrized accordingly. The angles α and κ characterizing $\hat{\underline{K}}$ are stored as an array, where the azimuthal and polar angles determined for the $\underline{c}$ vectors are used as array indices. In evaluating Eq. (7), the polar and azimuthal angles for a given direction $\hat{\underline{R}}(=\hat{\underline{c}})$ are then used to look up the corresponding $\hat{\underline{K}}$ direction from the stored array. Additional care has to be taken in directions where group velocity is multivalued—here, multiple look-up arrays are generated and applied which also allows one to take care of the different wave-decay properties in these directions [$\mathcal{O}(R^{-1/2})$ and $\mathcal{O}(R^{-5/6})$ instead of $\mathcal{O}(R^{-1})$[24]].

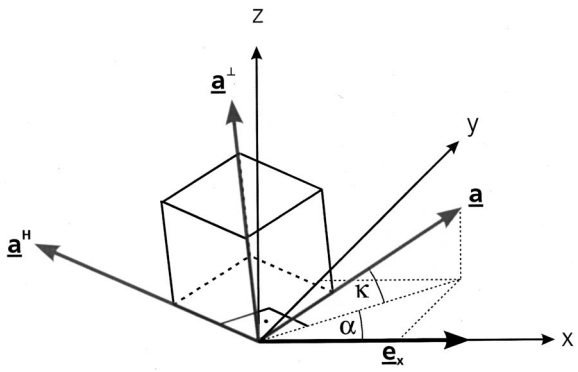


FIG. 1. Coordinate system and rotated (material) symmetry axes. Rotation transforms the reference vector $\underline{e}_x$ into $\underline{a}$, specified by α, the angle between its projection onto the $x-y$ plane and the x axis, and by κ, the angle between its $x-y$ projection and itself.

III. PLANE-WAVE RELATIONSHIPS

The dynamic behavior of a linear elastic, anisotropic medium can be described by the equation of motion for the displacement vector $\underline{u}$. Assuming harmonic time dependence $\sim e^{-j\omega t}$, where ω denotes the circular frequency, it can for a homogeneous solid be written in a general form according to

$$(\underline{\nabla}\cdot\underline{\underline{\underline{\underline{C}}}}\cdot\underline{\nabla})\cdot\underline{u}+\varrho\,\omega^2\underline{u}=-\underline{f}, \tag{8}$$

where ϱ is the mass density, $\underline{\nabla}$ is the gradient vector, and $\underline{f}$ accounts for the volume force density. The elastic properties of the homogeneous solid are described by the fourth-rank elastic tensor, which is for arbitrarily oriented orthotropic media given by

$$\begin{aligned}\underline{\underline{\underline{\underline{C}}}}^{OT}(\underline{a},\underline{a}^{H})=&(C_{33}-2C_{44})\underline{\underline{\underline{\underline{I}}}}\,\underline{\underline{\underline{\underline{I}}}}+C_{44}((\underline{\underline{\underline{\underline{I}}}}\,\underline{\underline{\underline{\underline{I}}}})^{1324}+(\underline{\underline{\underline{\underline{I}}}}\,\underline{\underline{\underline{\underline{I}}}})^{1342})+(C_{11}+C_{33}-2(C_{13}+2C_{55}))\underline{a}\,\underline{a}\,\underline{a}\,\underline{a}\\&+(C_{22}+C_{33}-2(C_{23}+2C_{44}))\underline{a}^{H}\underline{a}^{H}\underline{a}^{H}\underline{a}^{H}+(C_{13}+2C_{44}-C_{33})(\underline{\underline{\underline{\underline{I}}}}\,\underline{a}\,\underline{a}+\underline{a}\,\underline{a}\,\underline{\underline{\underline{\underline{I}}}})\\&+(C_{23}+2C_{44}-C_{33})(\underline{\underline{\underline{\underline{I}}}}\,\underline{a}\,\underline{a}^{H}+\underline{a}^{H}\underline{a}^{H}\underline{\underline{\underline{\underline{I}}}})+(C_{55}-C_{44})((\underline{\underline{\underline{\underline{I}}}}\,\underline{a}\,\underline{a})^{1324}+(\underline{a}\,\underline{a}\,\underline{\underline{\underline{\underline{I}}}})^{1324}\\&+(\underline{\underline{\underline{\underline{I}}}}\,\underline{a}\,\underline{a})^{1342}+(\underline{a}\,\underline{a}\,\underline{\underline{\underline{\underline{I}}}})^{1342})+(C_{66}-C_{55})(\underline{a}\,\underline{a}^{H}\underline{a}\,\underline{a}^{H}+\underline{a}^{H}\underline{a}\,\underline{a}^{H}\underline{a}+\underline{a}\,\underline{a}^{H}\underline{a}^{H}\underline{a}+\underline{a}^{H}\underline{a}\,\underline{a}\,\underline{a}^{H}).\end{aligned} \tag{9}$$

The derivation of this representation is outlined in Appendix B. The dyadic idemfactor appearing in Eq. (9) can be written—using dyadic products of the Cartesian unit vectors $\underline{e}_x$, $\underline{e}_y$ and $\underline{e}_z$—as

$$\underline{\underline{\underline{\underline{I}}}}=\{\delta_{ij}\underline{e}_i\underline{e}_j\}=\underline{e}_x\,\underline{e}_x+\underline{e}_y\,\underline{e}_y+\underline{e}_z\,\underline{e}_z=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}; \tag{10}$$

accordingly, the tetrad $\underline{\underline{\underline{\underline{I}}}}\,\underline{\underline{\underline{\underline{I}}}}$ is the dyadic product of $\underline{\underline{\underline{\underline{I}}}}$ with itself. Index changes, i.e., transpositions of elements in the elastic tensor are indicated by an upper indicial notation,[25] where, e.g.,

$$\underline{\underline{\underline{\underline{C}}}}^{1342}=\{C_{ijkl}\underline{e}_i\underline{e}_j\underline{e}_k\underline{e}_l\}^{1342}=C_{ijkl}\underline{e}_i\underline{e}_k\underline{e}_l\underline{e}_j=C_{iklj}\underline{e}_i\underline{e}_j\underline{e}_k\underline{e}_l\,. \tag{11}$$

Choosing the x axis as the initial spatial reference, the C_{ij} designate the respective nine elastic constants. Vector $\underline{a}$ indicates the position of the reference symmetry (x-) axis in 3D space after a rotation specified by two angles α and κ, while $\underline{a}^{H}$ ($\perp\underline{a}$) is parallel to the $x-y$ plane (Fig. 1).

Applying a three-dimensional Fourier transform with respect to $\underline{R}$ to the equation of motion (8) in terms of

$$\tilde{\underline{u}}(\underline{K},\omega)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\underline{u}(\underline{R},\omega)e^{-j\underline{K}\cdot\underline{R}}\mathrm{d}^3\underline{R}, \tag{12}$$

yields the dispersion equation

$$\tilde{\underline{\underline{W}}}(\underline{k},\omega)\cdot\tilde{\underline{u}}(\underline{k},\omega)=\tilde{\underline{f}}(\underline{K},\omega), \tag{13}$$

where the wave matrix is given by

$$\tilde{\underline{\underline{W}}}(\underline{K},\omega)=\underline{K}\cdot\underline{\underline{C}}\cdot\underline{K}-\varrho\,\omega^2\underline{\underline{I}}. \tag{14}$$

Insertion of the elastic stiffness tensor for OT media into Eq. (14) yields in abbreviated notation

$$\tilde{\underline{\underline{W}}}(\underline{K},\omega)\equiv\alpha\underline{\underline{I}}+\beta\underline{K}\,\underline{K}+\gamma\underline{a}\,\underline{a}+\delta(\underline{K}\,\underline{a}+\underline{a}\,\underline{K})+\varepsilon\underline{a}^{\mathrm{H}}\,\underline{a}^{\mathrm{H}}+\eta(\underline{K}\,\underline{a}^{\mathrm{H}}+\underline{a}^{\mathrm{H}}\,\underline{K})+\rho(\underline{a}\,\underline{a}^{\mathrm{H}}+\underline{a}^{\mathrm{H}}\,\underline{a}). \tag{15}$$

The Greek quantities in this equation depend on the elastic constants as well as on $\underline{a}$, $\underline{a}^{\mathrm{H}}$, and $\underline{K}$; the explicit expressions are given in Appendix C. The eigenvalues of Eq. (13) are determined by

$$\det\tilde{\underline{\underline{W}}}(\underline{K},\omega)=0; \tag{16}$$

from Eq. (15) it follows that

$$\det\tilde{\underline{\underline{W}}}(\underline{K},\omega)=(-\varrho\,\omega^2)^3+rK^2(-\varrho\,\omega^2)^2+sK^4(-\varrho\,\omega^2)+tK^6=0. \tag{17}$$

The quantities r, s, and t again depend on the C_{ij} and on the vectors $\underline{a}$, $\underline{a}^{\mathrm{H}}$, and $\hat{\underline{K}}$ as also listed in Appendix C. Writing Eq. (17) as

$$x^3+rK^2\,x^2+sK^4\,x+tK^6=0, \tag{18}$$

the substitution

$$y\equiv x+r/3\Leftrightarrow x\equiv y-r/3, \tag{19}$$

leads to the reduced equation

$$y^3+p\,y+q=0, \tag{20}$$

where $p=K^4(s-r^2/3)$ and $q=K^6(2r^3/27-rs/3+t)$. Using Cardano's formulas, the solutions of Eq. (20) follow according to

$$y_\alpha=\varepsilon_\alpha\,u+\varepsilon_{5-\alpha}\,v,\quad \alpha=1,2,3, \tag{21}$$

where

$$\varepsilon_{1,4}=1,\varepsilon_{2,3}=(-1\pm j\sqrt{3})/2, \tag{22}$$

$$u=K^2(\sqrt[3]{-q/2+\sqrt{D}}), \tag{23}$$

$$v=-K^2(p/3u), \tag{24}$$

$$D=K^{12}((p/3)^3+(q/2)^2). \tag{25}$$

Elementary manipulations finally provide the wave numbers as

$$K_\alpha^2=\varrho\,\omega^2/(r/3-(\varepsilon_\alpha\,u+\varepsilon_{5-\alpha}\,v)). \tag{26}$$

Furthermore, the modulus of phase velocity v_α is given by $v_\alpha=|\underline{s}_\alpha|^{-1}$, with slowness $\underline{s}_\alpha=\underline{K}_\alpha/\omega$.

The pertinent polarizations of the three bulk wave modes are obtained from the unit vectors $\hat{\underline{u}}_\alpha$, given by[26]

$$\hat{\underline{u}}_\alpha\hat{\underline{u}}_\alpha=\mathrm{adj}\,\tilde{\underline{\underline{W}}}/\mathrm{tr}\,(\mathrm{adj}\,\tilde{\underline{\underline{W}}}), \tag{27}$$

where ''adj'' and ''tr'' designate the adjoint and the trace of a matrix, respectively. Since for (strongly) anisotropic media the identification of waves according to specific dominant polarizations is meaningless, the three wave modes are designated according to their polarizations when propagated in certain symmetry directions. Thus, there are two quasishear waves and one quasipressure wave; accordingly $\alpha=qS1$, $qS2$, and qP will be used in the following.

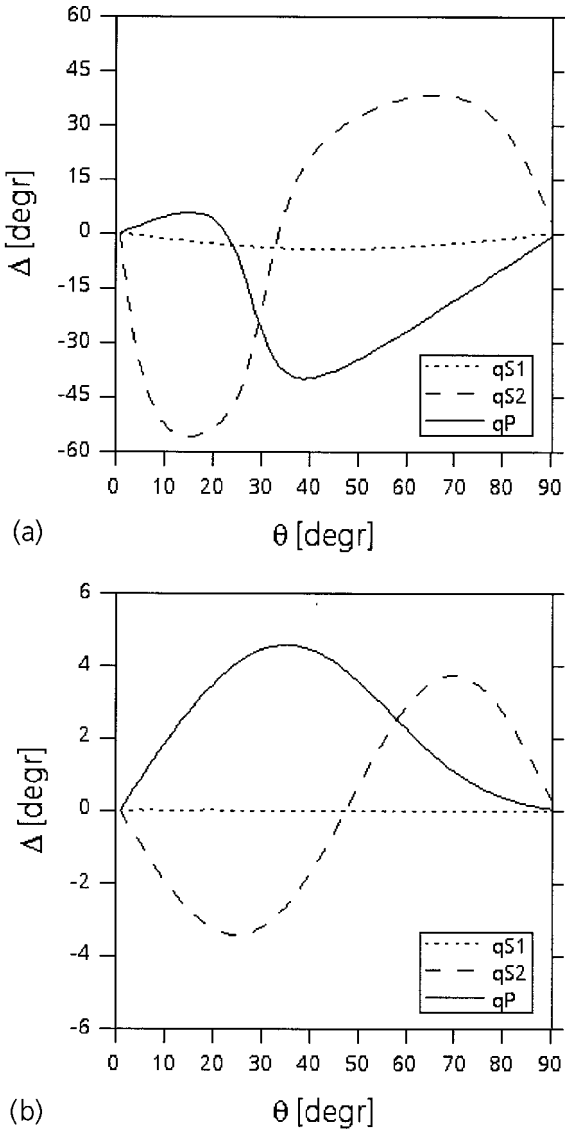


FIG. 2. Skewing angle Δ between group- and phase velocity plotted versus angle of incidence θ ($x-z$ plane, $\alpha=\kappa=0°$) for (a) the $[0_3/90]$ composite and (b) a quasi-isotropic composite. Note the different scales of the ordinates.

Applying the well-known definition of group velocity

$$\underline{c}=\partial\omega(\underline{K})/\partial\underline{K}, \tag{28}$$

and obeying that $\omega=K\,v$ yields

$$\underline{c}_\alpha=\left(\frac{\partial(K\,v)}{\partial\underline{K}}\right)_{\underline{K}_\alpha}=\left(v\,\frac{\partial K}{\partial\underline{K}}+K\frac{\partial v}{\partial\underline{K}}\right)_{\underline{K}_\alpha} \tag{29}$$

and then

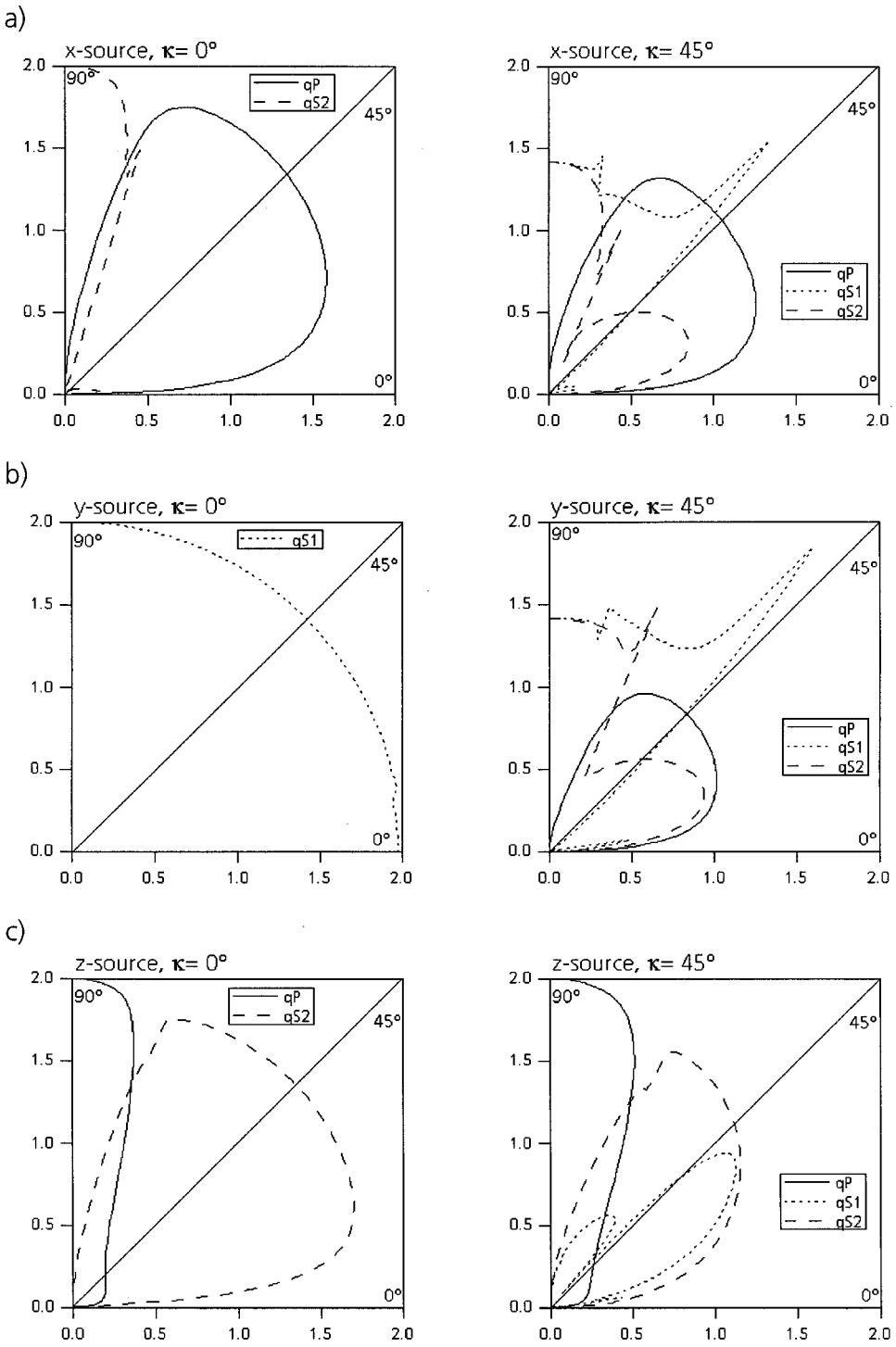


FIG. 3. Polar plots of the point-source directivities $g_{i\alpha}$ in the $x-z$ plane (in arbitrary units). The point sources are acting in (a) x-, (b) y-, and (c) z-direction on the stress-free surface of an orthotropic $[0_3/90]$ composite ($\alpha=0°,\kappa=0°,45°$).

$$\underline{\mathbf{c}}_\alpha = v_\alpha \hat{\underline{\mathbf{K}}} + (\underline{\mathbf{a}} - (\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})\,\hat{\underline{\mathbf{K}}})\,\frac{\partial v_\alpha}{\partial(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})} + (\underline{\mathbf{a}}^{\mathrm{H}} - (\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{K}}})\,\hat{\underline{\mathbf{K}}})\,\frac{\partial v_\alpha}{\partial(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{K}}})}. \tag{30}$$

The differentiations can be performed in a straightforward manner; the resulting expressions are omitted here.

IV. NUMERICAL EVALUATION EXAMPLES

In the following, numerical evaluation is conducted for a layered $[0_3/90]$ composite material. The elastic constants are[27] $C_{11}=94.0\,\mathrm{GPa}$, $C_{22}=13.0\,\mathrm{GPa}$, $C_{33}=34.0\,\mathrm{GPa}$, $C_{44}=3.6\,\mathrm{GPa}$, $C_{55}=7.2\,\mathrm{GPa}$, $C_{66}=4.2\,\mathrm{GPa}$, $C_{23}=9.1\,\mathrm{GPa}$, $C_{13}=8.2\,\mathrm{GPa}$, and $C_{12}=7.4\,\mathrm{GPa}$; the density ϱ used in the calculations is $\varrho=1.56\,\mathrm{g/cm^3}$. Basic information about wave propagation in anisotropic media is provided by the slowness- and group velocity characteristics, which can be obtained from Eqs. (26) and (30) in a staightforward manner.[28] However, more important from the practical point of view are quantitative measures of the deviation between energy flow and wave propagation direction. As an example, Fig. 2 displays the skewing angle diagrams plotted versus the angle of insonification in the $x-z$ plane for the $[0_3/90]$ com-

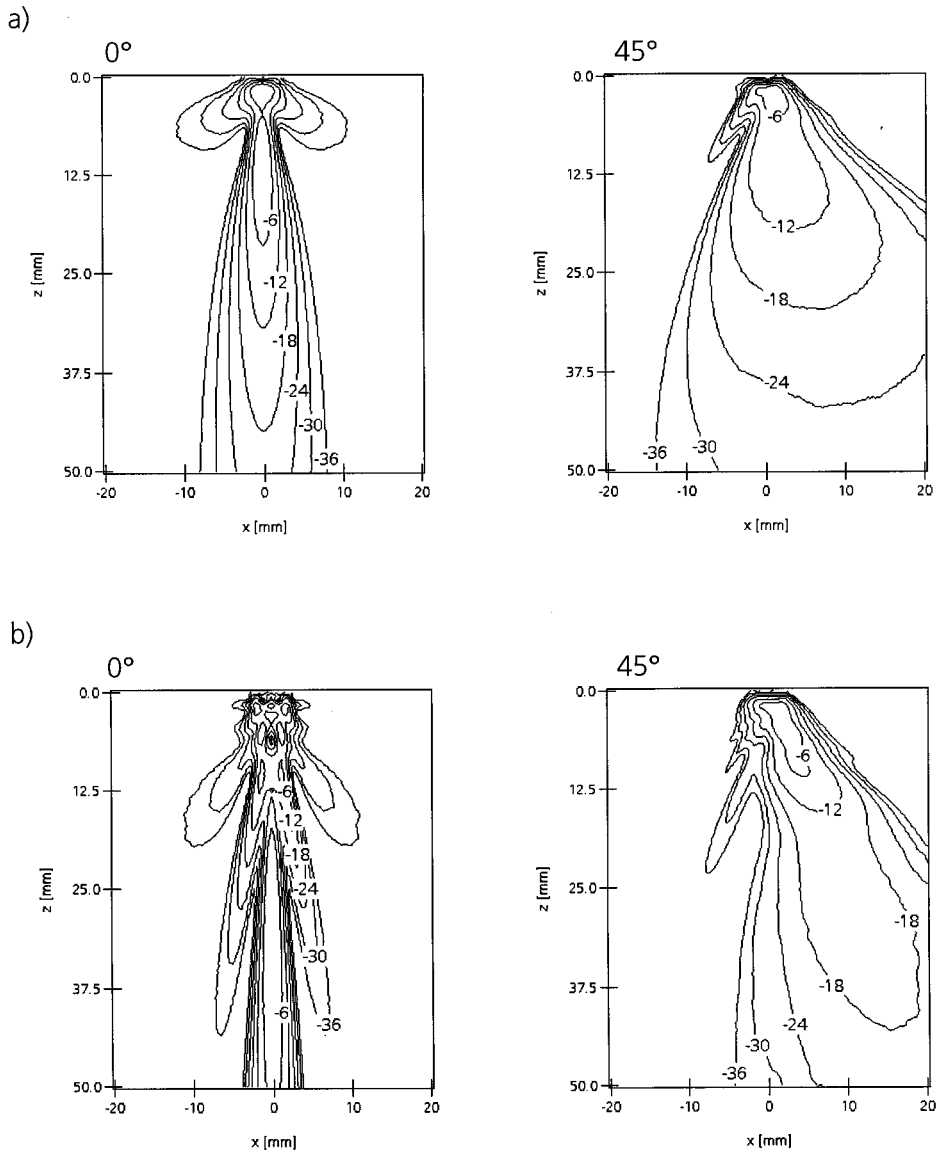


FIG. 4. Field patterns generated by a 6.3-mm qP transducer: frequencies are (a) 2.25 MHz and (b) 5 MHz ($\alpha = 0°, \kappa = 0°, 45°$). The logarithmic scaling ranges down to -36 dB as indicated.

posite and for a quasi-isotropic $[0/45/90/-45]$ composite.[27] While for the first material the amount of skewing ranges up to 60° for the qP wave, the quasi-isotropic composite shows only small effects of skewing for these insonification conditions.

In view of the excitation of the various wave types, polar diagrams of the point source directivity patterns $\underline{\mathbf{e}}_i \cdot \underline{\mathbf{g}}_\alpha$ in the $x-z$ plane are shown in Fig. 3. The point sources are located at the traction-free surface and act in the x-, y-, or z-direction. In calculating the directivities, different orientations of the $[0_3/90]$ composite material with respect to the surface have been assumed ($\alpha = 0°$, $\kappa = 0°$ and 45°). In the case $\alpha = \kappa = 0°$, point sources acting in the x- or z direction lead to qP- and $qS2$-wave excitation, while the $qS1$ wave is only generated by a y-acting point source (this is equivalent to the generation of the P-/SV waves and the SH wave, respectively, in the isotropic case). In the second case all three wave modes are excited when x-, y-, or z-directed point sources are considered, as also shown in Fig. 3.

For several material orientations, transducer-radiated (continuous wave) displacement fields are determined according to Eq. (7). Circular longitudinal-normal transducers of 2.25 and 5 MHz center frequency, respectively, and 6.3 mm in diameter are modeled, the point sources being approximately one-third of the smallest possible wavelength apart—at equal distance. The amplitude of the respective displacement vectors is displayed within a 40- by 50-mm^2 area in logarithmic scale. The results represent the dominant wave portion, i.e., the qP-wave mode for the probe applying forces in the z direction.

In Fig. 4, the two cases $\alpha = \kappa = 0°$ and $\alpha = 0°$, $\kappa = 45°$ are shown for the transducer frequencies 2.25 and 5.0 MHz, respectively. The—frequency-dependent—effects of beam skewing and spreading as well as the changes in the sidelobe structures are obvious. Figure 5 displays the on-axis response of the 2.25-MHz transducer for the orientations $\alpha = \kappa = 0°$ and $\alpha = 0°$, $\kappa = 90°$. As expected, the different elastic properties along the wave propagation direction lead to differences in the near-field lengths and the maximum amplitudes. Finally, Fig. 6 displays the field profiles at the near-field lengths for the same configurations. Here, the conditions along the main lobe are essentially the same for the $x-z$- and the $x-y$ plane, while differences exist in the sidelobe structures.

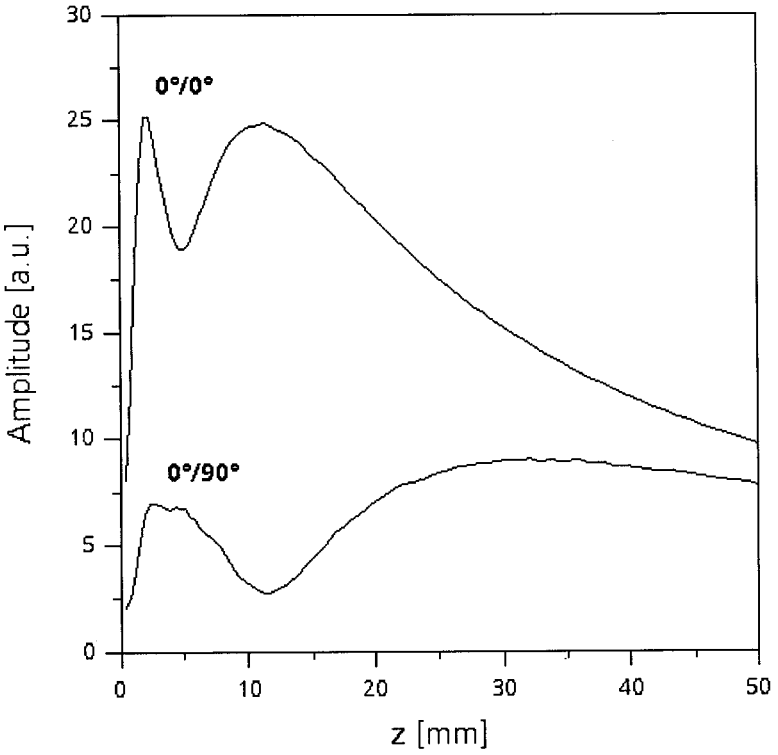


FIG. 5. On-axis response of the 2.25-MHz, 6.3-mm qP transducer for $\alpha=\kappa=0°$ and for $\alpha=0°$, $\kappa=90°$ (z is the on-axis distance to the transducer).

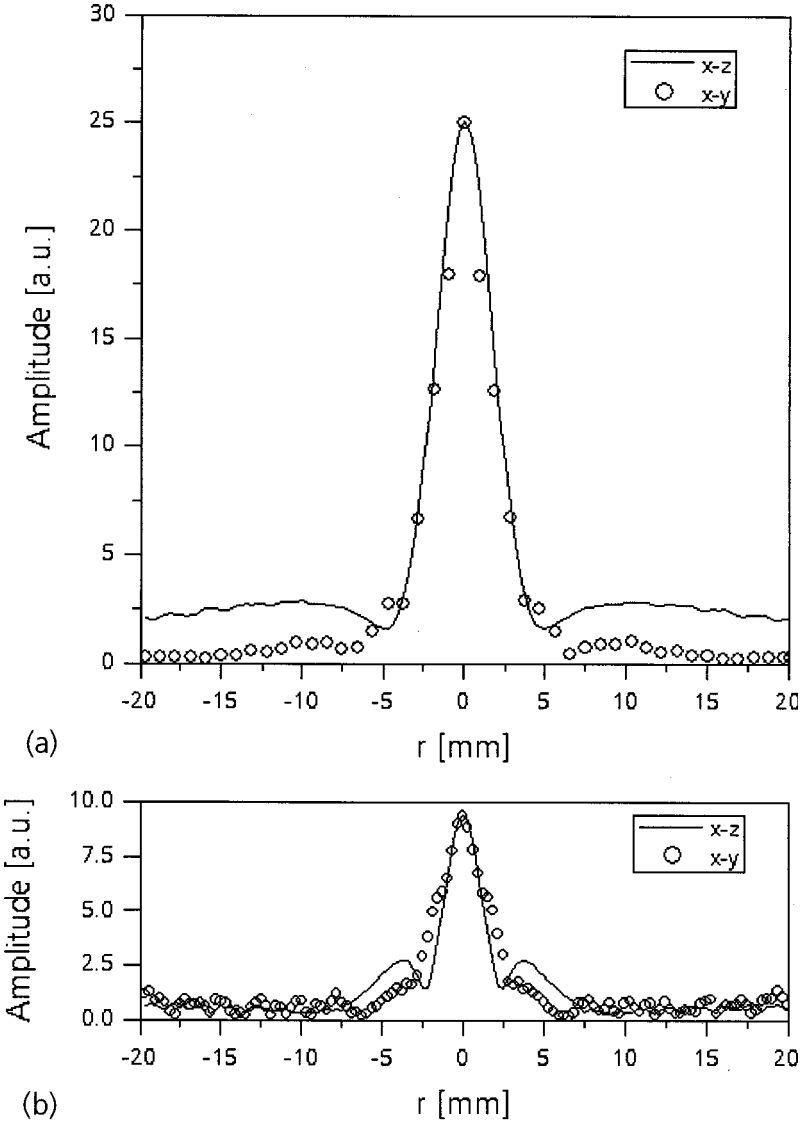


FIG. 6. Field profiles at the near-field lengths as generated by the 2.25-MHz 6.3 mm qP transducer for (a) $\alpha=\kappa=0°$ and (b) $\alpha=0°$, $\kappa=90°$ (r is the lateral coordinate).

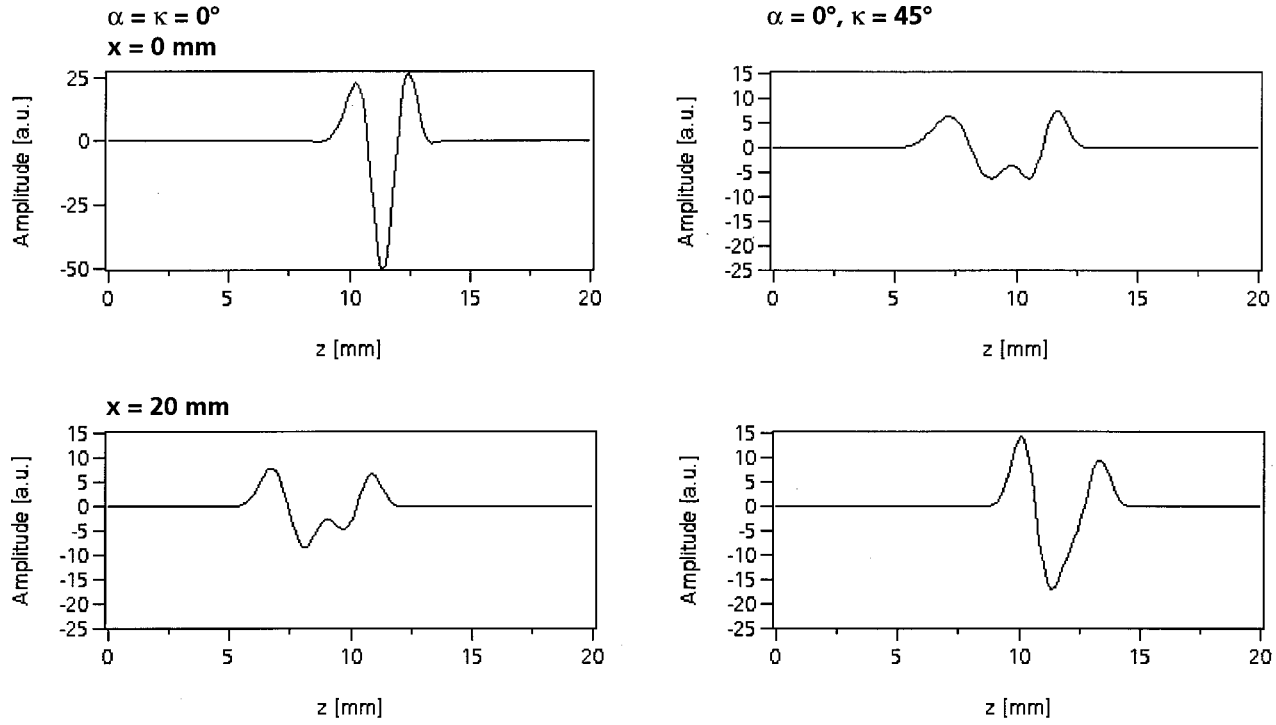


FIG. 7. Wavefront images of an RC2-qP-pulse in the $[0_3/90]$ composite for $\alpha=\kappa=0°$ (left) and for $\alpha=0°$, $\kappa=45°$ (right). The pulses are generated by the 2.25-MHz 6.3-mm transducer. The images are taken on-axis ($x=0$ mm) and off-axis ($x=20$ mm).

Time-dependent rf pulses are modeled by including the respective time function in Eq. (7). To account for a realistic experimental pulse, a raised cosine (RC2) time function is used according to

$$f_{RC2}(t)=[1-\cos(\omega t/2)]\cos(\omega t), \quad 0\leqslant t\leqslant 4\pi/\omega. \quad (31)$$

In the case of the $[0_3/90]$ configurations considered previously ($\alpha=0°$, $\kappa=0°,45°$), the pulses propagate along the paths illustrated in Fig. 4. For the $x-z$ plane, Fig. 7 shows time snapshots on the acoustic axis ($x=0$ mm) and off-axis ($x=20$ mm), calculated at $t=3\,\mu$s for the 2.25-MHz-transducer. For $\kappa=0°$ the pulse amplitude is maximal on the acoustic axis and fades smoothly in off-axis regions. In the second case, the pulse amplitude is maximal off-axis due to the beam skewing. The pulse appears more or less distorted on-axis, which is mainly due to the superposition of the pulses coming from the transducer edge (the same holds for $x=20$ mm in the $\kappa=0°$ case).

V. FINAL REMARKS

According to Eq. (7), any (continuous wave) displacement field can be determined using an equidistant distribution of point sources within a vibrating surface in accordance with the sampling theorem. The field patterns shown in Sec. IV have been calculated on a standard PC (Pentium II, 300 MHz) within a calculation time of less than half a minute. For obtaining a pulsed solution, the respective time function can be directly included into the calculation. Another straightforward computational approach is to calculate the harmonic solution at many frequencies and then numerically Fourier transform these data into the time domain.

In view of its computational efficiency, the presented semi-analytical method benefits from the reciprocity-based determination of the point source directivities described in Appendix A. Its applicability for anisotropic media follows from the general reciprocal relation for Green's functions by Burridge and Knopoff,[29] which allows one to show that the situation where the source is located at the surface and the receiver/observer at the depth is reciprocal to the situation and vice versa. In calculating such directivities, Wu, Nagy, and Adler[19] have obtained the same results using the reciprocity-based approach and Lamb's technique, respectively. Although the directivities are valid in the point-source far field, Eq. (7) also yields correct results within the transducer near field. This is illustrated in Fig. 8, where field profiles calculated at two-thirds of the near-field length are shown for a normal transducer. The results are taken from Ref. 30, where the semi-analytical method is compared with an exact numerical method presented previously by Roberts.[31]

The computational scheme presented in Sec. II can also be applied at interfaces: The particle displacement is calculated at the interface grid points and used as the prescribed displacement distribution at this interface. Taking into account the respective boundary conditions, this distribution is then propagated from the interface into the material as described above. The same principle holds for defects and other discontinuities with consideration of the appropriate point-source directivities, as presented previously.[32] Since orthotropy includes the higher symmetries of tetragonal, transversely isotropic, cubic, and isotropic, the results presented cover most of the materials of today's industrial interest.

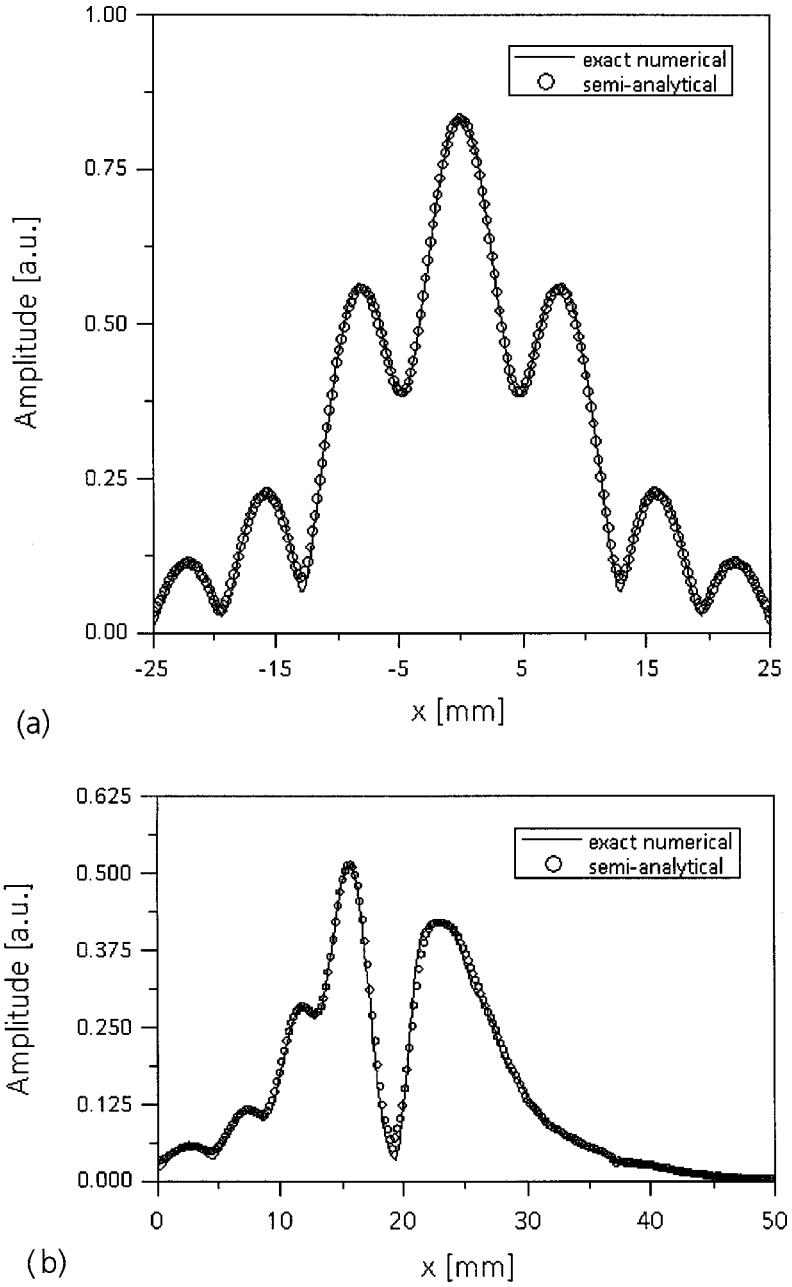


FIG. 8. Field profiles of a 2-MHz 20-mm qP transducer calculated for a transversely isotropic austenitic steel half-space with grains oriented at (a) $\alpha=\kappa=0°$ and (b) $\alpha=0°$, $\kappa=15°$. The profiles have been calculated in the $x-z$ plane at a distance of two-thirds of the near-field length ($z=86$ mm) using exact numerical calculation according to Ref. 31 and the semi-analytical method. The plots are normalized to the maximum amplitude at the near-field length ($z=126$ mm) in the case $\alpha=\kappa=0°$.

ACKNOWLEDGMENTS

The author would like to thank Professor Michael Krö-ning, director of IZFP, Saarbrücken, Germany for his interest in the subject and his steady support. This work was sponsored by the Deutsche Forschungsgemeinschaft DFG, which is gratefully acknowledged.

APPENDIX A: POINT-SOURCE DIRECTIVITY PATTERNS VIA RECIPROCITY

The derivation of the directional characteristics for normal and transverse point sources on the free surface of semi-infinite solids using the reciprocity theorem is briefly reviewed following Ref. 23. Considering wave propagation in the $x-z$ plane, a force $\underline{\mathbf{F}}$ is assumed to be applied at a point $\underline{\mathbf{R}}$ sufficiently remote from the origin. This force is directed parallel to the polarization vector of a given elastic wave mode ($qS1$, $qS2$, qP), which propagates in the direction of the radius vector pointing to the origin, so that only that particular type of wave is generated. The wave thus incident onto the free surface can be regarded as a plane wave in the neighborhood of the origin. Taking the reflected waves into consideration, the tangential ($i=x$, y) and normal ($i=z$) displacements at the origin are given by

$$u_i=\hat{\mathbf{u}}\cdot\mathbf{e}_i=U^I\hat{u}_i^I+U^{RqS1}\hat{u}_i^{RqS1}+U^{RqS2}\hat{u}_i^{RqS2}+U^{RqP}\hat{u}_i^{RqP}, \tag{A1}$$

where I designates the incident and R the reflected waves, respectively. From the reciprocity theorem it follows that the displacement of the given elastic wave mode at point $\underline{\mathbf{R}}$, which is generated by a force in the i direction applied at the origin, is also given by Eq. (A1). The particle displacement directivity for a point source on the free surface follows accordingly as

$$\underline{\mathbf{e}}_i \cdot \underline{\mathbf{g}}_\alpha = \hat{\underline{\mathbf{u}}} \cdot \underline{\mathbf{e}}_i . \tag{A2}$$

The determination of the reflected wave amplitudes for the OT media under concern in this article is described in Appendix D.

APPENDIX B: ELASTIC TENSOR REPRESENTATION

For an OT medium, which is characterized by three mutually orthogonal symmetry axes, nine elastic constants have to be considered. Replacing the first and the last two indices, respectively, of the tensor elements according to $11\rightarrow1$, $22\rightarrow2$, $33\rightarrow3$, $23\rightarrow4$, $31\rightarrow5$, $12\rightarrow6$ results in the well-known matrix representation

$$\begin{pmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{22} & C_{23} & 0 & 0 & 0 \\ C_{13} & C_{23} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{pmatrix}. \tag{B1}$$

Using Cartesian unit vectors, this representation can be put directly into an analytical formulation. A general representation of the elastic tensor for arbitrarily oriented OT media is then obtained by applying a transformation tensor $\underline{\underline{\mathbf{D}}}$ according to

$$\underline{\underline{\underline{\underline{\mathbf{C}}}}}^{\mathrm{gen}} = \underline{\underline{\mathbf{D}}} \cdot (\underline{\underline{\mathbf{D}}} \cdot \underline{\underline{\underline{\underline{\mathbf{C}}}}}^{\mathrm{cart}} \cdot \underline{\underline{\mathbf{D}}}) \cdot \underline{\underline{\mathbf{D}}}. \tag{B2}$$

Choosing the x direction as a spatial reference, $\underline{\underline{\mathbf{D}}}$ is to describe first a rotation through angle κ about the $-y$ axis, then a rotation about the z axis through angle α is to be performed (Fig. 1) according to

$$\underline{\underline{\mathbf{D}}} = \underline{\underline{\mathbf{D}}}_2 \cdot \underline{\underline{\mathbf{D}}}_1, \tag{B3}$$

$$\underline{\underline{\mathbf{D}}}_1 = \cos\kappa \underline{\underline{\mathbf{I}}} + (1-\cos\kappa)\underline{\mathbf{e}}_y \underline{\mathbf{e}}_y + \sin\kappa(\underline{\mathbf{e}}_z \underline{\mathbf{e}}_x - \underline{\mathbf{e}}_x \underline{\mathbf{e}}_z), \tag{B4}$$

$$\underline{\underline{\mathbf{D}}}_2 = \cos\alpha \underline{\underline{\mathbf{I}}} + (1-\cos\alpha)\underline{\mathbf{e}}_z \underline{\mathbf{e}}_z + \sin\alpha(\underline{\mathbf{e}}_y \underline{\mathbf{e}}_x - \underline{\mathbf{e}}_x \underline{\mathbf{e}}_y). \tag{B5}$$

In particular, it follows that

$$\underline{\underline{\mathbf{D}}} \cdot \underline{\mathbf{e}}_x \equiv \underline{\mathbf{a}} = \cos\alpha\cos\kappa \underline{\mathbf{e}}_x + \sin\alpha\cos\kappa \underline{\mathbf{e}}_y + \sin\kappa \underline{\mathbf{e}}_z , \tag{B6}$$

$$\underline{\underline{\mathbf{D}}} \cdot \underline{\mathbf{e}}_y \equiv \underline{\mathbf{a}}^{\mathrm{H}} = -\sin\alpha \underline{\mathbf{e}}_x + \cos\alpha \underline{\mathbf{e}}_y , \tag{B7}$$

$$\underline{\underline{\mathbf{D}}} \cdot \underline{\mathbf{e}}_z \equiv \underline{\mathbf{a}}^{\perp} = -\cos\alpha\sin\kappa \underline{\mathbf{e}}_x - \sin\alpha\sin\kappa \underline{\mathbf{e}}_y + \cos\kappa \underline{\mathbf{e}}_z ; \tag{B8}$$

thus, $\underline{\mathbf{a}}$ indicates the position of the reference symmetry axis in $3d$ space, while $\underline{\mathbf{a}}^{\mathrm{H}}(\perp \underline{\mathbf{a}})$ is parallel to the $x-y$ plane and $\underline{\mathbf{a}}^{\perp} = \underline{\mathbf{a}} \times \underline{\mathbf{a}}^{\mathrm{H}}$ (Fig. 1). The resulting formulation can be brought into a more convenient form by applying the orthogonality relationship

$$\underline{\underline{\mathbf{I}}} = \underline{\mathbf{a}}\,\underline{\mathbf{a}} + \underline{\mathbf{a}}^{\mathrm{H}}\underline{\mathbf{a}}^{\mathrm{H}} + \underline{\mathbf{a}}^{\perp}\underline{\mathbf{a}}^{\perp}, \tag{B9}$$

to eliminate vector $\underline{\mathbf{a}}^{\perp}$. After some algebraic manipulations, this leads to the general formulation given by Eq. (9) in Sec. III.

APPENDIX C: COEFFICIENTS OF THE OT-WAVE MATRIX AND ITS DETERMINANT

The abbreviated quantities in Eq. (15) are given by

$$\alpha = C_{44}K^2 + (C_{55}-C_{44})(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})^2 - \varrho\,\omega^2, \tag{C1}$$

$$\beta = (C_{33}-C_{44}), \tag{C2}$$

$$\gamma = (C_{55}-C_{44})K^2 + [C_{11}+C_{33}-2(C_{13}+2C_{55})](\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})^2 + (C_{66}-C_{55})(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\underline{\mathbf{K}})^2, \tag{C3}$$

$$\delta = [(C_{13}+C_{55})-(C_{33}-C_{44})](\underline{\mathbf{a}}\cdot\underline{\mathbf{K}}), \tag{C4}$$

$$\varepsilon = [C_{22}+C_{33}-2(C_{23}+2C_{44})](\underline{\mathbf{a}}^{\mathrm{H}}\cdot\underline{\mathbf{K}})^2 + (C_{66}-C_{55})(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})^2, \tag{C5}$$

$$\eta = [(C_{23}+C_{44})-(C_{33}-C_{44})](\underline{\mathbf{a}}^{\mathrm{H}}\cdot\underline{\mathbf{K}}), \tag{C6}$$

$$\rho = [(C_{33}-C_{44})-(C_{23}+C_{44})+(C_{12}+C_{66}) - (C_{13}+C_{55})](\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\underline{\mathbf{K}}). \tag{C7}$$

The coefficients in Eq. (17) are

$$r = r_0 + r_1(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^2 + r_2(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{K}}})^2, \tag{C8}$$

$$s = s_0 + s_1(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^2 + s_2(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{K}}})^2 + s_3(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^2(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{K}}})^2 + s_4(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^4 + s_5(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{K}}})^4, \tag{C9}$$

$$t = t_0 + t_1(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^2 + t_2(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^4 + t_3(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^6 + t_4(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{K}}})^2 + t_5(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{K}}})^4 + t_6(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{K}}})^6 + t_7(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^2(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{K}}})^2 + t_8(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^2(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{K}}})^4 + t_9(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^4(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{K}}})^2, \tag{C10}$$

where

$$r_0 = C_{33}+C_{44}+C_{55}, \tag{C11}$$

$$r_1 = (C_{66}-C_{44})+C_{11}-C_{33}, \tag{C12}$$

$$r_2 = (C_{66}-C_{55})+C_{22}-C_{33}, \tag{C13}$$

$$s_0 = C_{33}C_{44}+C_{33}C_{55}+C_{44}C_{55}, \tag{C14}$$

$$s_1 = C_{44}(C_{11}-C_{33})+C_{33}(C_{11}-C_{55})+(C_{66}-C_{44}) \times (C_{33}+C_{55})-C_{55}(C_{33}-C_{55})-(C_{13}+C_{55})^2, \tag{C15}$$

$$s_2 = (C_{33}+C_{44})(C_{66}-C_{55})+(C_{22}-C_{33})(C_{33}+C_{55}) - (C_{23}+2C_{44}-C_{33})(C_{23}+C_{33}), \tag{C16}$$

$$s_3 = (C_{66}-C_{55})[(C_{66}-C_{44})-(C_{33}-C_{55})] + (C_{22}-C_{33})(C_{11}-C_{33})+(C_{33}-C_{44})[2(C_{23} + 2C_{44}-C_{33})-(C_{66}-C_{55})]-2(C_{13}+C_{55})[(C_{66} - C_{55})+C_{12}-C_{13}]-[(C_{66}-C_{55})+C_{12}-C_{13}]^2 + (C_{23}+2C_{44}-C_{33})^2, \tag{C17}$$

$$s_4 = (C_{66}-C_{44})(C_{11}-C_{33})-(C_{33}-C_{55})(C_{11}-C_{55}) + (C_{13}+C_{55})^2, \tag{C18}$$

$$s_5=(C_{22}-C_{33})(C_{66}-C_{55})-(C_{33}-C_{44})[C_{22}-C_{33}-2(C_{23}+2C_{44}-C_{33})]+(C_{23}+2C_{44}-C_{33})^2, \tag{C19}$$

$$t_0=C_{33}C_{44}C_{55}, \tag{C20}$$

$$t_1=C_{33}C_{55}(C_{66}-C_{44})+C_{44}[C_{33}(C_{11}-C_{55})-C_{55}(C_{33}-C_{55})-(C_{13}+C_{55})^2], \tag{C21}$$

$$t_2=C_{44}[(C_{13}+C_{55})^2-(C_{33}-C_{55})(C_{11}-C_{55})]+(C_{66}-C_{44})[C_{33}(C_{11}-C_{55})-C_{55}(C_{33}-C_{55})-(C_{13}+C_{55})^2], \tag{C22}$$

$$t_3=(C_{66}-C_{44})[(C_{13}+C_{55})^2-(C_{33}-C_{55})(C_{11}-C_{55})], \tag{C23}$$

$$t_4=C_{33}C_{44}(C_{66}-C_{55})+C_{33}C_{55}(C_{22}-C_{33})-C_{55}[2(C_{33}-C_{44})+(C_{23}+2C_{44}-C_{33})]\times(C_{23}+2C_{44}-C_{33}), \tag{C24}$$

$$t_5=C_{33}(C_{22}-C_{33})(C_{66}-C_{55})-2(C_{23}+2C_{44}-C_{33})\times(C_{33}-C_{44})(C_{66}-C_{55})-C_{55}(C_{33}-C_{44})(C_{22}-C_{33}-2(C_{23}+2C_{44}-C_{33}))+[C_{55}-(C_{66}-C_{55})]\times(C_{23}+2C_{44}-C_{33})^2, \tag{C25}$$

$$t_6=(C_{66}-C_{55})[(C_{23}+2C_{44}-C_{33})^2-(C_{33}-C_{44})\times[C_{22}-C_{33}-2(C_{23}+2C_{44}-C_{33})]], \tag{C26}$$

$$t_7=-C_{44}(C_{66}-C_{55})(C_{33}-C_{55})+C_{33}(C_{66}-C_{55})(C_{66}-C_{44})+(C_{22}-C_{33})[C_{33}(C_{11}-C_{55})-C_{55}(C_{33}-C_{55})-(C_{13}+C_{55})^2]-C_{55}(C_{33}-C_{44})[-2(C_{23}+2C_{44}-C_{33})+(C_{66}-C_{55})]-2(C_{23}+2C_{44}-C_{33})(C_{33}-C_{44})(C_{11}-C_{55})-2C_{44}(C_{13}+C_{55})[(C_{66}-C_{55})+C_{12}-C_{13}]-C_{33}[(C_{66}-C_{55})+C_{12}-C_{13}]^2-C_{11}(C_{23}+2C_{44}-C_{33})^2+2(C_{13}+C_{55})(C_{12}+C_{66})(C_{23}+2C_{44}-C_{33})+2C_{55}(C_{23}+2C_{44}-C_{33})^2, \tag{C27}$$

$$t_8=-(C_{22}-C_{33})(C_{66}-C_{55})(C_{33}-C_{55})+2(C_{33}-C_{44})(C_{66}-C_{55})(C_{23}+2C_{44}-C_{33})-(C_{33}-C_{44})(C_{66}-C_{55})^2+(C_{22}-C_{33})(C_{13}+C_{55})^2-[C_{22}-C_{33}-2(C_{23}+2C_{44}-C_{33})](C_{33}-C_{44})(C_{11}-C_{55})+[C_{11}-C_{55}+(C_{66}-C_{55})]\times(C_{23}+2C_{44}-C_{33})^2+(C_{33}-C_{44})[(C_{66}-C_{55})+C_{12}-C_{13}]^2-2(C_{13}+C_{55})(C_{12}+C_{66})(C_{23}+2C_{44}-C_{33}), \tag{C28}$$

$$t_9=-(C_{66}-C_{44})(C_{66}-C_{55})(C_{33}-C_{55})+(C_{22}-C_{33})[-(C_{33}-C_{55})(C_{11}-C_{55})+(C_{13}+C_{55})^2]+[2(C_{23}+2C_{44}-C_{33})-(C_{66}-C_{55})](C_{33}-C_{44})(C_{11}-C_{55})+(C_{66}-C_{55})(C_{13}+C_{55})^2-2[(C_{66}-C_{55})+C_{12}-C_{13}](C_{55}-C_{44})(C_{13}+C_{55})-2(C_{13}+C_{55})(C_{12}+C_{66})(C_{23}+2C_{44}-C_{33})+(C_{11}-C_{55})(C_{23}+2C_{44}-C_{33})^2+(C_{33}-C_{55})[(C_{66}-C_{55})+C_{12}-C_{13}]^2. \tag{C29}$$

APPENDIX D: REFLECTION OF PLANE WAVES AT A FREE SURFACE

An arbitrarily oriented orthotropic half-space with the free surface being identical to the $x-y$ plane is considered. With the incident and the reflected waves given by

$$\underline{\mathbf{u}}^{R\alpha}(\underline{\mathbf{R}},\omega)=U^{I,R\alpha}\,\hat{\underline{\mathbf{u}}}^{I,R\alpha}\,e^{j\omega\,\underline{\mathbf{s}}^{I,R\alpha}\cdot\underline{\mathbf{R}}}, \tag{D1}$$

the amplitudes $U^{R\alpha}$ and the slownesses $\underline{\mathbf{s}}^{R\alpha}$ are to be determined. These quantities are obtained by considering the continuity of the slownesses ($\underline{\mathbf{s}}\cdot\underline{\mathbf{e}}_i$ continuous, $i=x,y$) and the condition of a stress-free boundary according to

$$\underline{\mathbf{e}}_z\cdot\underline{\underline{\mathbf{T}}}^{I}+\sum_{\alpha}\,\underline{\mathbf{e}}_z\cdot\underline{\underline{\mathbf{T}}}^{R\alpha}=\underline{\mathbf{0}}, \tag{D2}$$

where $\underline{\underline{\mathbf{T}}}$ designates the stress tensor. Considering wave propagation in the $x-z$ plane, which—due to the symmetry relations—is sufficient, Snell's law requires all projections of the $\underline{\mathbf{s}}$ vectors onto the interface to be identical ($s_x^I=s_x^{R\alpha}\equiv s_x$) and their y components to vanish ($s_y^I=s_y^{R\alpha}=0$). With these conditions, the $s_z^{R\alpha}$ components are left to be determined. Solving Eq. (17) for the slowness-z components yields a sixth order polynomial according to

$$\mathcal{P}(s_z^{R\alpha})=\sum_{i=0}^{6}\,C_i(s_z^{R\alpha})^{6-i}=0. \tag{D3}$$

Evaluation of the $s_z^{R\alpha}(s_x)$, which in some cases can have a nonvanishing imaginary part, is performed numerically using Bairstow's method[33] in a straightforward manner.

With the slowness vectors thus obtained, the reflection coefficients are determined from Eq. (D2). With $\underline{\nabla}\rightarrow j\omega\,\underline{\mathbf{s}}$, it follows that

$$\underline{\mathbf{e}}_z\cdot\underline{\underline{\mathbf{T}}}=\underline{\mathbf{e}}_z\cdot(\underline{\underline{\mathbf{C}}}:\underline{\nabla}\,\underline{\mathbf{u}})=j\omega\,U\underline{\mathbf{e}}_z\cdot(\underline{\underline{\mathbf{C}}}:\underline{\mathbf{s}}\,\hat{\underline{\mathbf{u}}})\equiv j\omega\,U\underline{\mathbf{t}}\cdot_z \tag{D4}$$

and further, that

$$\underline{\mathbf{t}}_{.z}=[((C_{13}+2C_{44}-C_{33})(\underline{\mathbf{s}}\cdot\hat{\underline{\mathbf{u}}})+(C_{11}+C_{33}-2(C_{13}+2C_{55}))(\underline{\mathbf{a}}\cdot\underline{\mathbf{s}})(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{u}}})+(C_{12}-C_{13}-C_{23}+2C_{44}-C_{33})$$
$$\times(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\underline{\mathbf{s}})(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{u}}}))(\underline{\mathbf{a}}\cdot\underline{\mathbf{e}}_z)+(C_{55}-C_{44})((\underline{\mathbf{a}}\cdot\underline{\mathbf{s}})(\hat{\underline{\mathbf{u}}}\cdot\underline{\mathbf{e}}_z)+(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{u}}})(\underline{\mathbf{s}}\cdot\underline{\mathbf{e}}_z))]\underline{\mathbf{a}}+(C_{66}-C_{55})((\underline{\mathbf{a}}^{\mathrm{H}}\cdot\underline{\mathbf{s}})(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{u}}})+(\underline{\mathbf{a}}\cdot\underline{\mathbf{s}})(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{u}}}))$$
$$\times(\underline{\mathbf{a}}\cdot\underline{\mathbf{e}}_z)\underline{\mathbf{a}}^{\mathrm{H}}+[C_{44}(\hat{\underline{\mathbf{u}}}\cdot\underline{\mathbf{e}}_z)+(C_{55}-C_{44})(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{u}}})(\underline{\mathbf{a}}\cdot\underline{\mathbf{e}}_z)]\underline{\mathbf{s}}+[C_{44}(\underline{\mathbf{s}}\cdot\underline{\mathbf{e}}_z)+(C_{55}-C_{44})(\underline{\mathbf{a}}\cdot\underline{\mathbf{s}})(\underline{\mathbf{a}}\cdot\underline{\mathbf{e}}_z)]\hat{\underline{\mathbf{u}}}$$
$$+[(C_{33}-2C_{44})(\underline{\mathbf{s}}\cdot\hat{\underline{\mathbf{u}}})+(C_{13}+2C_{44}-C_{33})(\underline{\mathbf{a}}\cdot\underline{\mathbf{s}})(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{u}}})+(C_{23}+2C_{44}-C_{33})(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\underline{\mathbf{s}})(\underline{\mathbf{a}}^{\mathrm{H}}\cdot\hat{\underline{\mathbf{u}}})]\underline{\mathbf{e}}_z. \tag{D5}$$

Defining the matrix

$$\underline{\underline{\mathbf{T}}}_{.z}{}^{R}\equiv(\underline{\mathbf{t}}_{.z}^{RqS1}\underline{\mathbf{t}}_{.z}^{RqS2}\underline{\mathbf{t}}_{.z}^{RqP}), \tag{D6}$$

and the vector

$$\underline{\mathbf{U}}^{R}\equiv(U^{RqS1}U^{RqS2}U^{RqP}), \tag{D7}$$

and using Eq. (D4), allows one to write Eq. (D2) as

$$U^{I}\underline{\mathbf{t}}_{.z}^{I}+\underline{\mathbf{U}}^{R}\cdot\underline{\underline{\mathbf{T}}}_{.z}{}^{R}=\underline{\mathbf{0}}. \tag{D8}$$

Solving this equation for the amplitudes of the reflected plane waves finally yields

$$\underline{\mathbf{U}}^{R}=-\,U^{I}\underline{\mathbf{t}}_{.z}^{I}\cdot(\underline{\underline{\mathbf{T}}}_{.z}{}^{R})^{-1}. \tag{D9}$$

Equation (D9) allows the numerical evaluation of the amplitudes $U^{R\alpha}$ for any orientation of the OT medium.

[1] J. A. Ogilvy, ''Ultrasonic Beam Profiles and Beam Propagation in an Austenitic Weld Using a Theoretical Ray Tracing Model,'' Ultrasonics **24**, 337–347 (1986).
[2] A. N. Norris, ''A Theory of Pulse Propagation in Elastic Anisotropic Solids,'' Wave Motion **9**, 509–532 (1987).
[3] A. H. Nayfeh and D. E. Chimenti, ''Free Wave Propagation in Plates of General Anisotropic Media,'' J. Appl. Mech. **56**, 881–886 (1989).
[4] B. Mandal, ''Reflection and Transmission Properties of Elastic Waves on a Plane Interface for General Anisotropic Media,'' J. Acoust. Soc. Am. **90**, 1106–1118 (1991).
[5] M. Spies, ''Elastic Waves in Homogeneous and Layered Transversely Isotropic Media: Plane Waves and Gaussian Wave Packets. A General Approach,'' J. Acoust. Soc. Am. **95**, 1748–1760 (1994).
[6] J. L. Rose, K. Balasubramaniam, and A. Tverdokhlebov, ''A Numerical Integration Green's Function Model for Ultrasonic Field Profiles in Mildly Anisotropic Media,'' J. Nondestruct. Eval. **8**, 165–179 (1989).
[7] B. P. Newberry and R. B. Thompson, ''A Paraxial Theory for the Propagation of Ultrasonic Beams in Anisotropic Solids,'' J. Acoust. Soc. Am. **85**, 2290–2300 (1989).
[8] M. Spies, ''Transducer Field Modeling in Anisotropic Media by Superposition of Gaussian Base Functions,'' J. Acoust. Soc. Am. **105**, 633–638 (1999).
[9] M. Spies, ''Transducer-Modeling in General Transversely Isotropic Media Via Point-Source-Synthesis. Theory,'' J. Nondestruct. Eval. **13**, 85–99 (1994).
[10] M. Rudolph, ''Ultrasonic Beam Models in Anisotropic Media,'' Ph.D. thesis, Iowa State University (1999).
[11] Q. C. Guo and J. D. Achenbach, ''Radiation of Ultrasound into an Anisotropic Solid,'' Ultrasonics **33**, 449–456 (1995).
[12] C. Y. Wang and J. D. Achenbach, ''Elastodynamic Fundamental Solutions for Anisotropic Solids,'' Geophys. J. Int. **118**, 384–392 (1994).
[13] C. Y. Wang and J. D. Achenbach, ''Three-Dimensional Time-Harmonic Elastodynamic Green's Functions for Anisotropic Solids,'' Proc. R. Soc. London, Ser. A **449**, 441–458 (1995).
[14] W. Lord, R. Ludwig, and Z. You, ''Developments in Ultrasonic Modeling with Finite Element Analysis,'' J. Nondestruct. Eval. **9**, 129–143 (1990).
[15] P. Fellinger, R. Marklein, K. J. Langenberg, and S. Klaholz, ''Numerical Modeling of Elastic Wave Propagation and Scattering with EFIT—Elastodynamic Finite Integration Technique,'' Wave Motion **21**, 47–66 (1995).
[16] A. Boström, ''Ultrasonic Probe Radiation and Crack Scattering in Anisotropic Media,'' Report, Division of Mechanics, Chalmers University of Technology, Göteborg, Sweden (1997).
[17] Y. H. Pao and V. Varatharajulu, ''Huygens' Principle, Radiation Conditions and Integral Formulas for the Scattering of Elastic Waves,'' J. Acoust. Soc. Am. **59**, 1361–1371 (1976).
[18] A. T. de Hoop, *Handbook of Radiation and Scattering of Waves* (Academic, London, 1995).
[19] K. Wu, P. B. Nagy, and L. Adler, ''Far-Field Radiation of a Vibrating Point Source in Anisotropic Media,'' J. Nondestruct. Eval. **10**, 71–78 (1991).
[20] M. Spies, ''Space-Time Far-Field Representation of Green's Function for Cross-Plane Shear Waves in General Transversely Isotropic Media,'' J. Acoust. Soc. Am. **102**, 733–740 (1997).
[21] M. Spies and M. Kröning, ''Accelerated Space-Time Modeling of Quasi-Longitudinal Waves in Austenitic Weld Structrues,'' in *Review of Progress in QNDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1997), Vol. 16, pp. 1175–1182.
[22] J. W. Strutt (3rd Baron Rayleigh), *Theory of Sound* (Dover Publications, New York, 1945), Vol. 1.
[23] K. Wu, P. B. Nagy, and L. Adler, ''Far Field Radiation of a Point Source on the Free Surface of Semi-Infinite Anisotropic Solids,'' in *Review of Progress in Quantitative NDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1990), Vol. 9A, pp. 149–156.
[24] V. T. Buchwald, ''Elastic Waves in Anisotropic Media,'' Proc. R. Soc. London, Ser. A **253**, 563–580 (1959).
[25] A. Ben-Menahem, and S. J. Singh, *Seismic Waves and Sources* (Springer, New York, 1981).
[26] F. I. Fedorov, *Theory of Elastic Waves in Crystals* (Plenum, New York, 1968).
[27] B. Hosten, ''Stiffness Matrix Invariants to Validate the Characterization of Composite Materials with Ultrasonic Methods,'' Ultrasonics **30**, 365–371 (1992).
[28] M. J. P. Musgrave, *Crystal Acoustics* (Holden-Day, San Francisco, 1970).
[29] R. Burridge and L. Knopoff, ''Body Force Equivalents for Seismic Dislocations,'' Bull. Seismol. Soc. Am. **54**, 1875–1888 (1964).
[30] M. Spies and R. A. Roberts (unpublished).
[31] R. A. Roberts, ''Ultrasonic Beam Transmission at the Interface Between an Isotropic and a Transversely Isotropic Solid Half-Space,'' Ultrasonics **26**, 139–147 (1988).
[32] M. Spies, ''Kirchhoff Evaluation of Scattered Elastic Wavefields in Anisotropic Media,'' J. Acoust. Soc. Am. **107**, 2755–2759 (2000).
[33] L. V. Atkinson, P. J. Harley, and J. D. Hudson, *Numerical Methods with* FORTRAN 77: *A Practical Introduction* (Addison-Wesley, Wokingham, 1989).

An approximate method for determining the maximum amplitude of higher-order duct modes

J. L. Horner
Department of Aeronautical and Automotive Engineering, Loughborough University, Loughborough LE11 3TU, United Kingdom

R. Lyons[a)]
Department of Civil and Building Engineering, Loughborough University, Loughborough LE11 3TU, United Kingdom

B. A. T. Petersson[b)]
Department of Aeronautical and Automotive Engineering, Loughborough University, Loughborough LE11 3TU, United Kingdom

(Received 12 July 2000; revised 26 March 2001; accepted 28 March 2001)

Previous investigations have used Hankel transforms to establish the amplitudes of the wave fields resulting from arbitrary angle impingement on a circular orifice in a baffle. Investigations have also been undertaken on the effects of model coupling between the higher modes in the orifice. In particular, these studies concentrated on establishing the contributions at the cut-on wave number for a particular mode. This work is extended to introduce a method of simply approximating the complex amplitude of the forward and backward waves in the duct at the cut-on wave number, based only on the modal wave number. The interest focuses on the cut-on wave number because the forcing functions for the various fields will be at a maximum at cut-on. This technique is established by observing the relationships between the peak values of the driving functions and the modal wave number. If simple approximations are established for the in-duct field, it is also possible to approximate the maximum amplitude of the scattered field. © *2001 Acoustical Society of America.*
[DOI: 10.1121/1.1373447]

PACS numbers: 43.20.Fn [ANN]

I. INTRODUCTION

A common problem in many disciplines is the transfer of acoustic energy from one domain to another via a device, an aperture, or both coupled together. Such an aperture is usually of symmetrical shape and has a finite depth, as if it were a short duct. It shall be referred to as an orifice within this study. At low frequencies, or small Helmholtz numbers, when the wavelength of the incident wave is much greater than the dimensions of the orifice, there are established approximations[1] that can be used to determine the reflected and transmitted fields. Transmission line theory may be invoked and the orifice system assigned a four-pole description. When the wavelength of the incident field approaches the dimensions of the orifice higher-order duct modes will propagate and any solution must include the effects of coupling between the higher-order modes.[2,3] Thus the impedance approach is no longer directly applicable. Although there is considerable literature that evaluates the fully coupled problem, none provide simple approximate solutions. This article presents a technique, based on a Hankel transform solution for the transmission of sound through a circular orifice,[4] to obtain, simply, an approximation for the wave amplitudes at cut-on for the duct/orifice. These amplitudes would be the maximum response for any mode and thus render conservative predictions for the system. In this case the system is assumed to be an orifice in a rigid wall separating two free acoustic fields.

The approach is an extension of previous work[5] on the effect of coupling between higher-order modes in such a system. It has been demonstrated that at the cut-on wave number for a particular mode, all modal contributions, apart from the mode of interest, are zero. This work determined the contributions to the fields using fully coupled numerically calculated integrals. The following method will utilize approximations to the full integral and show that it is possible to establish the maximum in-duct amplitudes for a particular cut-on wave number using only a single mode. The technique is limited to predictions only at certain wave numbers and it is assumed that the peak responses for a system occur at the cut-on of the various higher modes. This implies that it is possible to treat the influence of aperture devices as part of transmission line elements (four-poles).[6]

II. DESCRIPTION OF ACOUSTIC FIELD

Consider a circular orifice of radius R in a rigid wall (Fig. 1). For simplicity the following analysis is undertaken in terms of velocity potentials, from which the impedance may be obtained.[2] Following Ref. 4, using similar notation, the velocity potential for a plane wave of arbitrary angle, θ, may be given by

[a)]Electronic mail: r.lyons@lboro.ac.uk
[b)]Now at Institut fur Technische Akustik, Technische Universitat Berlin, Einsteinufer 25, D-10587 Berlin, Germany.

 0001-4966/2001/110(1)/80/6/$18.00

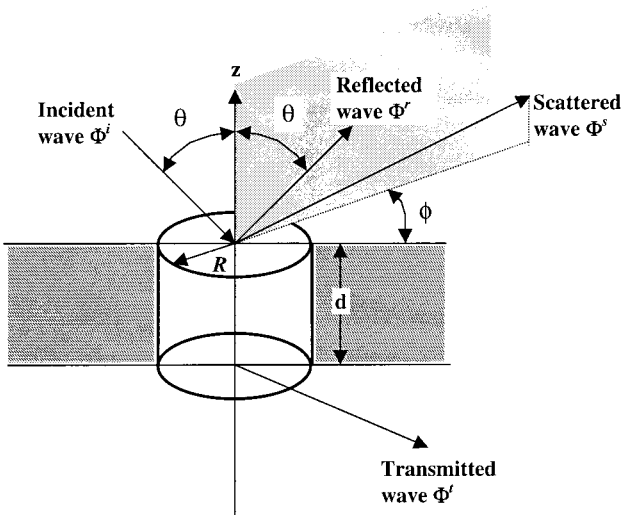


FIG. 1. Coordinate system for circular orifice.

$$\Phi^i(r,\phi,z)=e^{-ikz\cos\theta}\sum_{m=-\infty}^{\infty} i^m J_m(kr\sin\theta)e^{im\phi}, \tag{1}$$

where k is the free wave number. A time base of $e^{-i\omega t}$ is suppressed throughout for brevity. The resulting reflected and scattered velocity potentials are given by

$$\Phi^r(r,\phi,z)=e^{ikz\cos\theta}\sum_{m=-\infty}^{\infty} i^m J_m(kr\sin\theta)e^{im\phi},$$
$$\Phi^s(r,\phi,z)=\sum_{m=-\infty}^{\infty} e^{im\phi}\int_0^{\infty}\tilde{\Phi}^{s_m}(\zeta)J_m(\zeta r)e^{i\kappa z}\zeta\,d\zeta, \tag{2}$$

where $\tilde{\Phi}^{s_m}(\zeta)$ is the angle distribution factor, the scattered wave number, $\kappa=\sqrt{k^2-\zeta^2}$, and ζ is the recoil wave number. Inside the orifice the field is given over the depth d by

$$\Phi^d(r,\phi,z)=\sum_{m=-\infty}^{\infty} e^{im\phi}\sum_{n=1}^{\infty}[b_n^m\sin k_z(z+d) + c_n^m\cos k_z(z+d)]J_m(k_n r), \tag{3}$$

where the wave number in the z direction is $k_z=\sqrt{k^2-k_n^2}$ and k_n is the modal wave number. The transmitted velocity potential is given by

$$\Phi^t(r,\phi,z)=\sum_{m=-\infty}^{\infty} e^{im\phi}\int_0^{\infty}\tilde{\Phi}^{t_m}(\zeta)J_m(\zeta r)e^{-i\kappa(z+d)}\zeta\,d\zeta, \tag{4}$$

where $\tilde{\Phi}^{t_m}(\zeta)$ is the angle distribution factor for the transmitted field. Assuming a rigid wall surrounding the orifice, the velocity potential gradients may be summed for the orifice area and regions outside. Applying a Hankel transform to the two expressions determined from the summed velocity potential gradients yields the following expressions:[4,7]

$$\tilde{\Phi}^{s_m}(\zeta)=\frac{-i}{\kappa}\sum_{n=1}^{\infty} k_z(-c_n^m\sin k_z d + b_n^m\cos k_z d)I^m(k_n,\zeta), \tag{5}$$

$$\tilde{\Phi}^{t_m}(\zeta)=\frac{i}{\kappa}\sum_{n=1}^{\infty} k_z b_n^m I^m(k_n,\zeta), \tag{6}$$

in which[8]

$$I^m(k_n,\zeta)=\frac{R}{k_n^2-\zeta^2}[k_n J_{m+1}(k_n R)J_m(\zeta R) - \zeta J_m(k_n R)J_{m+1}(\zeta R)], \quad k_n\neq\zeta,$$
$$=\frac{R^2}{2}[J_m^2(k_n R)-J_{m-1}(k_n R)J_{m+1}(k_n R)], \quad k_n=\zeta. \tag{7}$$

Also, by equating the velocity potentials on the incident side of the orifice, it may be shown that[4]

$$(c_p^m\cos k_z d+b_p^m\sin k_z d)I^m(k_p,k_p) + \sum_{n=1}^{\infty} ik_z(b_n^m\cos k_z d - c_n^m\sin k_z d)I_{np}^m = 2i^m I^m(k\sin\theta,k_p), \tag{8}$$

where

$$I_{np}^m=\int_0^{\infty}\kappa^{-1}I^m(k_n,\zeta)I^m(k_p,\zeta)\zeta\,d\zeta, \tag{9}$$

(m,n) is the mode of interest, and (m,p) is any other higher mode of the duct. Equating velocity potentials and gradients of velocity potential, respectively, on the outlet side of the orifice gives, in accordance with Ref. 4,

$$c_p^m I^m(k_p,k_p)-\sum_{n=1}^{\infty} ik_z b_n^m I_{np}^m=0. \tag{10}$$

Thus Eqs. (8) and (10) must be solved for p number of modes in order to determine the coupled in-duct wave amplitudes b_n^m and c_n^m. Previous work[5] has shown that at the cut-on wave number ($\zeta=k_n$ and $k\sin\theta=k_n$) the contributions from other higher modes to the amplitude are zero due to the orthogonality of functions $I^m(k_n,\zeta)$, Eq. (7). As $I^m(k_n,\zeta)$ is zero at all other cut-on wave numbers, then the only contribution to the amplitude of the scattered field, Eq. (5), at the wave number, $k_n=\zeta$, is the single (m,n) mode. At any wave numbers other than a cut-on wave number, the function $I^m(k_n,\zeta)$ is nonzero and all cut-on modes contribute to the scattered and transmitted fields. Thus, the sum in both Eq. (8) and (10) vanishes, leaving only terms related to the single (m,n) mode present. Therefore, if Eqs. (8) and (10) are solved for a single uncoupled (m,n) mode, assuming $\zeta=k_n$, then b_n^m and c_n^m are given by

$$b_n^m=\frac{I^m(k_n,k_n)\gamma_n^m}{\Delta}, \tag{11}$$

$$c_n^m=\frac{ik_z I_{nn}^m\gamma_n^m}{\Delta}, \tag{12}$$

where $\Delta=(k_z^2 I_{nn}^{m\,2}+(I^m(k_n,k_n))^2)\sin k_z d+2ik_z I_{nn}^m I^m(k_n,k_n)\times\cos k_z d$, $\gamma_n^m=2i^m I^m(k\sin\theta,k_n)$, and the forcing function for the uncoupled mode, I_{nn}^m, is determined from Eq. (9).

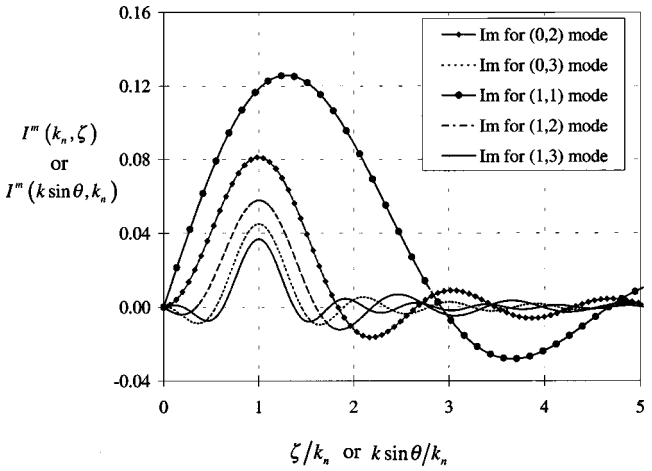


FIG. 2. $I^m(k_n,\zeta)$ and $I^m(k\sin\theta,k_n)$ for various modes.

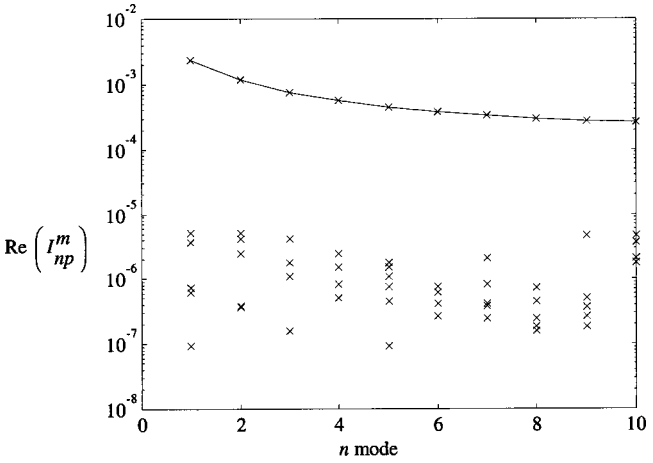


FIG. 3. Real part of the forcing function I^m_{np}; —×— I^m_{nn} data, × I^m_{np} data.

Thus to calculate b^m_n and c^m_n simply, it is necessary to evaluate the forcing terms, I^m_{nn}, $I^m(k_n,k_n)$, and γ^m_n at the cut-on wave number of the mode of interest. It is assumed that the three forcing functions will be at a maximum at cut-on.

III. FORCING FUNCTIONS $I^m(k_n,k_n)$ AND $I^m(k\sin\theta,k_n)$

Using a Hankel transform approach allows the three-dimensional problem of scattering from an orifice to be reduced to a two-dimensional one, where the system may be considered to be driven in the $z-\phi$ plane. Thus the function $I^m(k_n,\zeta)$ represents the excitation as a function of the recoil wave number, ζ, and the function $I^m(k\sin\theta,k_n)$ represents the excitation as a function of the projected wave number, $k\sin\theta$. If both functions are plotted against a normalized wave number (Fig. 2), it may be seen that the functions are identical for each (m,n) mode. Previously an expression was determined for the peak value of the forcing function using simple curve fitting and was of the form[5]

$$I^m(k_n,\zeta)\simeq aRk_n^{-b}, \tag{13}$$

where

$$a=0.30,\quad b=0.98,\quad n>1,$$

$$a=0.25,\quad b=1.10,\quad n=1.$$

The modes where $n=1$, those with no circumferencial nodal lines, peak at a wave number greater than $\zeta=k_n$ (Fig. 2) and the peak value is controlled by different constants.[5] This is due to the omission of any end correction effects in the model.

IV. FORCING FUNCTION I^m_{nn}

The forcing function I^m_{np} is determined from the integral given in Eq. (9) and represents the contribution of a mode (m,p) to the ''driving field'' for the scattering of the mode of interest (m,n). Thus when fully considering the resulting scattered field for the (m,n) mode, a contribution must be included for every other mode. When calculating a single mode contribution, I^m_{nn}, the mode of interest (m,n) only is considered. The integral, in Eq. (9), is over ζ, the recoil wave number, and represents a summation over all possible wave numbers. As the scattered wave number, κ, for a given driving wave number, k, will be imaginary for the nonpropagating or ''shadow forming'' (Ref. 2, Chap. 8) region of the scattered field, the result of the integral will be complex. It can be argued that the directivity of the scattered field is determined from the real part of the scattered wave number. When evaluating Eq. (9) numerically, if the upper integration limit is not sufficiently great, the resulting values of b^m_n and c^m_n, evaluated using a full modal sum, will be in error. It is possible to consider that the integrand in Eq. (9) couples two directivity patterns together, leading to a decrease in the nonpropagating part of the field. If the upper limit of integration does not fully consider all propagating scattered wave numbers, the acoustic energy will be spread over an incorrect area and hence the amplitudes will be in error.

Figures 3 and 4 show the real and imaginary parts of I^m_{np} plotted against n for the $m=1$ set of modes, $n=1$ to 10. Other values of m were investigated and indicate that the observations hold for all m values. For the set of data shown, the radius $R=1.0$ m and the duct length $d=0.1$ m. This particular aspect ratio of the duct was chosen to allow a significant number of modes to propagate in the duct. In Fig. 3,

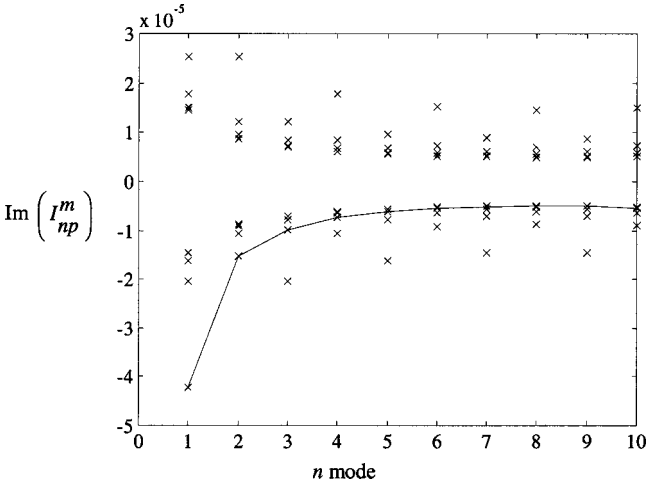


FIG. 4. Imaginary part of the forcing function I^m_{np}; —×— I^m_{nn} data. × I^m_{np} data.

TABLE I. Relative error in real part of I^m_{nn}.

Mode	k_n	Relative error in real part of I^m_{nn} (%)
2,1	5.13	3.4
0,2	6.38	0.4
2,2	11.21	3.6
0,3	11.71	−3.8
2,3	16.63	4.0
0,4	16.96	0.03
2,4	21.96	5.7
0,5	22.21	6.7
2,5	27.21	9.5

shown logarithmically to indicate differences, the curve represents the values of I^m_{nn}, the real part of the forcing function with only a single mode contributing to the field. The other data points represent the 90 values of I^m_{np}, which is the real part of the forcing function for a mode (m,n) with a contribution from another mode (m,p). It can be observed that the I^m_{nn} data set is some three orders of magnitude greater than the I^m_{np} data set. Also, it can be observed that the I^m_{nn} data set forms a curve reducing in magnitude as n increases. From examining the values of the real part of I^m_{nn} for other aspect ratios and mode numbers, it is found that for each m set, curves could be plotted through the real part of I^m_{nn}. Thus using simple curve fitting allows relationships to be established between mode numbers and the real part of I^m_{nn}. As for the previous forcing function, $I^m(k_n,k_n)$, the controlling parameters for the $n=1$ set are different from the $n\neq 1$ set. Also the $n=1$ set is a function of the modal wave number, k_n. However, the $m=0$ set is purely a function of the mode number n and the data set $m\neq 0$, $n\neq 1$ is a function of both the m and n mode numbers. It should be remembered that the $m=0$ data set represents forcing functions for those modes with only circumferential nodal lines, hence the sole dependency is on n.

It is worth reiterating that the main aim of this work is to find a simple method of approximating the complex amplitude of the waves in the duct at cut-on. By applying simple curve fitting techniques to various sets of data for the real part of I^m_{nn}, it is possible to establish the following empirical relationships. There are three different governing relationships depending on the mode number.

(1) For the $m=0$, $n>1$ set of modes,

$$\mathrm{Re}(I^0_{nn}) = 4.00\times 10^{-3} R^2 n^{-1.30}. \tag{14a}$$

(2) For the $m\neq 0$, $n=1$ set of modes,

$$\mathrm{Re}(I^m_{11}) = 4.86\times 10^{-3} R^2 (k_n,R)^{-1.12}. \tag{14b}$$

(3) For the $m\neq 0$, $n\neq 1$ set of modes,

$$\mathrm{Re}(I^m_{nn}) = 2.54\times 10^{-3} R^2 m^{-0.70} n^{\alpha},$$
$$\alpha = -1.08 m^{-0.39}. \tag{14c}$$

Table I indicates the relative error between the real part of I^m_{nn} calculated using Eqs. (7) and (9) and the real part calculated using Eqs. (14a)–(14c). For this example $R=0.6$ m and $d=0.2$ m, to ensure cut-on of higher modes. Note that it is not possible to determine the real part of I^m_{nn} for the (0, 1) mode using Eq. (14) as $k_n=0$. As the amplitude of this mode is simple to predict, it is not a disadvantage to the proposed method. Additionally, it may be observed that the relative error increases with increasing modal wave number, suggesting that this is a possible trend.

Figure 4 shows the imaginary part of I^m_{np} for the modes (1,1) to (1,10). The data points joined with a line represent the (n,n) values and the other points are the (n,p) values. It should first be noted that the imaginary part of the function is two orders of magnitude smaller than the real part. For the cases considered in this article, free fields are present on both sides of the orifice. Also the imaginary part has both positive and negative values with the (n,n) values being always negative. The consequence of the imaginary part of I^m_{nn} being negative is a wholly positive real part in the denominator of Eqs. (11) and (12), implying that the wave amplitudes are always complex.

As the real part of the forcing function I^m_{nn} dominates by more than an order of magnitude, it is assumed that only the real part needs to be evaluated for an approximate solution of a free field situation. This is also based on the view that the real part of I^m_{np} represents the propagating part of the scattered field, not the ''shadow forming'' part. The implication of treating the approximate value of I^m_{nn} as wholly real is a loss of the phase information associated with the function. In the following section, the error introduced by this approximation will be examined.

V. PREDICTION OF MAXIMUM IN-DUCT WAVE AMPLITUDES

Equations (11) and (12) are expressions for the amplitudes of the in-duct waves uncoupled from any other mode. In previous investigations,[5] it was found that at the cut-on wave number of the mode, only the mode of interest contributes to the in-duct amplitude. The function $I^m(k_n,\zeta)$, Eq. (7), is a maximum at $\zeta=k_n$ and is zero at $\zeta=k_p$. Thus in the expression for the amplitude of the scattered field, Eq. (5), only the (m,n) mode will be nonzero at $\zeta=k_n$. Thus at recoil wave numbers other than those equal to modal wave numbers, all modes contribute to the system response. At recoil wave numbers that equal modal wave numbers, only a single mode contributes to the total response. As cut-on for a mode occurs when there is coincidence between the driving wave number and the recoil and modal wave numbers, then the three forcing functions will be a maximum. Thus it is possible to predict the maximum of the three forcing functions using Eqs. (13) and (14). Once the three maximum forcing functions are known, it is possible to substitute the values into Eqs. (11) and (12) to predict the approximate maximum in-duct amplitudes.

Table II shows, as an example, the comparison between the approximate and fully coupled solutions of the in-duct wave amplitudes at cut-on for the mode of interest. Approximate values were determined using Eqs. (11)–(14) and are shown as a ratio of the response at cut-on from a fully coupled solution using Eqs. (7)–(10). For this comparison the radius $R=0.6$ m and the depth $d=0.2$ m. The fully coupled solution used ten modes to calculate the modal sum. From observation the fully coupled calculation appeared

TABLE II. Ratio of approximate–fully coupled solutions for modes (2,1)–(2,5) (Rounded to two decimal places).

	$b_n^m\left(\frac{\text{Approximate}}{\text{Fully coupled}}\right)$				$c_n^m\left(\frac{\text{Approximate}}{\text{Fully coupled}}\right)$			
Mode	Modulus	Phase	Real	Imaginary	Modulus	Phase	Real	Imaginary
2,1	0.98	1.04	0.92	1.00	0.93	0.92	0.93	1.00
2,2	1.13	1.02	1.04	1.14	1.03	1.03	1.03	1.01
2,3	1.10	1.02	0.21	1.10	1.04	1.04	1.02	1.00
2,4	1.09	1.02	1.18	1.07	1.01	1.01	1.02	1.00
2,5	1.13	1.02	1.16	1.06	1.01	1.00	1.01	1.00

stable after five modes were added to the sum. The driving wave number was set to $k=50\,\mathrm{m}^{-1}$ for all modes. The table provides the ratio of both the real and imaginary parts and the modulus and phase of the wave amplitudes. Thus a unity value represents full agreement between the approximate and full solutions.

Table II is for the set of modes (2,1–5). This set of data is representative of the results found using the approximate method. Apart from the real part of b_n^m of mode (2,3), the error is within acceptable bands. For the b_n^m wave for mode (2,3), the difference appears to be some 80% of the fully coupled calculation. On inspection of the data for this mode, it was found that the real part of b_n^m, determined from the full calculation, was about an order of magnitude smaller than the real parts of other modes, with the wave amplitude being almost wholly imaginary. The approximate method also yielded a small real part for this wave, but some 80% in error. The imaginary part of b_n^m (2,3) mode has an acceptable 10% error, leading to the modulus of the wave amplitude of the (2,3) mode to have an error of 10%. For both the real and imaginary parts of the c_n^m wave the error is small with the (2,1) mode having an error of approximately 7% and the other four modes having an error of less than 3%.

This data set was selected to show an example where apparently large errors in the real part of the solution from the approximate calculations lead to acceptable errors in both the modulus and phase of the amplitudes. It was observed from all data sets investigated that the errors in the calculation of c_n^m are always smaller than the errors in the calculation of b_n^m. The integral, Eq. (9), present in the calculation of c_n^m reduces the error in this term.

VI. PREDICTION OF MAXIMUM SCATTERED WAVE AMPLITUDES

Equation (5) determines the amplitude of the scattered field and the result of Eq. (5) is substituted into Eq. (2) to determine the spatial characteristic of the scattered field. In Eq. (5), the resulting scattered wave amplitude is the result of summing over n modes, the product of the magnitude of the in-duct waves at the scattered face and the forcing function $I^m(k_n,\zeta)$. As stated earlier, this forcing function, for a mode (m,n), is zero when the recoil wave number equals the modal wave number of any other p mode, and the maximum at cut-on may be approximated using Eq. (13). The approximate amplitudes of the in-duct waves at cut-on may be predicted using Eqs. (11)–(14). Thus at the cut-on wave number, Eq. (5) may be rewritten as follows. Due to the orthogonality properties of $I^m(k_n,k_n)$, the summation has been removed:

$$\tilde{\Phi}^{s_m}=\frac{2ik_zI^m(k_n,k_n)^2}{\kappa\Delta}(-I_{nn}^m\sin k_zd - iI^m(k_n,k_n)\cos k_zd), \tag{15}$$

where Δ is as stated for Eqs. (11) and (12).

Table III shows the comparison, for the $m=0$ and $m=2$ modes, of the predictions for the real part of the scattered wave amplitude using Eq. (15) and a fully coupled solution. As in the previous comparisons, the fully coupled solution was determined using ten modes in the sum. All dimensions are the same as for the previous calculation. Again a value of unity would indicate exact agreement. Only the real part of the scattered wave amplitude is investigated due to the real part dominating over the imaginary part at cut-on in this free field situation. Obviously if the scattered field occurred in a reverberant space, the imaginary part of the amplitude would also be considered. It may be observed that very good agreement is obtained for the $m=0$ set of data, with the maximum difference [mode (0,5)] being approximately 10%. For the $m=2$ data set, the maximum difference [mode (2,3)] is 22%. As discussed in the previous section, this mode has significant error in the approximate solution of the forward wave. It is worth noting that the higher modes, which show greater differences between the approximate and fully coupled solutions, contain less energy than the lower modes and therefore the larger differences here are not critical.

TABLE III. Ratio of approximate to fully coupled solutions (rounded to two decimal places).

Mode	Real part of $\tilde{\Phi}^{s_m}\left(\frac{\text{Approximate}}{\text{Fully coupled}}\right)$
0,2	1.04
0,3	0.98
0,4	1.06
0,5	1.10
2,1	0.80
2,2	1.19
2,3	1.22
2,4	1.08
2,5	1.04

VII. DISCUSSION

In previous investigations it has been established that at the cut-on wave number for a higher mode in a duct, only the mode of interest contributes to the field at that wave number. By following a Hankel transform solution for the fields resulting from the impingement of a wave on a circular orifice in a hard wall, it is possible to obtain an expression for the uncoupled higher mode wave amplitudes in the duct. These wave amplitudes depend on the dimensions of the duct, the modal and axial wave numbers, and three forcing functions. The forcing functions can be considered to represent the modal in-duct field, the modal external field, and the nett modal cross-coupling. By investigating each of the driving terms individually it was found that simple approximate expressions could be obtained for the peak load, at cut-on, for each (m,n) mode. For the radial and axial driving terms, the approximate expression was a function of duct radius and modal wave number. Although the modal wave number may be considered to only drive the radial field, the axial and radial forcing functions are identical at cut-on and hence both may be described using the same expression. The approximate expression for the driving term, dependent on the scattered field, is a function of both (m,n), duct radius and the modal wave number.

When the maximum values for the three driving terms are substituted into the expressions for the uncoupled in-duct wave amplitudes it is possible to obtain an approximate complex amplitude for a particular cut-on mode. As the amplitude of any mode will be greatest at cut-on, the approximate value represents the maximum possible amplitude. Comparison of the approximate solutions with the solutions from a fully coupled calculation shows good agreement. The maximum error occurred in the forward wave, b_n^m, due to the lack of a net modal cross-coupling term in the numerator.

It is possible to extend the simple approximate method to the scattered field and obtain an approximate expression for the amplitude of that field. The size of the error in that calculation directly relates to the size of the error in the in-duct wave amplitudes. Although only the scattered field is considered in this investigation, it is possible to extend the analysis to obtain an approximate expression for the transmitted field.

The advantage of the method presented above is that it allows easy calculation of the maximum possible in-duct modal wave amplitudes for a system. By obtaining maximum values for each mode, it is possible, for a particular driving wave number, to establish the dominant contribution to the part of the field of interest. It should be noted that the method only works for modes that are cut-on in the orifice and this method is currently developed only for free fields or highly reverberant fields: While it is appreciated that in a diffuse field, contributions from both the real and imaginary parts in the orifice field can be influential, the fact that there is an order of magnitude difference between those parts means that the true field will clearly be dominated by the real part.

[1] A. D. Pierce, *Acoustics: An Introduction to its Physical Principles and Applications* (McGraw-Hill, New York, 1981), Chap. 7.

[2] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton U. P., Princeton, NJ, 1986), Chap. 9.

[3] F. P. Mechel, *Schallabsorber, Band I, Äussere Schallfelder; Wechselwirkungen* (Hirzel, Leipzig, 1989), Chap. 5.

[4] K. H. Jun and H. J. Eom, ''Acoustic scattering from a circular aperture in a thick hard screen,'' J. Acoust. Soc. Am. **98**, 2324–2327 (1995).

[5] J. L. Horner, R. Lyons, and B. A. T. Peterson, ''Approximations for modal coupling in scattered fields from orifices,'' J. Acoust. Soc. Am. **108**, 488–493 (2000).

[6] R. Lyons, P. C. Macey, and J. L. Horner, ''Acoustic performance of finite length apertures using finite element analysis,'' in Collected papers Proc. 137th Meeting of the Acoustical Society of America, Technische Universitat Berlin, CD-ROM ISBN 3-9804568-5-4 (1999).

[7] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Sec. 8.5.

[8] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products* (Academic, New York, 1980).

Effectiveness of the continuous wavelet transform in the analysis of some dispersive elastic waves

Yoon Young Kim[a] and Eung-Hun Kim
School of Mechanical and Aerospace Engineering, and Institute of Advanced Machinery Design, Seoul National University, Shinlim-Dong San 56-1, Kwanak-Gu, Seoul 151-742, Korea

(Received 23 May 2000; accepted for publication 16 April 2001)

Although there have been many investigations employing the continuous wavelet transform for the analysis of dispersive waves, they seem to lack theoretical justifications for the effectiveness of the continuous wavelet transform (CWT) over other time–frequency analysis tools such as the short-time Fourier transform (STFT). The goal in this paper is to offer theoretical and experimental justifications for its effectiveness by comparing the performance of CWT and STFT in terms of their time–frequency analysis capabilities of certain dispersive elastic waves. The waves in consideration are elastic flexural waves generated by an impact in a solid circular cylinder. The ridge analysis procedure is employed to estimate instantaneous frequencies by CWT and STFT. Although in the present investigation we are focused on a limited class of dispersive waves, it gives an insight into the effectiveness of CWT for the analysis of other types of dispersive wave systems. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1378348]

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

I. INTRODUCTION

In recent years, the wavelet transform receives much attention in all areas of engineering. In particular, the continuous wavelet transform has become useful as a time–frequency analysis tool of some wave signals.[1–5] Before the continuous wavelet transform is available, the short-time Fourier transform[6] and the Wigner–Vile Distribution (WVD)[7] are mainly used for the time–frequency analysis of wave signals. For instance, STFT was applied for the elastic wave analysis in strings, beams, and shells[8] and WVD was employed for the identification of structure-born noise components.[9]

STFT and WVD have some advantages over CWT, but these techniques have some limitations. The resolution achieved by STFT in the time–frequency plane is independent of the location in the time–frequency plane so that STFT is inappropriate for the analysis of a wave signal whose instantaneous frequency varies rapidly. WVD has excellent time–frequency resolution, but the smoothing of WVD is necessary to eliminate cross-term effects. If a signal arrives at the time when the interference of two other signals by WVD appears, the smoothing often prohibits the signal from being identified especially when the signal has a small magnitude (see Kim and Kim[10]).

In order to overcome the limitations of STFT and WVD, CWT has been applied in elastic wave problems (Kishimoto *et al.*;[11] Inoue *et al.*;[12] Kim and Kim[10]). CWT can measure the time–frequency variation of a wave signal like STFT. However, CWT has a different time–frequency resolution: it uses a shorter time support for the analysis of high-frequency components and a longer time support for the analysis of low-frequency components. Kishimoto *et al.*[11] and Inoue *et al.*[12] employed CWT for the dispersion analysis of flexural waves in a beam and suggested that CWT was effective in identifying impact locations in beams. Gaul *et al.*[13] applied CWT to a plate case. Jeong *et al.*[14] utilized CWT to identify fracture source locations. Kim and Kim[10] utilized CWT as a signal analysis tool for damage detection in a beam and presented some evidence showing that CWT is the only successful time–frequency tool for their problem investigated.

These investigations have suggested that CWT is a very useful time–frequency analysis tool in various situations. Obviously, the success of CWT comes from the varying time–frequency resolution of CWT. However, none of these investigations has examined theoretically why the continuous wavelet transform outperforms other time–frequency transforms.

Our objective in this paper is to investigate theoretically why CWT must be used for the analysis of some dispersive waves. To this end, we examine theoretically the performance of CWT and STFT with respect to the capability to trace a flexural wave signal consisting of rapidly varying frequency components. The specific problem we will consider in this paper is the damage detection problem studied by Kim and Kim.[10]

For the theoretical investigation, we first look into how well windowed Fourier ridges and wavelet ridges capture instantaneous frequencies of a wave signal. Since the ridges can trace the instantaneous frequencies only when certain conditions are met, we investigate the effect of the conditions on the performance of CWT and STFT. Numerical results based on the theoretical analysis are given for the problem dealing with the measurement of dispersive flexural waves near an impact location. The experimental results for this problem have been given by Kim and Kim.[10] Although only a specific dispersive elastic wave is considered in the present theoretical analysis, we give an insight into the effectiveness of CWT in the time–frequency analysis of other types of dispersive wave systems.

[a]Electronic mail: yykim@snu.ac.kr

 0001-4966/2001/110(1)/86/9/$18.00

II. TIME–FREQUENCY ANALYSIS: CWT VERSUS STFT

In this section we compare two linear transforms that have been used for time–frequency analysis of wave signals. These transforms analyze the frequency evolution of a one-dimensional time signal in the time–frequency plane. The Wigner–Vile distribution (WVD), a quadratic transform, has excellent resolution in the time–frequency plane, but it suffers from interferences by cross terms. The serious problem by the interferences, when WVD is applied for damage detection, has been demonstrated.[10] Subsequently, we will be mainly concerned with two linear transforms, CWT and STFT, which do not have inherent interference problems. In this section we will review briefly the underlying properties of CWT and STFT and then discuss how accurately these transforms can trace rapidly varying wave signals. See Mallat[15] for more detailed accounts of these transforms.

A. Short-time Fourier transform (STFT)

The short-time Fourier transform (STFT), often called the windowed Fourier transform, was introduced by Gabor[6] to measure localized frequency components of sound. The short-time Fourier transform of a function f belonging to a finite-energy signal space $L^2(t)$ is defined as

$$Sf(u,\xi)=\int_{-\infty}^{+\infty} f(t)g^*_{u,\xi}(t)dt=\int_{-\infty}^{+\infty} f(t)g(t-u)e^{-i\xi t}\,dt, \tag{1}$$

with

$$g_{u,\xi}(t)=e^{i\xi t}g(t-u), \tag{2}$$

where * denotes the complex conjugate. The transform $Sf(u,\xi)$ can measure the behavior of f in the neighborhood of time u and frequency ξ. The function $g(t)$ in Eq. (1) is the window function, which is assumed to be real and symmetric. Thus $\hat{g}(\omega)$ is also real and symmetric.

For a given window $g(t)$ with the time spread σ_t and frequency spread σ_ω, the time and frequency spreads of the modulated translated window $g_{u,\xi}(t)$ are independent of (u,ξ). This means that the short-time Fourier transform has the same resolution across the time–frequency plane. If the resolution covered by σ_t and σ_ω is denoted by a rectangle, which is usually referred to as a Heisenberg box, the tiling of STFT in the time–frequency plane may be represented by Fig. 1. Mallat illustrates the effectiveness of STFT based on a Gaussian window $g(t)=e^{-t^2/2\sigma^2}$ in dealing with signals having slowly varying instantaneous frequencies such as linear and quadratic chirps.

The instantaneous frequency is defined as the positive derivative of $\phi(t)$,

$$\omega_{\text{inst}}(t)=\frac{d\phi(t)}{dt}, \tag{3}$$

where $\phi(t)$ is the modulating phase for a real signal $f(t)$ having an amplitude $a(t)$,

$$f(t)=a(t)e^{i\phi(t)} \quad [a(t)\geqslant 0]. \tag{4}$$

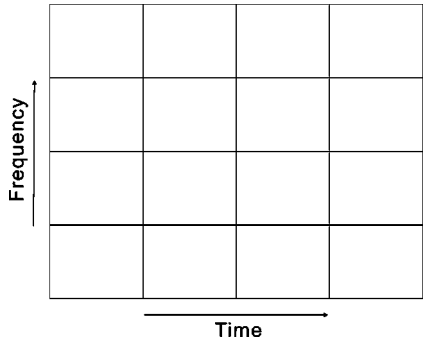


FIG. 1. The tiling of the short-time Fourier transform in the time–frequency plane.

B. Continuous wavelet transform (CWT)

The definition of the continuous wavelet transform[15] is

$$Wf(u,s)=\int_{-\infty}^{+\infty} f(t)\psi^*_{u,s}(t)dt=\int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{s}}\psi^*\left(\frac{t-u}{s}\right)dt, \tag{5}$$

with

$$\psi_{u,s}(t)=\frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right). \tag{6}$$

The function $\psi(t)$ is called a mother wavelet satisfying the admissibility condition

$$\int_{-\infty}^{+\infty}\frac{|\hat{\psi}(\omega)|^2}{|\omega|}d\omega<\infty, \tag{7}$$

where $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(t)$.

The existence of the integral in Eq. (7) requires that

$$\hat{\psi}(0)=0, \tag{8a}$$

$$\int_{-\infty}^{+\infty}\psi(t)dt=0. \tag{8b}$$

Equation (5) indicates that the mother wavelet $\psi(t)$ is translated by the translation parameter u and dilated by the scaling parameter s when a signal $f(t)$ is analyzed.

In order to see the advantage of CWT over STFT, we need to examine the tiling of CWT in the time–frequency plane illustrated in Fig. 2. To this end, we assume that the center frequency of $\hat{\psi}(\omega)$ is η and that the time and frequency spreads of $\psi(t)$ are σ_t and σ_ω, respectively. Since

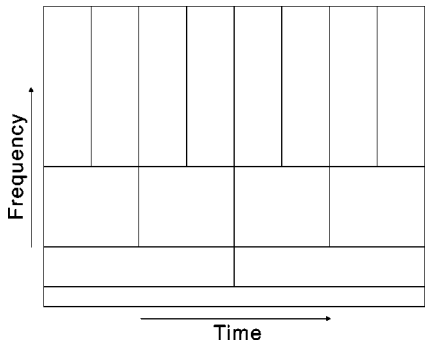


FIG. 2. The tiling of the continuous wavelet transform in the time–frequency plane.

$\psi_{u,s}(t)$ involves scaling by s, we can show that the time and frequency spreads of $\psi_{u,s}(t)$ are $s\sigma_t$ and σ_ω/s, respectively. We can also show that the center of the corresponding Heisenberg box is located at $(u, \eta/s)$.

In analyzing the frequency evolution of a signal using CWT, one must utilize analytic wavelets such as the Gabor wavelet or the harmonic wavelet.[16] The Fourier transform $\hat{\psi}(\omega)$ of an analytic wavelet $\psi(t)$ is identically zero at $\omega = 0$. We remark that although the harmonic wavelet has a good frequency resolution, it is not suitable for the analysis of signals having rapidly varying components because it has a large support in the time domain. On the other hand, the Gabor wavelet has the smallest Heisenberg box and can be adjusted to have a shorter time support and thus we will consider only the Gabor wavelet in the present work.

The Gabor wavelet $\psi(t)$ is a complex-valued modulated Gaussian function, defined as

$$\psi(t) = e^{i\eta t} g(t), \tag{9a}$$

$$g(t) = \frac{1}{(\sigma^2\pi)^{1/4}} e^{-t^2/2\sigma^2}, \tag{9b}$$

where η is the center frequency of $\hat{\psi}(\omega)$ and σ is a measure of the spread of $\psi(t)$. The shape of the Gabor wavelet is controlled by the product G_s of η and σ,

$$G_s = \sigma\eta. \tag{10}$$

This parameter, which may be called the Gabor shaping factor,[10] affects significantly the time resolution of the wavelet and thus the performance of the corresponding wavelet transform. Although Eq. (8a) requires that $G_s \to \infty$, the condition $G_s \gg 1$ suffices for actual numerical computation. The Gabor wavelet satisfying this condition can be an approximate analytic wavelet. Kim and Kim[10] have reported the importance of an optimal selection of G_s in the analysis of dispersive waves. In this work, we will investigate theoretically the effect of G_s on the performance of the Gabor wavelet transform.

C. Ridge method for instantaneous frequency estimation

The spectrogram $P_s f(u,\xi) = |Sf(u,\xi)|^2$ and the normalized scalogram $(\xi/\eta) P_w f(u,\xi) = |W_f(u,s)|^2/s$ (with $\xi = \eta/s$) measure the energy of f in a time–frequency neighborhood of (u,ξ). The ridge algorithm[17] computes instantaneous frequencies from the local maxima of spectrogram and scalogram. The result of the ridge analysis provides criteria to select windows for STFT and mother wavelets for WT.

Some results[15] by the ridge analysis, which are needed for the present investigation, are summarized first. For the subsequent analysis, the signal $f(t)$ is assumed to take the following form:

$$f(t) = a(t)\cos\phi(t), \tag{11}$$

where $a(t)$ and $\phi(t)$ are an amplitude and a time-varying phase of $f(t)$. The instantaneous frequency ω_{inst} is defined by Eq. (3).

1. STFT

It is convenient to introduce a scaling parameter s to adjust the size of window $g(t)$ such that

$$g_{s,u,\xi}(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{i\xi t}. \tag{12}$$

For the subsequent analysis, it is assumed that the support of a real symmetric window is $[-\frac{1}{2}, \frac{1}{2}]$, and $|\hat{g}(\omega)| \leqslant \hat{g}(0)$ for all frequencies ω.

Using $g_{s,u,\xi}(t)$ as the windowed Fourier atom, the following result can be found:

$$\begin{aligned} Sf(u,\xi) &= \int_{-\infty}^{+\infty} f(t) g^*_{s,u,\xi}(t)\, dt \\ &= \frac{\sqrt{s}}{2} a(u) e^{i(\phi(u)-\xi u)} (\hat{g}(s[\xi - \dot{\phi}(u)]) + \epsilon(u,\xi)), \end{aligned} \tag{13}$$

where $\dot{(\ )}$ denotes the differentiation with respect to its own argument.

The corrective term, or the error term, $\epsilon(u,\xi)$, is bounded by

$$|\epsilon(u,\xi)| \leqslant \epsilon_{a,1} + \epsilon_{a,2} + \epsilon_{\phi,2} + \sup_{|\omega| \geqslant |s\dot{\phi}(u)|} |\hat{g}(\omega)|, \tag{14}$$

with

$$\epsilon_{a,1} \leqslant \frac{s|\dot{a}(u)|}{|a(u)|}, \tag{15a}$$

$$\epsilon_{a,2} \leqslant \sup_{t \in [u - s/2, u + s/2]} \frac{s^2|\ddot{a}(t)|}{|a(u)|}, \tag{15b}$$

If $s|\dot{a}(u)|/|a(u)| \leqslant 1$,

$$\epsilon_{\phi,2} \leqslant \sup_{t \in [u - s/2, u + s/2]} s^2|\ddot{\phi}(t)|. \tag{16}$$

The last term in Eq. (14) can be negligible if

$$s|\dot{\phi}(u)| \geqslant \Delta\omega, \tag{17}$$

where $\Delta\omega$ is the bandwidth of the selected window $\hat{g}(\omega)$.

At ridges, i.e., at $\xi(u) = \dot{\phi}(u)$, the spectrogram $P_s(u,\xi) = |Sf(u,\xi)|^2$ becomes maximum and $\epsilon_{a,1}$ is given by

$$\epsilon_{a,1} = \frac{s|\dot{a}(u)|}{|a(u)|} |\hat{g}(2s\dot{\phi}(u))|. \tag{18}$$

This term is negligible when Eq. (17) is satisfied. Thus at ridge points,

$$P_s(u,\xi) = |Sf(u,\xi)|^2 = \frac{s}{2} a^2(u) (\hat{g}(0) + \epsilon_{a,2} + \epsilon_{\phi,2})^2. \tag{19}$$

In order to extract instantaneous frequencies accurately by using the spectrogram, the following conditions must be met:

$$\epsilon_{a,2} \leqslant \sup_{t \in [u - s/2, u + s/2]} \frac{s^2|\ddot{a}(t)|}{|a(u)|} \ll 1, \tag{20a}$$

$$\epsilon_{\phi,2} \leqslant \sup_{t\in[u-s/2,u+s/2]} s^2|\ddot{\phi}(t)| \ll 1, \tag{20b}$$

$$\frac{\Delta\omega}{s|\dot{\phi}(u)|} \ll 1. \tag{20c}$$

2. CWT

As in STFT, $Wf(u,s)$ for a signal $f(t)$ given by Eq. (11) can be expressed as

$$Wf(u,s) = \int_{-\infty}^{+\infty} f(t)\psi^*_{u,s}(t)dt = \frac{\sqrt{s}}{2}a(u)e^{i\phi(u)}(\hat{g}(s[\xi-\dot{\phi}(u)]) + \epsilon(u,\xi)), \tag{21}$$

where the corrective term $\epsilon(u,\xi)$ is given by Eq. (14). Following the same treatment used for the ridge analysis of $Sf(u,\xi)$, the normalized scalogram at the ridge points, $\xi = \dot{\phi}(u)$, is given by

$$\frac{\xi}{\eta}P_w f(u,\xi) = \frac{|Wf(u,s)|^2}{s} = \frac{1}{4}a^2(u)\left|\hat{g}\left(\eta\left[1-\frac{\dot{\phi}(u)}{\xi}\right]\right) + e(u,\xi)\right|^2, \tag{22}$$

where

$$s = \frac{\eta}{\xi}.$$

The corrective terms $\epsilon_{a,2}$ and $\epsilon_{\phi,2}$ become negligible when

$$\frac{s^2|\ddot{a}(u)|}{|a(u)|} = \frac{\eta^2}{|\dot{\phi}(u)|^2}\frac{|\ddot{a}(u)|}{|a(u)|} \ll 1, \tag{23a}$$

$$s^2|\ddot{\phi}(u)| = \eta^2\frac{|\ddot{\phi}(u)|}{|\dot{\phi}(u)|^2} \ll 1, \tag{23b}$$

$$\frac{\Delta\omega}{s|\dot{\phi}|} = \frac{\Delta\omega}{\eta} \ll 1. \tag{23c}$$

Comparing Eq. (20b) and Eq. (23b), the wavelet transform is useful in capturing very rapidly varying instantaneous frequencies as long as the corresponding instantaneous frequencies are large. In the next section, we will examine the detection of near-field dispersive elastic wave signals where this property of the wavelet transform plays a key role.

III. ANALYSIS OF WAVES GENERATED BY AN IMPACT

In order to show the effectiveness of CWT over STFT, we consider a damage detection problem in which bending waves are generated by an impact.

A. Damage detection problem

Figure 3(a) shows a simply-supported beam having a cross section. The beam is excited by a steel ball dropped at the center of the beam marked by A. The experimentally estimated duration of the contact time between the ball and the beam is about 85 μsec. Bending strains generated by the impact are measured by strain gages at the sampling rate of 1 μsec at point B.

The beam has a small cut at location C and Fig. 3(b) shows the shape and size of the cut. Kim and Kim[10] proposed a damage diagnostic technique based on the continuous wavelet transform and in this paper we give the theoretical background for the success of the wavelet transform.

B. Wave propagation in beams

Before carrying out the wavelet transform on the strain signals measured at B, we consider the wave theory in a beam and give some fundamental results that are necessary for the subsequent analysis.

Although the exact Pochhammer–Chree equation[18–20] is well known, the Timoshenko beam theory[21] will be employed to analyze low-frequency waves, say, waves having frequencies below 30 kHz. Since the measured strain waves at location B [see Fig. 3(a)] are mostly below 30 kHz, the

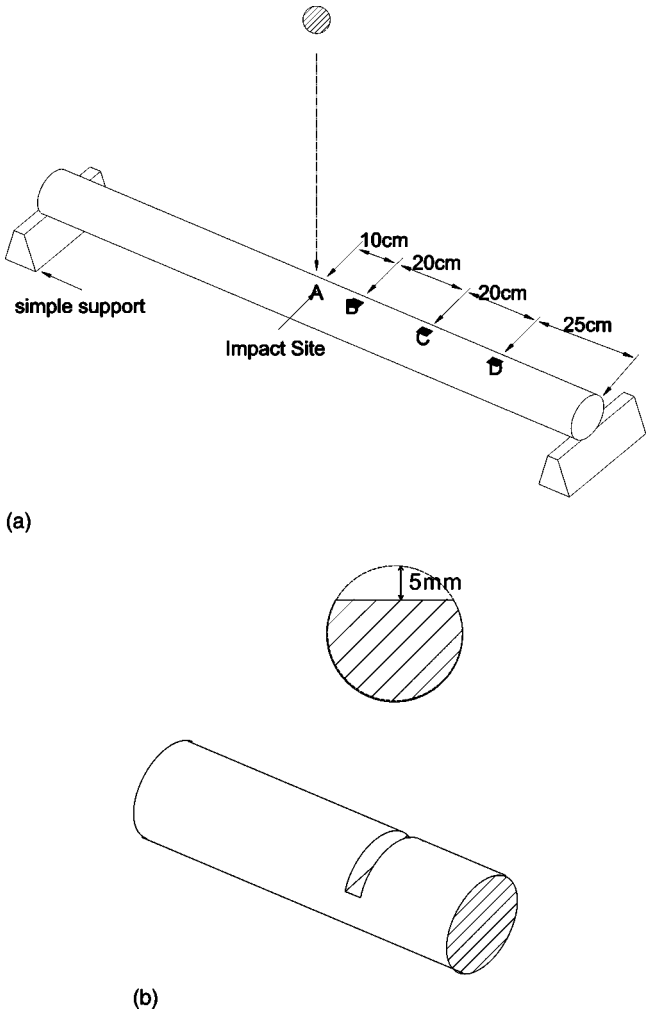


FIG. 3. Experimental setup for a solid circular beam excited by a steel ball at the center. (Diameter $d=2$ cm, Young's modulus $E=117.2$ GPa, density $\rho=8.90\times10^3$ kg/m^3). (a) Experimental setup; (b) close view near C.

use of the Timoshenko beam theory can be justified (see, e.g., Achenbach,[22] Miklowitz,[23] and Doyle[24]).

The governing wave equations by the Timoshenko beam theory are given by

$$KAG\left[\frac{\partial^2 v}{\partial x^2}-\frac{\partial \psi}{\partial x}\right]=\rho A\frac{\partial^2 v}{\partial t^2}, \tag{24a}$$

$$EI\frac{\partial^2 \psi}{\partial x^2}+KGA\left[\frac{\partial v}{\partial x}-\psi\right]=\rho I\frac{\partial^2 \psi}{\partial t^2}, \tag{24b}$$

where t and x are the time and the axial coordinate. The transverse displacement of the neutral axis and the rotation of the normal of a beam denoted by v and ψ, respectively. The density ρ, Young's modulus E, and the shear modulus G denote the material properties of a beam. The area and the moment of inertia are designated by A and I, respectively. The shear correction factor K is taken to be 0.8.

Assuming the solution of Eq. (24) in propagating wave form,

$$v=v_0 e^{-i(kx-\omega t)}, \tag{25a}$$

$$\psi=\psi_0 e^{-i(kx-\omega t)}, \tag{25b}$$

the following dispersion relation relating the wave number k and the frequency ω can be found (see, e.g., Doyle[24]):

$$k=k_1(\omega) \quad \text{or} \quad k_2(\omega), \tag{26a}$$

where

$$k_1(\omega)=\sqrt{\frac{1}{2}\left[\left(\frac{1}{c_s}\right)^2+\left(\frac{1}{c_b}\right)^2\right]\omega^2+\sqrt{\left(\frac{\omega}{c_b q}\right)^2+\frac{1}{4}\left[\left(\frac{1}{c_s}\right)^2-\left(\frac{1}{c_0}\right)^2\right]^2\omega^4}}, \tag{26b}$$

$$k_2(\omega)=\sqrt{\frac{1}{2}\left[\left(\frac{1}{c_s}\right)^2+\left(\frac{1}{c_b}\right)^2\right]\omega^2-\sqrt{\left(\frac{\omega}{c_b q}\right)^2+\frac{1}{4}\left[\left(\frac{1}{c_s}\right)^2-\left(\frac{1}{c_0}\right)^2\right]^2\omega^4}}. \tag{26c}$$

In Eq. (26), c_b, c_s, and q are defined as

$$c_b\equiv\sqrt{\frac{E}{\rho}}, \quad c_s\equiv\sqrt{\frac{KG}{\rho}}, \quad q\equiv\sqrt{\frac{I}{A}}. \tag{27}$$

The arrival time of a wave signal is governed by the group velocity C_g, defined as

$$C_g\equiv\frac{d\omega}{dk}. \tag{28}$$

Since flexural waves in a beam are dispersive, the group velocity is frequency dependent.

C. Phase analysis of strain signals

Consider the measurement of the strain signal at B that is generated by an impact at A [see Fig. 3 (a)]. For damage as well as source location identification, it is very important to find accurately the arrival time of each frequency component comprising a strain signal. Before applying either STFT or CWT on the measure signal, it is necessary to analyze the phase of the measure strain signal. In this section we determine theoretically the phase of the measured strain signal generated by an impact.

Since it can be assumed that the impact at A generates an ideal impulse, the Fourier transform $\hat{F}$ of the impact force $F(t)$ is approximately constant. Using Eqs. (24)–(25) and the local balance condition, one obtains

$$\psi(x,t)=\frac{\hat{F}}{2EI(k_1^2-k_2^2)}[e^{-ik_1x}-e^{-ik_2x}]e^{i\omega t}. \tag{29}$$

For frequencies below 30 kHz, k_2 is purely imaginary and thus the contribution of e^{-ik_2x} to ψ is negligible at some distance away from the impact location. Furthermore, one can approximate in this frequency range

$$k_2\sim ik_1 \quad (k_1 \text{ is positive real valued}). \tag{30}$$

Therefore, we may replace Eq. (29) by the following equation in order to simplify the subsequent analysis:

$$\psi(x,t)\sim\frac{\hat{F}}{4EIk_1^2(\omega)}e^{-i(k_1x-\omega t)}. \tag{31}$$

The captured signal at location B that is apart by L from A is now expressed as

$$\epsilon^B(t)=r\frac{d\psi}{dx}\bigg|_{x=L}=\frac{r\hat{F}}{4EIk(\omega)}e^{-i(kL-\omega t+\pi/2)}, \tag{32}$$

where k in Eq. (32) implies k_1. Note that $\epsilon^B(t)$ in Eq. (32) actually denotes the contribution of a wave component having a circular frequency ω. The radius of the beam is denoted by r in Eq. (32).

If the frequency of the strain signal arriving at $t=t_a$ is denoted by $\omega(t_a)$, the corresponding phase $\phi(t_a)$ and magnitude $a(t_a)$ of the strain in the form of Eq. (11) can be written as

$$\phi(t_a)=\omega(t_a)t_a-k(\omega(t_a))L-\frac{\pi}{2}, \tag{33a}$$

$$a(t_a)=\frac{r\hat{F}}{4EIk(\omega)}. \tag{33b}$$

The time derivatives of the magnitude $a(t)$ are given by

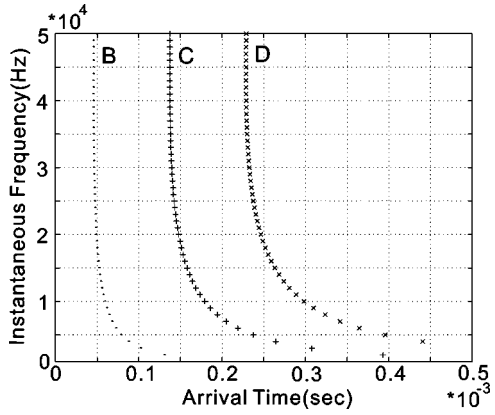


FIG. 4. The instantaneous frequency as the function of the arrival time in a beam shown in Fig. 3. (The Timoshenko beam theory is used.)

$$\dot{a}(t)=\frac{-r\hat{F}}{4EIk^2}\frac{t}{L}\dot{\omega}, \tag{34a}$$

$$\ddot{a}(t)=\frac{r\hat{F}}{4EI}\left[\frac{2}{k^3}\frac{t^2}{L^2}\dot{\omega}-\frac{1}{k^2}\frac{1}{L}\dot{\omega}-\frac{1}{k^2}\frac{t}{L}\ddot{\omega}\right]. \tag{34b}$$

The instantaneous frequency $d\phi(t_a)/dt_a$ of the measured strain becomes

$$\frac{d\phi(t_a)}{dt_a}=\frac{d\omega(t_a)}{dt_a}t_a+\omega(t_a)-\left.\frac{dk}{d\omega}\right|_{\omega=\omega_a}\frac{d\omega(t_a)}{dt_a}L$$

$$=\frac{d\omega(t_a)}{dt_a}\left(t_a-\left.\frac{dk}{d\omega}\right|_{\omega=\omega_a}L\right)+\omega(t_a). \tag{35}$$

Since the arrival time of a wave component having $\omega=\omega_a$ can be expressed as

$$t_a=\frac{L}{C_g(\omega_a)}=\left.\frac{L}{d\omega/dk}\right|_{\omega=\omega_a}=\left.\frac{dk}{d\omega}\right|_{\omega=\omega_a}L, \tag{36}$$

Eq. (35) can be simplified as

$$\frac{d\phi(t_a)}{dt_a}=\omega(t_a). \tag{37}$$

Dropping the subscript a in Eq. (37), we obtain

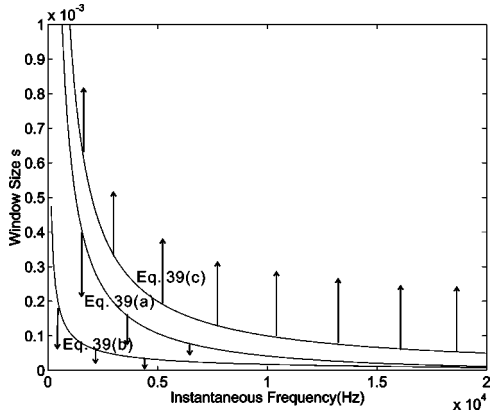


FIG. 5. The allowable window size for the applicability of STFT. (No window size satisfies all the necessary condition for the frequency range of interest.)

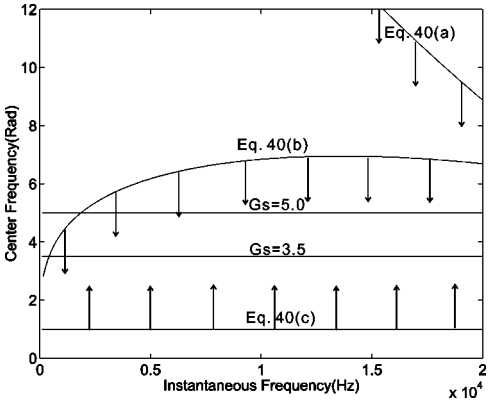


FIG. 6. The allowable range of the center frequency η for CWT.

$$\dot{\phi}(t)=\omega(t), \tag{38a}$$

$$\ddot{\phi}(t)=\dot{\omega}(t), \tag{38b}$$

where the instantaneous frequency $\omega(t)$ of the arriving signal in Eq. (35) must be found by solving Eq. (36) for a given arrival time.

Figure 4 plots the theoretical instantaneous frequencies $\dot{\phi}(t)$ as the function of the arrival time t at locations B, C, and D, which are 10, 30, and 50 cm away from the impact site A.

D. The applicability of STFT and CWT

To investigate the applicability of STFT and CWT to the problem in consideration, it is remarked that the bending wave components are accurately measured below 30 kHz because of the excitation and measurement method employed. Thus, we are mainly concerned with the instantaneous frequency components in the range between 5 and 20 kHz.

We propose to assess the applicability of STFT and CWT by examining how well the condition given by Eq. (20) for STFT and the condition given by Eq. (23) for CWT are satisfied. The theoretical analysis below will clearly show the justification for the choice of CWT over STFT for the problem in consideration.

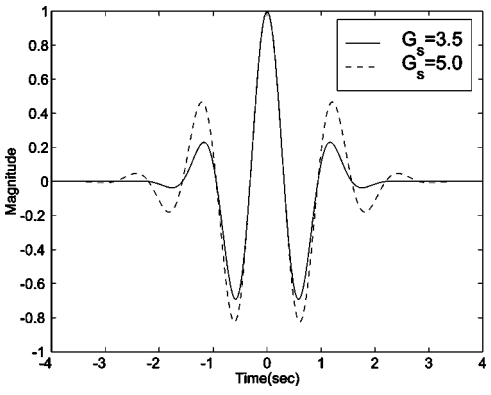


FIG. 7. The real parts of the Gabor wavelets with $G_s=3.5$ and $G_s=5.0$ (Morlet wavelet).

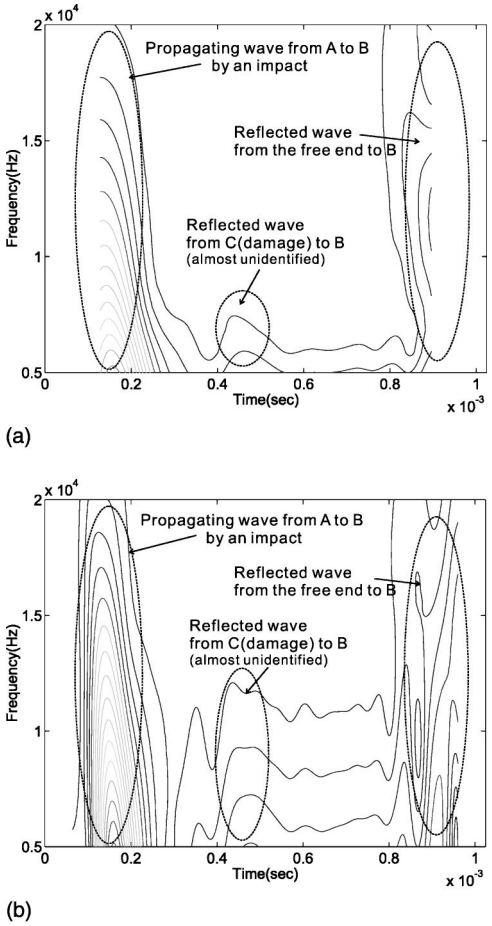


FIG. 8. The spectrogram for the strain signal measured at B in a beam shown in Fig. 3. (a) $s=256\times10^{-6}$ s; (b) $s=128\times10^{-6}$ s.

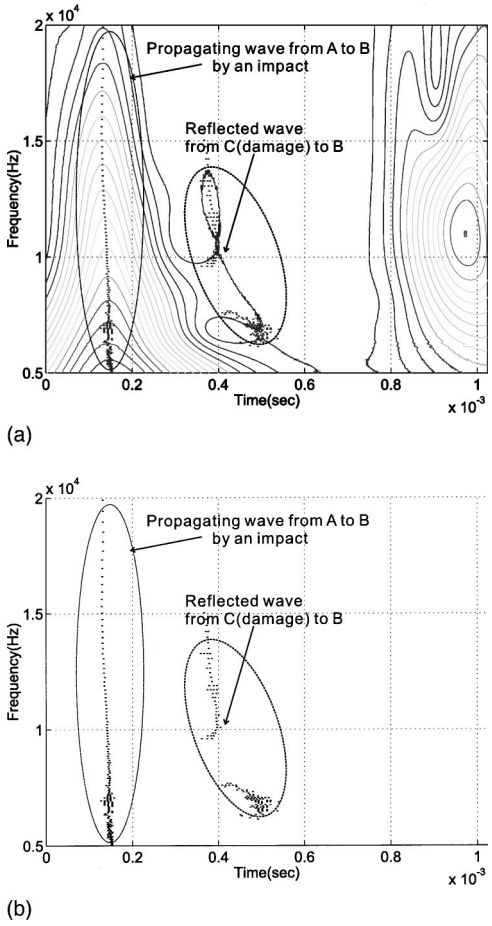


FIG. 9. The scalogram for the strain signal measured at B in a beam shown in Fig. 3 with $G_s=5.0$. (a) The normalized scalogram; (b) the corresponding ridges.

1. STFT

In order to extract the instantaneous frequency accurately from the ridges of STFT, the conditions stated by Eqs. (20) must be met. To look for the size s of a window $g(t)$, Eqs. (20) are rewritten as

$$s \ll \sqrt{\frac{|a(t_a)|}{|\ddot{a}(t_a)|}}, \tag{39a}$$

$$s \ll \sqrt{\frac{1}{|\ddot{\phi}(t_a)|}}, \tag{39b}$$

$$s \gg \frac{1}{|\dot{\phi}(t_a)|}, \tag{39c}$$

where the bandwidth $\Delta\omega$ of $\hat{g}(\omega)$ is assumed to be $O(1)$. (For instance, $\Delta\omega=1.44$ for the Hanning window whose support is $[-\frac{1}{2},\frac{1}{2}]$.)

Figure 5 shows the three curves representing Eqs. (39), where the abscissa is the instantaneous frequency. It is clear that there is no window size s satisfying simultaneously the conditions given by Eqs. (39) for the frequency range of interest. As a result, STFT cannot accurately trace the time-varying instantaneous frequencies of the strain signal.

2. CWT

To obtain satisfactory results from CWT, the conditions stated by Eqs. (23) must be fulfilled. To look for the center frequency η satisfying Eqs. (23), these equations are rewritten as

$$\eta \ll \sqrt{\frac{|a(u)||\dot{\phi}(u)|^2}{|\ddot{a}(u)|}}, \tag{40a}$$

$$\eta \ll \frac{|\dot{\phi}(u)|}{\sqrt{|\ddot{\phi}(u)|}}, \tag{40b}$$

$$\eta \gg 1. \tag{40c}$$

When a Gabor wavelet is employed for CWT, we may set $\sigma=1$ without the loss of generality and thus η is equal to G_s.

Using Eqs. (40), one can plot Fig. 6 that shows the allowable region for the center frequency η of the mother wavelet. Figure 6 clearly shows the advantage of CWT: unlike STFT, there exists a range of η satisfying all the necessary conditions that must be satisfied for successful applications of CWT in the extraction of instantaneous frequencies. The effect of the values of G_s on the performance of CWT is clear; the smaller η is, the larger the range of instantaneous frequencies to be analyzed becomes. However, G_s must be

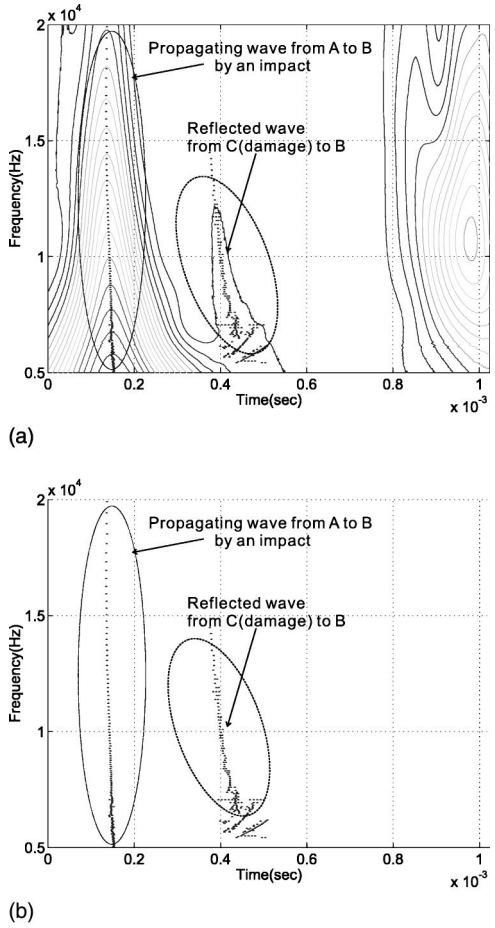


FIG. 10. The scalogram for the strain signal measured at B in a beam shown in Fig. 3 with $G_s=3.5$. (a) The normalized scalogram; (b) the corresponding ridges.

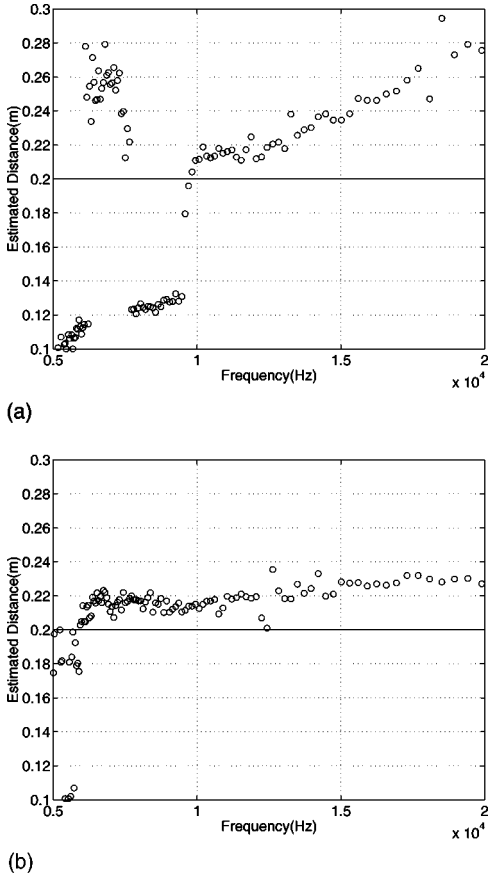


FIG. 11. Estimated distance from the measurement location B to the damage location using the ridges of the normalized scalogram ($d_{\text{exact}}=0.2$ m). (a) With $G_s=5.0$ ($d_{\text{estimated}}=0.235$ m); (b) with $G_s=3.5$ ($d_{\text{estimated}}=0.212$ m).

sufficiently large enough not to violate the admissibility condition stated by Eq. (7). The minimum numerical value of G_s that satisfies the admissibility condition is approximately 3.5.

Figure 7 compares the real parts of the Gabor wavelets with $G_s=3.5$ and $G_s=5.0$. The shorter time support for $G_s=3.5$ is advantageous for better time localization. The effective support of the Gabor wavelet is taken as $[-4,4]$ for actual numerical implementation. Most existing investigations[11–14] have used the Morlet wavelet, which is the Gabor wavelet with $G_s=5.0$. However, Kim and Kim[10] have proposed to use the Gabor wavelet with $G_s=3.5$ for damage detection and illustrated the advantage of using a smaller value of G_s.

The key message of the present theoretical analysis is that (1) there are certain types of wave signals that must be analyzed by CWT and, that (2) choosing appropriate wavelets affect significantly the performance of CWT.

E. Experimental verification

To test the performance of STFT and CWT, the near-field signal measured at location B (see Fig. 3) is analyzed by STFT and CWT. The beam in consideration has a small cut at C.

Figure 8 shows the spectrogram with (a) $s=256\times10^{-6}$ and (b) $s=128\times10^{-6}$. Because the conditions stated by Eq. (39) is not simultaneously satisfied with any value of the window size s, it is difficult to extract instantaneous frequency information accurately from the result by STFT. In particular, it is almost impossible to tell the presence of the damage from the spectrogram.

On the other hand, the application of CWT to the same signal gives quite satisfactory results: see Figs. 9 and 10. It is clear that the reflected wave from the damaged location B, whose magnitude is quite small, can be now detected by CWT. Although CWT with $G_s=5.0$ can capture the reflected wave from the damage, it is obvious that CWT with $G_s=3.5$ captures the reflected wave more accurately. To see the effect of the value of G_s (or η) on the analysis accuracy, we compare the predicted distances from B to the damage location C in Fig. 11.

In order to estimate the distance d, we used

$$d(\omega)=C_g(\omega)\frac{\Delta t(\omega)}{2}, \tag{41}$$

where Δt is the difference between the arrival times of the propagating wave from A and the first reflected wave. Note that the group velocity $C_g(\omega)$ is found from the experimentally measured strain signals. Though d must be independent of frequency, there always exists the scattering of d over frequencies.

The rms (root-mean-square) values of the estimated distance d were obtained using the data in the frequency range

between 10 and 20 kHz. As expected, the use of $G_s = 3.5$ yields a better estimate for the distance between B and C. It is worth noting that scattering near the low frequency in Fig. 11(a) is quite severe. This is because not all the conditions stated by Eqs. (40) are satisfied for $G_s = 5.0$ in this frequency range: see Fig. 6.

IV. CONCLUSIONS

In this paper, we have theoretically showed that there are some dispersive waves whose time–frequency distributions must be analyzed by CWT. The specific problem in consideration was an elastic flexural wave generated by an impact in a solid circular cylinder. During this investigation, we have also showed the importance of choosing appropriate wavelets to enhance the performance of CWT. The experimental results confirmed the present theoretical analysis.

[1] T. Önsay and A. G. Haddow, ''Wavelet transform analysis of transient wave propagation in a dispersive medium,'' J. Acoust. Soc. Am. **95**, 1441–1449 (1994).
[2] L. Carin, L. B. Felsen, D. Kralj, S. U. Pillai, and W. C. Lee, ''Dispersive modes in the time domain: Analysis and time–frequency representation,'' IEEE Microwave Guid. Wave Lett. **4**, 23–25 (1994).
[3] L. Carin, L. B. Felsen, D. Kralj, H. S. Oh, W. C. Lee, and S. U. Pillai, ''Wave-oriented signal processing of dispersive time–domain scattering data,'' IEEE Trans. Antennas Propag. **45**, 592–600 (1997).
[4] A. Abbate, J. Koay, J. Frankel, S. C. Schroeder, and P. Das, ''Application of wavelet transform signal processor to ultrasound,'' in *Ultrasonics Symposium*, 1994.
[5] A. Abbate, J. Frankel, and P. Das, ''Wavelet transform signal processing for dispersion analysis of ultrasonic signals,'' in *Ultrasonics Symposium*, 1995.
[6] D. Gabor, ''Theory of communication,'' J. IEE **93,** 429–457 (1946).
[7] E. P. Wigner, ''On the quantum correction for thermodynamic equilibrium,'' Phys. Rev. **40**, 749–759 (1932).
[8] C. H. Hodges, J. Power, and J. Woodhouse, ''The use of the sonogram in structural acoustics and an application to the vibrations of cylindrical shells,'' J. Sound Vib. **101**, 203–218 (1985).
[9] T. J. Wahl and J. S. Bolton, ''The application of the Wigner distribution to the identification of structure-borne noise components,'' J. Sound Vib. **163**, 101–122 (1993).
[10] Y. Y. Kim and E. H. Kim, ''A new damage detection method based on a wavelet transform,'' in *Proceedings of 18th IMAC*, 2000, pp. 1207–1212.
[11] K. Kishimoto, H. Inoue, M. Hamada, and T. Shibuya, ''Time frequency analysis of dispersive waves by means of wavelet transform,'' ASME Trans. J. Appl. Mech. **62**, 841–846 (1995).
[12] H. Inoue, K. Kishimoto, and T. Shibuya, ''Experimental wavelet analysis of flexural waves in beams,'' Exp. Mech. **36**, 212–217 (1996).
[13] L. Gaul and S. Hurlebaus, ''Identification of the impact location on a plate using wavelets,'' Mech. Syst. Signal Process. **12**, 783–795 (1998).
[14] H. Jeong and Y.-S. Jang, ''Fracture source location in thin plate using the wavelet transform of dispersive waves,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **47**, 612–619 (2000).
[15] S. Mallat, *A Wavelet Tour of Signal Processing* (Academic, New York, 1998).
[16] D. E. Newland, ''Harmonic wavelets in vibrations and acoustics,'' in *Royal Society Discussion Meeting*, 1999.
[17] N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinent, P. Tchamitchian, and B. Torrésani, ''Asymtotic wavelet and Gabor analysis: Extraction of instantaneous frequencies,'' IEEE Trans. Inf. Theory **38**, 644–664 (1992).
[18] L. Pochhammer, ''Über die Fortpflanzungsgeschwindigkeiten kleiner Schwingungen in einem unbegrenzten istropen kreiszylinder,'' J. Reine Angew. Math. **81**, 324–336 (1976).
[19] C. Chree, ''The equation of an isotropic elastic solid in polar and cylindrical coordinates, their solutions and applications,'' Trans. Cambridge Philos. Soc. **14**, 250 (1889).
[20] Y. Y. Kim, Ph.D. thesis, Stanford University, 1988.
[21] S. P. Timoshenko and J. N. Goodier, *Theory of Elasticity*, 3rd ed. (McGraw-Hill, New York, 1970).
[22] J. D. Achenbach and R. J. Brind, ''Scattering of surface waves by a sub-surface crack,'' J. Sound Vib. **76**, 43–56 (1981).
[23] J. Miklowitz, *The Theory of Elastic Waves and Waveguides* (North-Holland, New York, 1978).
[24] J. F. Doyle, *Wave Propagation in Structures*, 2nd ed. (Springer-Verlag, New York, 1997).

Numerical modeling of finite-amplitude sound beams: Shock formation in the near field of a cw plane piston source

V. A. Khokhlova
Department of Acoustics, Physics Faculty, Moscow State University, Moscow 119899, Russia

R. Souchon
INSERM–Unité 281, 151 Cours Albert Thomas, 69424 Lyon Cedex 03, France

J. Tavakkoli
Institute of Biomaterials and Biomedical Engineering, University of Toronto, 4 Taddle Creek Road, Toronto, Ontario M5S 3G9, Canada

O. A. Sapozhnikov
Department of Acoustics, Physics Faculty, Moscow State University, Moscow 119899, Russia

D. Cathignol
INSERM–Unité 281, 151 Cours Albert Thomas, 69424 Lyon Cedex 03, France

(Received 15 May 2000; accepted for publication 5 March 2001)

Two theoretical models and the corresponding numerical codes for the description of nonlinear acoustic beams radiated from intense cw sources in water are presented. In the first model, diffraction effects are included using the Rayleigh integral, whereas nonlinearity and thermoviscous absorption are accounted for in a quasi-plane approximation. The simulations are performed in the time domain using the code previously developed for single-pulse propagation in medium having arbitrary frequency-dependent absorption. The second model is based on the Khokhlov–Zabolotskaya–Kuznetsov equation, which, contrary to the first model, accounts for diffraction in the parabolic approximation. The simulations are performed in the frequency domain using a novel algorithm that has been developed. A variable number of harmonics, which follows the nonlinear broadening of the wave spectrum are employed in the algorithm to speed up calculations. In order to prove the validity and the accuracy of the two codes developed, the simulation of diffraction and nonlinear effects in the near field of an intense ultrasound circular piston source in water is performed. The results of modeling obtained by both codes are compared with each other and with known experimental data, and are found to be in a good agreement. Frequency-domain code is then used for detailed study of the strongly nonlinear regime of propagation, when shocks are developed in the waveform close to the source. It is demonstrated that diffraction plays a major role in shock formation. Development of two shocks in each cycle and their further collision is predicted. It is also shown that nonlinear propagation and shock formation result at some distance in the two times excess of peak positive pressure in comparison with the maximum value obtained in the case of linear propagation. The beam total power decay due to formation of shocks as a function of the propagation distance is compared with the intensity in a plane wave propagation without diffraction. It is shown that nonlinear energy decay starts earlier for the beam, but decreases slower over longer distances. [DOI: 10.1121/1.1369097]

PACS numbers: 43.25.Cb, 43.25.Jh [MAB]

I. INTRODUCTION

Investigation of acoustic fields radiated from intense sources requires adequate description of nonlinear, absorption, and diffraction phenomena. Most of the problems important for practical applications, for example, in medicine, cannot be solved analytically, and numerical simulation is needed. Various computational algorithms have been developed in the last two decades to simulate acoustic fields of the real sources, related to the problems of hydroacoustics,[1–4] intense noise,[5] the sonic boom problem,[6] and medical acoustics.[7–11] Numerical solutions are available either in the time[3,6–9] or frequency domain.[1,2,4,10–12] Moreover, some algorithms operate utilizing a combination of the time and frequency domain approaches.[3,5,13] An overview of the modern computational techniques related to the description of intense acoustic fields can be found in a recent book on nonlinear acoustics.[14]

The algorithms become very time consuming at high intensities, when a very fine time grid or a large number of harmonics are required to model thin shocks developed in the waveform due to acoustic nonlinearity. To reduce computational time, various approaches based on an *a priori* known spatial or temporal property of the acoustic field have been proposed lately. Optimization of the spatial grid that follows the geometry of the beam, either focused or unfocused, enables us to decrease the number of spatial samples necessary for calculations.[7,15–18] A spherically convergent or divergent grid is employed,[7,15,16] as well as the transforma-

tion of spatial and temporal coordinates based on the analytic result for the linear Gaussian beam.[17–19] Artificial absorption within the thin layer close to the edge of the lateral numerical grid permits us to reduce the spatial window in the lateral coordinates and to avoid reflections from the edge of the grid.[16] To smoothen the shock structure and thus to reduce the number of harmonics retained in the algorithms, an artificial viscosity rapidly increasing with frequency is included.[10,13] Contrary to the smoothening of the shocks, asymptotic approaches have been used in which the shock front of finite thickness is replaced by a discontinuity. For this purpose, the weak shock theory can be used in the time domain[5,20] or a known high-frequency asymptotic result for shocks can be incorporated in the frequency domain algorithm.[11,17,21–23] This problem becomes of particular complexity when two- or three-dimensional[24,25] waves are considered.

A new time-domain code has been developed recently for simulation of the focused nonlinear beams in a pulsed regime.[9] Diffraction, nonlinearity, arbitrary frequency-dependent absorption, and dispersion are taken into account in the code. These effects are treated independently using the method of fractional steps with a second-order operator-splitting algorithm. The main advantages of the model compared with the existing time-domain methods are an arbitrary frequency-dependent absorption and full diffraction formulation, the latter being particularly important for studying intense sources with high focusing gains. A good agreement has been demonstrated between the results of simulations and experiments for a focused pulsed piezoelectric source used in lithotripsy and ultrasound surgery experimental research. In this article the code is applied to the description of the cw field.

As an alternative numerical approach, a newly developed frequency-domain algorithm based on the KZK equation[26,27] is presented in this article. It is known that the frequency domain approach works well at low intensities, when the wave spectrum consists of a small number of harmonics. As it was previously mentioned, for higher intensities, when shock fronts are developed in the waveform, the number of spectral components becomes bigger and the simulations are very time consuming. Application of an artificial viscosity for higher frequencies reduces the number of harmonics necessary to be retained in the algorithm, but makes it impossible to adequately describe the fine structures of the waveform such as shock fronts and possible spikes. The main feature of the code developed is that it operates without artificial viscosity and is optimized to reduce the calculation time in the presence of shocks.

One of the goals of the present article is to validate the new frequency-domain code and the extension of the time-domain code to cw regime. The validation is performed via comparison of the numerical results obtained by the two codes with each other and with available experimental data. As a model problem to study numerically, a near field of a cw circular plane piston source is investigated here. Due to diffraction, the structure of the acoustic field is highly oscillating in space and nonlinear interactions take place quite differently compared with the case of one-dimensional plane wave propagation. This classical problem has been studied in various papers mostly in relatively weak nonlinear regimes without shocks in the waveform.[2,4,28,29] Furthermore, stronger nonlinear effects, when the shock fronts are developed within the near field of the piston, were investigated both experimentally and numerically.[30] The experimental data from that paper are taken as a benchmark result to compare with the numerical simulations of the present work. The study of Ref. 30 is limited to the axial and lateral behavior of the first several harmonics of the fundamental frequency. However, the evolution of the waveform itself, in particular, the formation and further propagation of the shocks, may be of more interest in this regime. Other characteristics, such as peak positive and negative pressures, spatial distribution of wave intensity, and total acoustic power decay due to absorption at the shocks, are also of interest but not studied yet.

The article consists of three parts. Two numerical models and corresponding codes are presented in the first part. In the second part, the results of numerical simulations are compared with each other and with the experimental data. Finally, a frequency-domain algorithm is employed to study specific features of the waveform distortion and evolution of shocks in the near field of the cw plane piston source.

II. NUMERICAL MODELS

In this study, the existing experimental data are used for comparison with the numerical results. The experimental data are for a circular plane piston source with an aperture radius $a=2.35$ cm operating at a frequency of $f_0=1$ MHz. Acoustic waveforms were measured by a broadband PVDF membrane hydrophone in water as the axial distances from the source varied from 5 to 40 cm, and the lateral distances varied up to 5 cm. The amplitudes of the first five harmonics were registered in that experiment.[30] The following parameters for water as the propagation medium are used here for numerical simulations: sound speed $c_0=1500$ m/s, ambient density $\rho_0=1$ g/cm^3, nonlinear parameter $\beta=3.5$, absorption coefficient at 1 MHz $\alpha_0=b\omega_0^2/2c_0^3\rho_0=2.533\times10^{-4}$ cm^{-1}, and $\omega_0=2\pi f_0$ is the angular frequency, which corresponds to the value of dissipative parameter $b=4.33\times10^{-2}$ g·s^{-1} cm^{-1}.

For the purpose of simulations and further analysis of the results, the following dimensionless variables are introduced: particle velocity $V=u/u_0$, where u_0 is the normal velocity amplitude of the source surface, acoustic pressure $P=p/p_0$, where $p_0=u_0c_0\rho_0$ is the characteristic initial pressure amplitude, time $T=\omega_0\tau$ where $\tau=t-z/c_0$ is the retarded time, and wave number $K=ka$. Propagation distance along the axis z is normalized by the Rayleigh distance for the fundamental frequency as $Z=z/z_d$, where $z_d=ka^2/2=115.6$ cm, and $k=\omega_0/c_0$. Lateral coordinate r is normalized by the source radius: $R=r/a$. Finally, two more dimensionless parameters N (nonlinearity) and A (absorption) are defined as

$$N=\frac{z_d}{z_n}=\frac{\beta\omega_0^2a^2p_0}{2c_0^4\rho_0} \quad \text{and} \quad A=\frac{z_d}{z_a}=\frac{b\omega_0^3a^2}{4c_0^4\rho_0}.$$

The parameters N and A relate different characteristic lengths to each other. N relates the Rayleigh distance z_d to the shock formation distance for the case of plane wave propagation where $z_n = c_0^3 \rho_0 / \beta \omega_0 p_0$, and A relates z_d to the absorption length $z_a = \alpha_0^{-1}$. In accordance with the given source amplitude in experiments $p_0 = 2$, 6.4, and 14.3 bars (1 bar$\cong$0.1 MPa),[30] the shock formation distances are $z_n = 76.7$, 24, and 10.7 cm, and the values of the dimensionless nonlinear parameter are $N = 1.5$, 4.8, and 10.8, respectively. The dimensionless absorption parameter value is $A = 0.0293$. The dimensionless wave number is $K = 98.4$.

A. Time-domain code

1. Basic equations

An acoustic wave radiated from a plane source is considered as propagating in the z direction in the sense that the wave parameters at any plane $z = z_2$ can be calculated from those given at some precedent plane $z = z_1 < z_2$. In the limiting case of an infinitesimal distance $\Delta z = z_2 - z_1$, a differential equation of evolution type can therefore be written in the time domain:

$$\frac{\partial u}{\partial z} = \hat{L} u. \tag{1}$$

Here $u = u(x, y, z, \tau)$ is the wave variable, for instance the z component of the particle velocity, x and y are the transverse coordinates, τ is the retarded time, and the operator $\hat{L}$ accounts for various physical effects changing the waveform. In the presence of multiple effects, if each is fairly weak, the operator $\hat{L}$ can be expressed as a sum of corresponding ''one-effect'' operators $\hat{L}_i$ as $\hat{L} u = \sum_i \hat{L}_i u$. When nonlinearity, absorption, and diffraction are of importance, the evolution equation has the following general form:

$$\frac{\partial u}{\partial z} = \hat{L}_N u + \hat{L}_A u + \hat{L}_D u, \tag{2}$$

where the operators $\hat{L}_N$, $\hat{L}_A$, and $\hat{L}_D$ account for the corresponding effects. For instance, in the case of a plane progressive wave traveling in thermoviscous quadratically nonlinear medium, Eq. (2) reduces to the Burgers equation

$$\frac{\partial u}{\partial z} = \hat{L}_N u + \hat{L}_A u, \tag{3}$$

where $\hat{L}_N$ and $\hat{L}_A$ are nonlinear and absorption operators:

$$\hat{L}_N u = \frac{\beta}{c_0^2} u \frac{\partial u}{\partial \tau}, \tag{4}$$

$$\hat{L}_A u = \frac{b}{2 \rho_0 c_0^3} \frac{\partial^2 u}{\partial \tau^2}. \tag{5}$$

If the wave is not plane, diffraction plays a role in changing the acoustic field. In an ideal linear medium, the z component of the particle velocity, u, and the acoustic pressure, p, at the current z plane can be obtained by the Rayleigh integral from the distribution of the wave variable u over any precedent plane:[9]

$$u(\mathbf{r}, t) = \frac{1}{2\pi} \int \int_S \left\{ \left[\frac{1}{c_0 R} \cdot \frac{\partial u(\mathbf{r}', t - R/c_0)}{\partial t} + \frac{u(\mathbf{r}', t - R/c_0)}{R^2} \cos \gamma \right] \right\} ds', \tag{6}$$

$$p(\mathbf{r}, t) = \frac{\rho_0}{2\pi} \int \int_S \frac{\partial u(\mathbf{r}', t - R/c_0)}{\partial t} \frac{ds'}{R}. \tag{7}$$

Here the vector $\mathbf{r}'$ represents coordinates of the surface element ds' in the plane surface S, $R = |\mathbf{r} - \mathbf{r}'|$ is the distance between ds' and the observation point, γ is the angle between vector $\mathbf{R} = \mathbf{r} - \mathbf{r}'$ and the z axis, and therefore $\cos \gamma = \partial R / \partial z$. The diffraction operator in the corresponding evolution equation $\partial u / \partial z = \hat{L}_D$ can be derived directly from Eq. (6). The expression for $\hat{L}_D$ is not presented here because the direct use of the Rayleigh integral is more suitable in numerical modeling.

The presented model is based on the evolution equation (2) with the operators $\hat{L}_N$, $\hat{L}_A$, and $\hat{L}_D$ given by Eqs. (4)–(6). Concerning the operator $\hat{L}_A$ for absorption, a minimum-phase filter model was developed that enables us, in addition to the thermoviscous absorption, Eq. (5), which is proportional to the square of frequency, to take into account the effects of arbitrary frequency-dependent absorption and dispersion. Note that in this time-domain approach the diffraction term is exact, whereas the nonlinearity and absorption are accounted for in the approximation of plane wave propagation.

2. Numerical algorithm

The numerical time-domain algorithm for solving Eq. (2) is based on the method of fractional steps with a second-order operator-splitting procedure. The particle velocity u is calculated plane-by-plane with a step Δz_1 along the z axis. In each step, the right-hand side of Eq. (2) is split into two operators $\hat{L}_1 = \hat{L}_N + \hat{L}_A$ and $\hat{L}_2 = \hat{L}_D$, so that the plane wave propagation effects (nonlinearity and absorption) and the diffraction effect are applied sequentially. At the first fractional step, the Burgers equation $\partial u / \partial z = \hat{L}_1$ is solved from the current plane z to the next plane $z + \Delta z_1$. The obtained result is then used as a boundary condition for the second fractional step, at which the solution of the evolution equation $\partial u / \partial z = \hat{L}_2$ is calculated by applying its exact solution, Eq. (6). The integration in Eq. (6) is taken over the surface of the plane z numerically.[9] This procedure gives the particle velocity u for each point of the plane $z + \Delta z_1$. The pressure p, if needed for output, is derived from Eq. (7).

As far as the exact solution is used for the diffraction term, there is no restriction on the smallness of the integration interval Δz_1 to account for the diffraction effect accurately. The value Δz_1 therefore can be chosen fairly big if all other effects, nonlinearity and absorption, weakly contribute to the evolution of the wave over this interval. Although the exact solution is available for the Burgers equation as well,[31] it is not suitable for simulations, and the equation is solved numerically. Smaller steps thus are required to obtain stable and accurate solutions to nonlinearity and absorption over

the step Δz_1. The interval Δz_1 is divided into smaller substeps Δz_2 which are passed sequentially to get a solution over the whole step Δz_1. An additional operator splitting procedure is applied for nonlinearity $\hat{L}_N$ and absorption $\hat{L}_A$ at each substep Δz_2. The simple wave equation $\partial u/\partial z = \hat{L}_N$ is solved using an analytical implicit solution and linear interpolation procedure to keep a uniform time grid spacing.[7,9] The computational step Δz_2 is taken small enough to avoid multivalued solutions caused by nonlinear waveform steepening. The absorption equation is solved in the time domain using the standard convolution procedure with the impulse response of a minimum-phase digital filter.

To achieve the second-order accuracy of the operator-splitting algorithm over the main step Δz_1, a spatial shift $\Delta z_1/2$ is introduced between steps used for the plane wave operator $\hat{L}_1 = \hat{L}_N + \hat{L}_A$ and the diffraction operator $\hat{L}_2 = \hat{L}_D$.[9]

3. Parameter values for operation

The algorithm operates in physical dimension variables. The axial step is chosen as $\Delta z_1 = 19$ mm for diffraction and $\Delta z_2 = 0.23$ mm for nonlinearity and attenuation. The surface elements ds' for covering the area for the Rayleigh integral are chosen with a constant radial step size, Δr. The angular step, $\Delta\Phi$, is chosen so that the width of each element is approximately equal to the step Δr. The radial step is taken smaller for higher nonlinearity to accurately account for higher frequencies involved in the calculation of the Rayleigh integral. This leads to $\Delta r = 47\ \mu$m (500 points per piston radius, i.e., about 4×10^5 surface elements ds' on the piston) for the weakly nonlinear case $N = 1.5$, up to $\Delta r = 11.75\ \mu$m (2000 points per piston radius, i.e., about 6×10^6 surface elements) for the case with the strongest nonlinearity, $N = 10.8$. Throughout all calculations, the time window is set to 6 μs, representing six cycles sampled at $\Delta t = 5.86$ ns (1024 points). The radius $r_{\max}$ at each step Δz_1 varies as described in a previous paper,[9] so that the points farther than $r_{\max}$ from the axis have no effect on the signal in the time window considered (due to the finite speed of sound propagation). The value of $r_{\max}$ therefore varies between one and three times the piston radius, a. The amplitudes of harmonics are calculated using the fast Fourier transform (FFT) of one cycle time window. The time window is shifted point by point over the stable part of the signal and the results for the amplitudes obtained are averaged. In order to reduce the noise content of the signals, the first window starts 30 samples after the end of the transient part of the waveform. The transient time is derived from the difference of the shortest and longest distance between the point considered and the piston source.

B. Frequency-domain code

1. Basic equations

In this approach propagation of an intense acoustic beam radiated by the piston circular source is modeled by the Khokhlov–Zabolotskaya–Kuznetsov (KZK) equation,[26,27] which can be written in terms of the axial component of the particle velocity as

$$\frac{\partial}{\partial\tau}\left(\frac{\partial u}{\partial z} - \frac{\beta}{c_0^2}u\frac{\partial u}{\partial\tau} - \frac{b}{2c_0^3\rho_0}\frac{\partial^2 u}{\partial\tau^2}\right) = \frac{c_0}{2}\Delta_\perp u. \tag{8}$$

Here $\Delta_\perp = \partial^2/\partial r^2 + r^{-1}\partial/\partial r$ is the transversal Laplacian. This equation is also of the evolution type, as is Eq. (2), with the same nonlinear and absorption terms included in the time-domain approach, which has been given earlier in Eqs. (4) and (5), but with different diffraction operator. From Eq. (8), the evolution equation for diffraction can be written as

$$\hat{L}_D u = \frac{c_0}{2}\Delta_\perp \int^\tau u\, d\tau', \tag{9}$$

which accounts for diffraction in parabolic approximation. Such an approximation is known to give fairly accurate results for quasi-plane waves (acoustic beams).[32] Equation (8) in terms of dimensionless variables introduced earlier can be written as

$$\frac{\partial}{\partial T}\left(\frac{\partial V}{\partial Z} - NV\frac{\partial V}{\partial T} - A\frac{\partial^2 V}{\partial T^2}\right) = \frac{1}{4}\Delta_\perp V. \tag{10}$$

Here $V = u/u_0$, acoustic pressure p can be calculated in a parabolic approximation from the particle velocity u as $p = c_0\rho_0 u$, and the Laplacian operator is introduced now in terms of dimensionless transversal coordinate R as $\Delta_\perp = \partial^2/\partial R^2 + R^{-1}\partial/\partial R$. In the frequency domain we seek a solution to Eq. (10) in a form of the Fourier series expansion

$$V(Z,R,T) = \sum_{n=-\infty}^{\infty} C_n(Z,R)\exp(-inT), \tag{11}$$

where $C_n(Z,R)$ is the complex amplitude of the nth harmonic ($-\infty < n < \infty$). Substitution of the solution (11) into Eq. (10) yields a set of coupled nonlinear differential equations for the complex amplitudes C_n:

$$\frac{\partial C_n}{\partial Z} = -\frac{in}{2}\sum_{k=-\infty}^{\infty} C_k C_{n-k} - An^2 C_n + \frac{i}{4n}\Delta_\perp C_n, \tag{12}$$

where $C_{-n} = C_n^*$, and C_n^* denotes the complex conjugate of C_n. The first term on the right-hand side of Eq. (12) is a convolution which accounts for the nonlinear interaction of harmonics. The second term governs thermoviscous absorption proportional to the square of frequency. Note that implementation of the spectral approach enables us to easily extend the model to include arbitrary frequency-dependent absorption and dispersion. Diffraction effects are accounted for by the third term.

The numerical solution for harmonics C_n can be used to study the main characteristics of the acoustic field. The dimension values of harmonic amplitudes are expressed as $A_n = 2p_0|C_n|$. The waveform is reconstructed from Eq. (11). The mean intensity for the quasiplane wave propagation, Eq. (1), has its main component in the z direction $I = \langle p^2\rangle/c_0\rho_0$, which is the same as for a plane wave propagation.[2] In terms of complex harmonic amplitudes C_n the normalized intensity can be written as

$$\tilde{I}(Z,R) = 4\sum_{n=1}^{\infty}|C_n|^2, \tag{13}$$

where $\tilde{I}=I/I_0$ and $I_0=p_0^2/2c_0\rho_0$ is the time-averaged source intensity.

The total power of the beam at some distance z from the source is obtained by integration of the mean intensity over the surface of the corresponding plane: $W(z)=2\pi\int_0^\infty I(z,r)r\,dr$, or, in the normalized form,

$$\tilde{W}(Z)=2\int_0^\infty \tilde{I}(Z,R)R\,dR, \tag{14}$$

where $\tilde{W}=W/W_0$ and $W_0=\pi a^2 I_0$ is the total power of the source.

2. Numerical algorithm

The numerical solution of Eq. (12) is obtained for the amplitudes C_n of harmonics using an operator splitting procedure similar to that in the time-domain algorithm discussed earlier in Sec. II A. A uniform spatial grid with the step ΔR is used for sampling the solution C_n over the lateral coordinate R. For the given absorption parameter $A=0.0293$ and the biggest nonlinear parameter $N=10.8$ the number of harmonics needed for adequate description of the waveform when shock fronts are developed should be of the order of $n_{\max}\approx\pi N/A\approx 10^3$. This estimation follows from the analytic result for the width of the stationary shock wave governed by the Burgers equation.[31] Direct use of such a high number of harmonics makes the calculations very time consuming. On the other hand, the shocks are developed starting only from some axial distance from the source and mostly in the area close to the beam axis. A variable number of harmonics is used, therefore, in the algorithm for various coordinates Z and R in order to reduce the computational time.

Let $C_n^m(Z)$ be the amplitude of the nth harmonic at the axial distance Z and at the lateral distance $R_m=m\cdot\Delta R$, $m=0,1,...,m_{\max}^{(n)}$. Here the number $m_{\max}^{(n)}$ denotes the maximum number of lateral points for the nth harmonic. The amplitudes $C_n^m(Z)$ are calculated plane by plane following the propagation over the step ΔZ from the distance Z to the next location, $Z+\Delta Z$.

At the first fractional step of the operator splitting procedure, the effect of nonlinearity is accounted for. A fourth-order Runge–Kutta method is employed,[33] independently for each lateral index m, to solve the set of nonlinear coupled equations for the first $n_{\max}^{(m)}$ number of harmonics:

$$\frac{\partial C_n^m}{\partial Z}=-\frac{in}{2}\sum_{k=-n_{\max}^{(m)}}^{n_{\max}^{(m)}} C_k^m C_{n-k}^m. \tag{15}$$

The amplitudes of the harmonics with $n>n_{\max}^{(m)}$ are assumed to be equal to zero. The variable number of harmonics $n_{\max}^{(m)}(Z)$ is used according to the width of the current wave spectrum. At the beginning, $Z=0$, the waveform is smooth. A fairly small number of harmonics, $n_{\max}^{(m)}(0)$, is thus chosen for all lateral coordinates R_m. At each axial step the absolute value of the last harmonic amplitude at all lateral points m is compared with a certain small threshold value, ε. If at least one of these amplitudes exceeds the threshold ε, then the number of harmonics is increased, for all m, by Δn up to some maximum number $\tilde{n}_{\max}^{(m)}$, so that $n_{\max}^{(m)}(Z+\Delta Z)=\min\{n_{\max}^{(m)}(Z)+\Delta n,\tilde{n}_{\max}^{(m)}\}$. As far as nonlinear broadening of the spectrum is more pronounced close to the axis of the beam, the value of $\tilde{n}_{\max}^{(m)}$ is chosen differently depending on the distance off the axis, R_m.

The second fractional step is made to account for the thermoviscous absorption effect. Exact analytic result for the corresponding equation $\partial C_n^m/\partial Z=-An^2C_n^m$ is used for each lateral index m and harmonic index n independently:

$$C_n^m(Z+\Delta Z)=C_n^m(Z)\exp(-An^2\Delta Z). \tag{16}$$

The diffraction effects are accounted for at the third fractional step. The parabolic equation $\partial C_n/\partial Z=i\Delta_\perp C_n/4n$ is approximated by a finite-difference implicit backward algorithm[33] for each harmonic index n independently:

$$\frac{C_n^m(Z+\Delta Z)-C_n^m(Z)}{\Delta Z}=\frac{i}{4n\Delta R^2}\left\{\left[1-\frac{1}{2m}\right]C_n^{m-1}(Z+\Delta Z)-2C_n^m(Z+\Delta Z)+\left[1+\frac{1}{2m}\right]C_n^{m+1}(Z+\Delta Z)\right\}. \tag{17}$$

The boundary condition on the axis is specified using the axial symmetry of the beam, which yields $\partial C_n/\partial R=0$ for $R=0$. The right-hand side of Eq. (17) for $m=0$ thus can be rewritten as $i(C_n^1-C_n^0)/n\Delta R^2$. For the maximum lateral index the requirement that the amplitude of the harmonic at the next lateral point, $m_{\max}^{(n)}+1$, equals to zero is employed in Eq. (17). The value of $m_{\max}^{(n)}$ and corresponding size of the spatial window in the radial direction $R_{\max}^{(n)}=m_{\max}^{(n)}\Delta R$ are different for different harmonics. As diffraction divergence is less pronounced for higher harmonics, a smaller number of lateral points is used for them.

Finally, to avoid numerical artifact of reflection from the boundary $R_{\max}^{(n)}$, artificial absorption is introduced for the last Δm points near the boundary as an additional, fourth, fractional step:

$$C_n^m(Z+\Delta Z)=C_n^m(Z)\exp(-\alpha_n^m\Delta Z), \tag{18}$$

where absorption coefficient α_n^m is equal to zero for $m<m_{\max}^{(n)}-\Delta m$ and increases quadratically with m for the rest of the Δm points: $\alpha_n^m=\alpha_*[1+(m-m_{\max}^{(n)})/\Delta m]^2$, $m_{\max}^{(n)}-\Delta m\leqslant m\leqslant m_{\max}^{(n)}$.

The waveform at each spatial point is reconstructed from Eq. (11), where the summation is performed for harmonics with numbers $|n|\leqslant n_{\max}^{(m)}$. The mean intensity at the mth lateral point is calculated from Eq. (13) as

$$\tilde{I}^{(m)}=4\sum_{n=1}^{n_{\max}^{(m)}}|C_n^m|^2. \tag{19}$$

The mean power of the beam, Eq. (14), is calculated numerically using the standard trapezoidal method:

$$\tilde{W}=\frac{\Delta R^2}{2}\sum_{m=1}^{m_{\max}}(2m-1)\cdot(\tilde{I}^{(m-1)}+\tilde{I}^{(m)}), \tag{20}$$

where summation is performed up to $m_{\max}=\max(m_{\max}^{(n)})$.

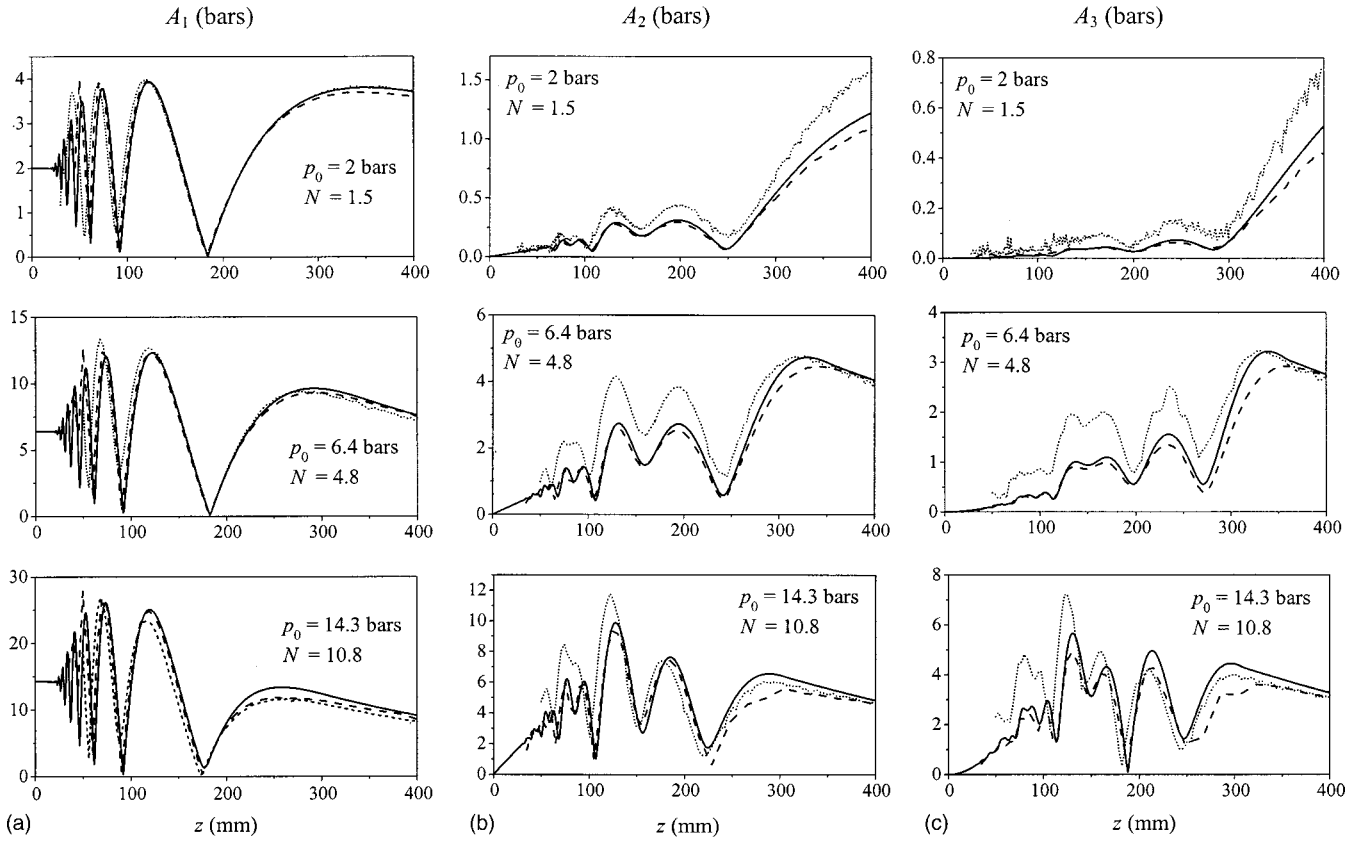


FIG. 1. Comparison of experimental data (···) and numerical results obtained by the frequency-domain (———) and the time-domain (---) codes. Axial propagation curves for the first (a), second (b), and third (c) harmonic amplitudes, A_1, A_2, and A_3, for various pressure amplitudes at the source.

3. Parameter values for operation

The lateral grid step is equal to $\Delta R = 2.5 \times 10^{-3}$, which corresponds to 400 points per piston radius. The size of the spatial window in the radial direction $R_{\max}^{(n)} = m_{\max}^{(n)} \Delta R$ is chosen depending on the harmonic's number n in the following way:

$$R_{\max}^{(n)} = \begin{cases} 5, & 1 \leqslant n \leqslant 30, \\ 4, & 31 \leqslant n \leqslant 100, \\ 3, & 101 \leqslant n \leqslant 300, \\ 2, & 301 \leqslant n \leqslant 800, \\ 1, & 801 \leqslant n \leqslant 1000. \end{cases} \tag{21}$$

The axial grid step is taken smaller for higher nonlinearity N to provide necessary stability and accuracy of the solution. The minimum step size used in the calculations is $\Delta Z = 2.5 \times 10^{-5}$. The calculations are performed up to $Z_{\max} = 0.35$, according to the range of experimental data. The initial number of harmonics equals to $n_{\max}^{(m)}(0) = 30$, independently on the index m. The number of harmonics is increased by $\Delta n = 100$ when the absolute value of the last harmonic amplitude exceeds the threshold $\varepsilon = 2 \times 10^{-4}$. The maximum possible number of harmonics $\tilde{n}_{\max}^{(m)}$ for a given radial distance R_m corresponds to Eq. (21), so that $\tilde{n}_{\max}^{(m)} = 1000$ ($0 \leqslant m \leqslant 400$), 800 ($401 \leqslant m \leqslant 800$), 300 ($801 \leqslant m \leqslant 1200$), 100 ($1201 \leqslant m \leqslant 1600$), and 30 ($1601 \leqslant m \leqslant 2000$). The boundary layer with artificial absorption for all harmonics consists of $\Delta m = 50$ points. The corresponding absorption parameter equals $\alpha_* = 500$.

III. RESULTS AND DISCUSSION

A. Comparison of numerical results and experimental data

The results of numerical simulations of the near field of the plane piston source are presented here and compared with experimental data reported in a previous paper.[30] Shown in Fig. 1(a) are dimension pressure amplitudes A_n of the first three harmonics, $n = 1,2,3$, along the acoustic beam axis for various initial pressure amplitude $p_0 = 2$, 6.4, and 14.3 bars at the source. Axial distributions of the first harmonic, Fig. 1(a), are in a good agreement for all levels of excitation. Some discrepancy between the numerical results is observed at the distances very close to the source because of different approaches used in calculating the diffraction effects. It is seen, however, that highly oscillating structure of the field is adequately governed by numerical models starting from the small distances from the source, $z \sim 60$ mm, so that the last three minimums and maximums in axial distributions of the fundamental frequency are accurately captured. The very last diffraction maximum for the fundamental frequency is positioned at the distance 25 to 35 cm from the source, closer to the transducer for a higher level of excitation. The very last diffraction minimum is positioned at 18.5 cm from the source for weakly nonlinear propagation, $p_0 = 2$ bars, and also slightly moves towards the transducer, in about 1 cm for the highest source level $p_0 = 14.3$ bars.

The characteristic shock formation length for the chosen

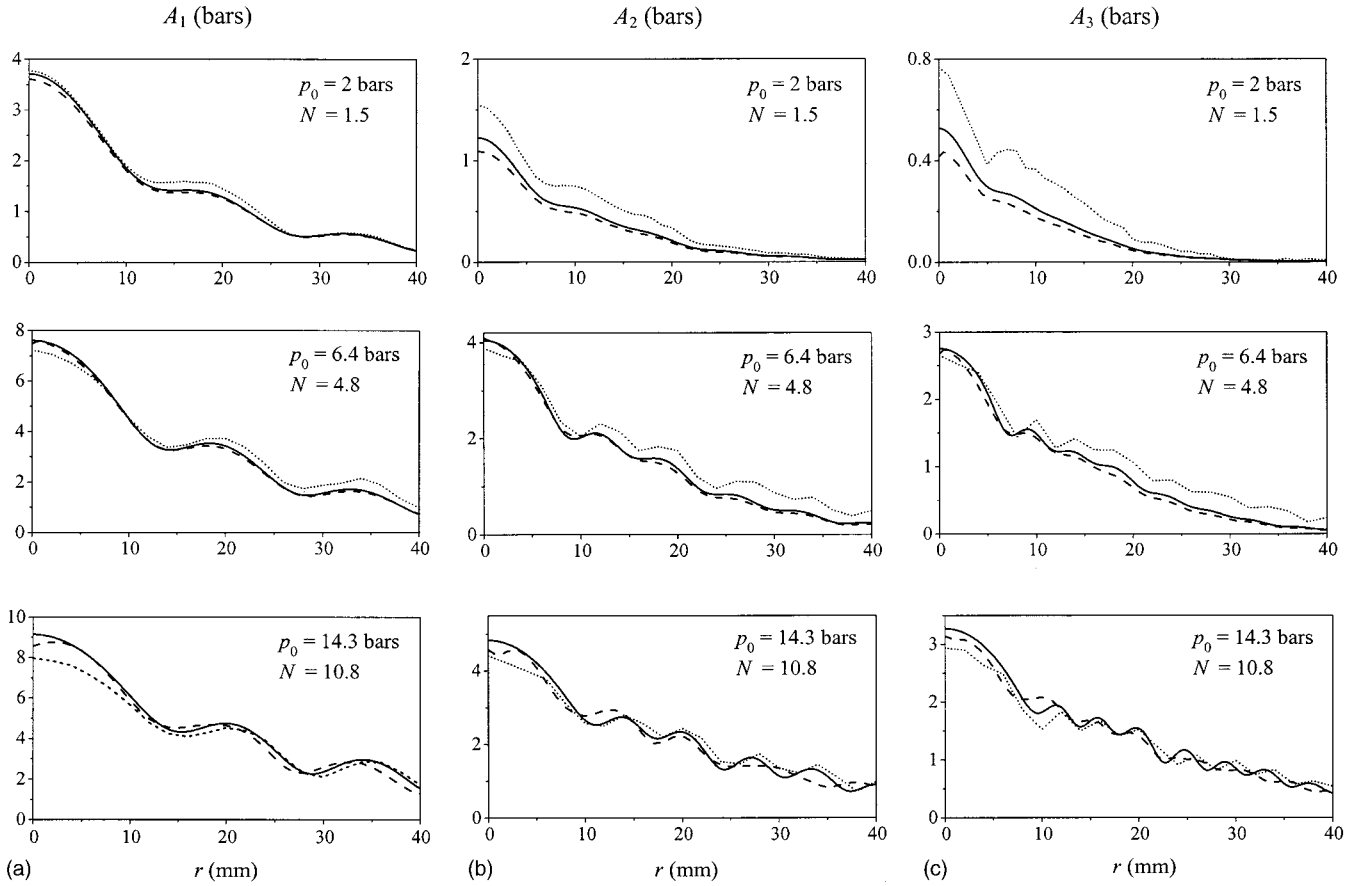


FIG. 2. Comparison of experimental data (···) and numerical results obtained by the frequency-domain (———) and the time-domain (---) codes. Radial distribution of the first (a), second (b), and third (c) harmonic amplitudes, A_1, A_2, and A_3, at $z=400$ mm for various pressure amplitudes at the source.

three levels of excitation is equal to $z_n=76.7$, 24, and 10.7 cm. In the first case of the small initial amplitude, the propagation is weakly nonlinear and the shock will form in the smooth far field of the beam, beyond the distances considered here. In the latest two cases the propagation is highly nonlinear and the shock front will form within the oscillating part of the beam. Strong nonlinear interactions in a regular but oscillating acoustic field are of particular interest as they might be realized quite differently compared with those in plane waves or far field of weakly nonlinear beams.

Axial distributions of the second and third harmonics, Figs. 1(b) and (c), also have oscillating structure. For each harmonic, the structure of the oscillations is qualitatively the same for all the source amplitudes. The positions of maximums and minimums, especially for the very last maximum and minimum, move closer to the source for higher nonlinearities. The amplitudes of oscillations are much higher in the case of stronger nonlinear effects. The numerical solutions correlate very well with each other, although some slight discrepancy is observed between the measured and calculated amplitudes of the harmonics. This discrepancy between experiment and numerical simulations has been observed and explained in a previous paper.[30] Figure 2 shows the beam patterns (lateral distributions), measured and calculated numerically at the longest distance $z=400$ mm from the source. Again, the numerical calculations correlate well with the experimental data.

A very good agreement between experiment and numerical solutions, obtained both by the time-domain and the frequency-domain codes, demonstrates the validity and accuracy of the two algorithms to model the first harmonics in the near field of the piston source for both quasilinear and strongly nonlinear wave propagation. However, the spectrum of the nonlinear wave, especially with shocks, is wide and the errors in numerical simulations are more pronounced for higher harmonics. Although the results of the numerical simulations for the first harmonics are close to each other, the high-frequency part of the spectrum might be quite different. Temporal waveforms simulated directly in the time domain or reconstructed from the results of the modeling in the frequency domain consist of all the harmonics and thus are more informative in terms of comparison of the codes. Unfortunately, there are no experimental waveforms available, therefore only the results of simulations are presented here. Shown in Fig. 3 are pressure-time waveforms calculated by the two codes at various distances z from the source and for various values of nonlinear parameter N, which is proportional to initial pressure amplitude p_0. The results are in a very good agreement. The axial waveforms calculated by the time-domain code are noisy for higher nonlinearities; however, the waveforms in average are very close to each other. Two interesting features can be observed for the waveform calculated for $N=10.8$ at the distance $z=255$ mm. It is clearly seen that the waveform calculated by the frequency domain code contains two thin shocks in a cycle and the positive peak pressure is about two times higher than that

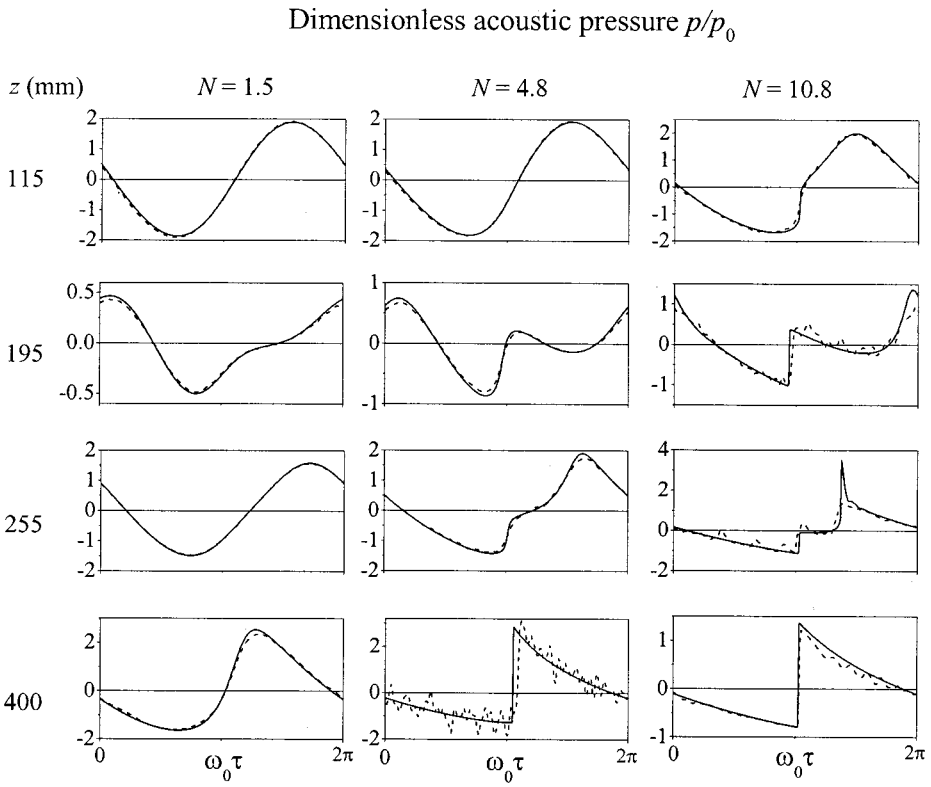


FIG. 3. Comparison of numerical results. Axial waveforms calculated using frequency domain (———) and time-domain (---) codes at various distances z from the source and for various initial pressure amplitudes p_0.

observed for the quasilinear propagation, $N=1.5$. To our knowledge, formation of two shocks within one period of an initially harmonic wave and strong excess of the peak positive pressure has never been observed earlier in nonlinear acoustic beams and will be addressed in more detail in the next section of this article.

In conclusion to this section it should be noted that the comparison of the experimental data available for harmonics in the near field of the piston circular source and the results of numerical simulations obtained by the two different numerical codes shows a good agreement between them. The calculated waveforms also correlate very well with each other. The time domain code previously developed for the pulsed regime of nonlinear beams is thus validated here to operate well for continuous waves. However, the calculations of the entire acoustic field in the time domain are rather time consuming because of a long time window, fine time grid, and mostly because of the necessity to calculate the Rayleigh integral for each spatial point. Calculation time was respectively 18.7 h, 14.6 days and about 45 days for $N=1.5$, $N=4.8$, and $N=10.8$ on a Compaq XP-1000 workstation (Digital Alpha 21264 533 MHz chip), corresponding to approximately 6 days, 4 months, and 13 months on a Pentium III 350 MHz. This is for the full z axis (up to 400 mm) with 5-mm diffraction step.

The newly developed and optimized for the specific problem frequency-domain code presented here is less time consuming and more accurate for the description of a continuous wave. The computational time varied from 7.5 min for the case of $N=1.5$ to 12.5 h for $N=4.8$ and 32.5 h for $N=10.8$ on a Compaq XP-1000. This time corresponds to approximately 1 h for $N=1.5$, 4 days for $N=4.8$, and 10 days for $N=10.8$ on a Pentium III 350 MHz. The frequency-domain algorithm will be used here for further detailed study of nonlinear interactions in the near field of the piston source operating in a highly nonlinear regime.

B. Shock wave formation in the near field of the beam

It is seen from Fig. 3 that nonlinear waveform distortion and formation of shocks in the oscillating near field of the beam happen in a quite different way compared to that known for the plane wave propagation. Shown in Fig. 4 are the pressure-time waveforms calculated at various dimensionless distances $Z=z/z_d$ from the source for initial pressure amplitude $p_0=14.3$ bars ($N=10.8$) with and without accounting for diffraction effects. Axial waveforms are presented in the first column. The second and the third columns correspond to the waveforms at $R=0.5$ and 1 source radius off-axis positions. Consequent development of two shocks on the axis of the beam is seen from the first column. The first shock forms at nearly the same distance $Z\sim 0.1$ and has nearly the same phase $T_s=\omega_0\tau=\pi$ in time as the shock in the case of the plane wave propagation. Variations in asymmetry of the shock position regarding the zero level $p/p_0=0$, which is due to diffraction, lead to relative movement of the shock forward and backwards compared with the shock position for the plane wave. The distance $Z=0.15$ corresponds to the last minimum for the first harmonic and to the maximum of the second harmonic. The waveform therefore consists of the two oscillations of different shape. One oscillation is smooth and one is already shocked. The smooth maximum then grows and moves faster, appearing in the right part of the waveform ($Z=0.18$). Further propagation of the wave leads to development of the second shock with a

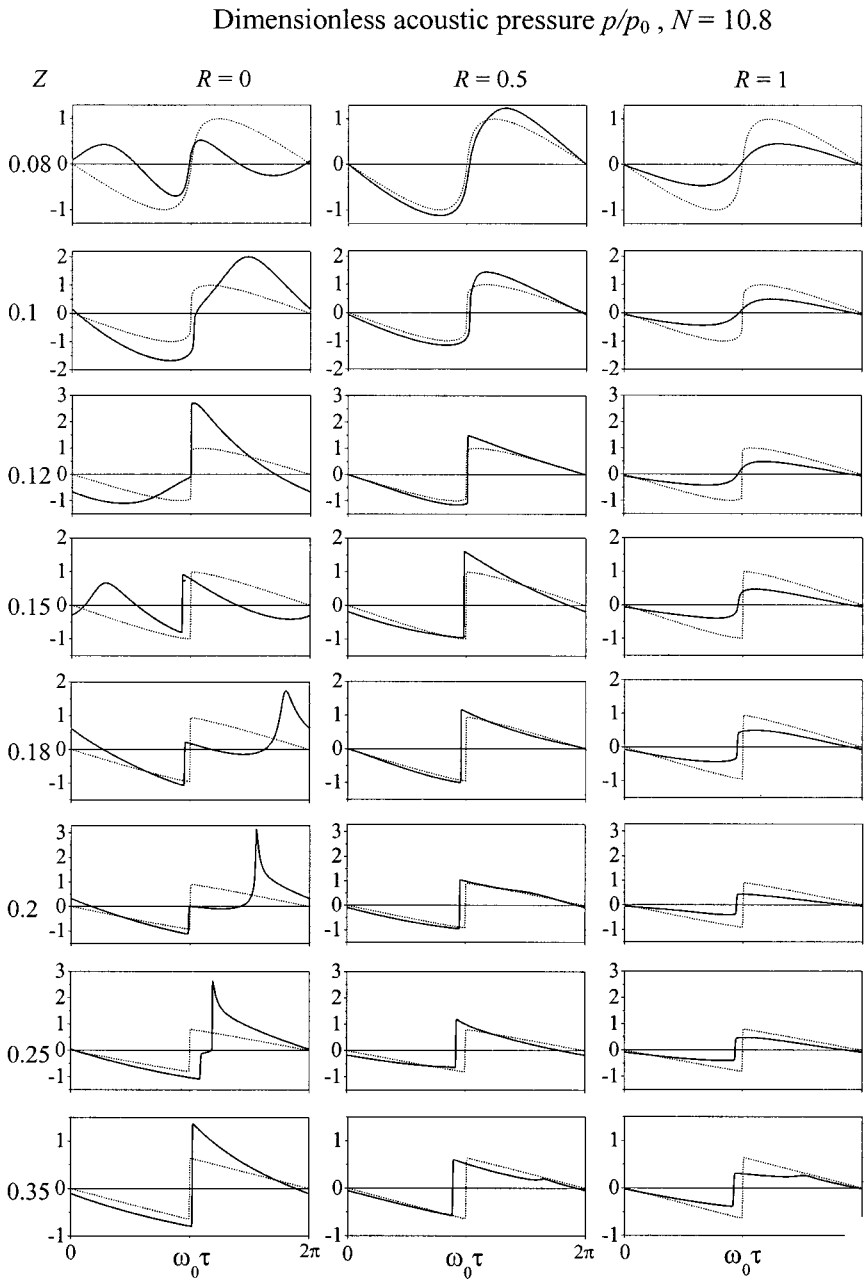


FIG. 4. Waveforms at various distances $Z=z/z_d$ from the source on axis, $R=0$, and radially at $R=0.5$, $R=1$ source radius off axis (solid lines) for $N=10.8$. Dotted lines are waveforms that correspond to the plane wave propagation without diffraction.

narrow sharp peak at the front of the second maximum ($Z=0.20$), relative movement of these two shocks towards each other ($Z=0.25$), collision of shocks, and formation of sawtooth wave ($Z=0.35$). The appearance and growth of the second maximum can be explained in terms of the edge wave coming to the axis in a different phase. It is seen that the amplitude of the resulting shock at $Z=0.35$ is even higher than that for the plane wave at the same distance.

Contrary to the axial waveform transformation, the waveforms off-axis contain only one shock and look qualitatively similar to those computed for the plane propagation. Very small difference is observed at the radial distance $R=0.5$ off the axis. Strong diffraction smoothening of the beam close to the edge, $R=1$, results in less amplitude of the wave compared to the plane wave. Nonlinear effects are therefore less pronounced, and the shock forms at longer distances from the source and has smaller amplitude.

The relative movement and interaction of two shocks in one cycle are shown in more detail in Fig. 5 where the axial waveforms at various distances are presented for $N=10.8$. The waveforms are compared with those computed for the case of linear diffraction ($N=0$, dotted lines). It is seen that at the distances when two shocks are developed, the maximum asymmetry of the waveform is observed. The peak positive pressure, despite the nonlinear attenuation at the shocks, exceeds the initial pressure p_0 more than 3.5 times ($Z=0.22$) compared to 2 times excess which can be observed in the near field of the linear beam, $N=0$. On the contrary, the peak negative pressure is less than that for the linear case. Close-up view of the waveforms close to the second positive peak (within the area limited by dashed lines in Fig. 5) is shown in Fig. 6 for various distances Z. The waveforms are presented both on the axis, $r=0$, and at small distances off-axis $r=0.15$, 0.3, and 0.45 mm. The structure of this part of the waveforms on- and off-axis is very different at the distances $Z=0.2$ and 0.22. Very strong excess of

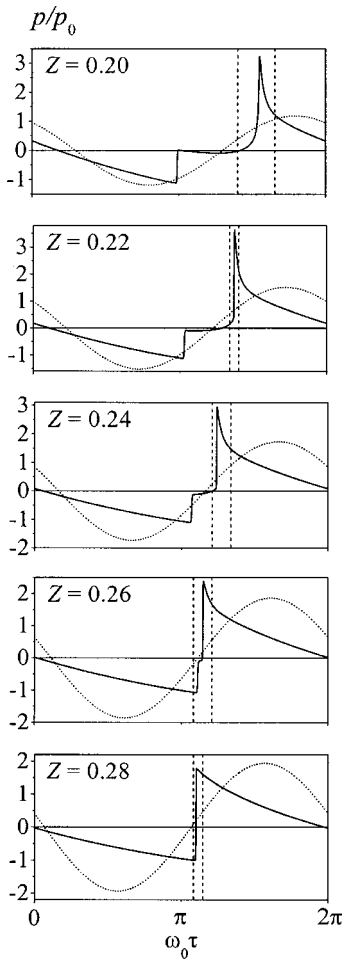


FIG. 5. Axial waveforms at various distances $Z=z/z_d$ from the source for $N=10.8$ (solid lines). Dotted lines are axial waveforms corresponding to the diffraction of the linear beam for $N=0$.

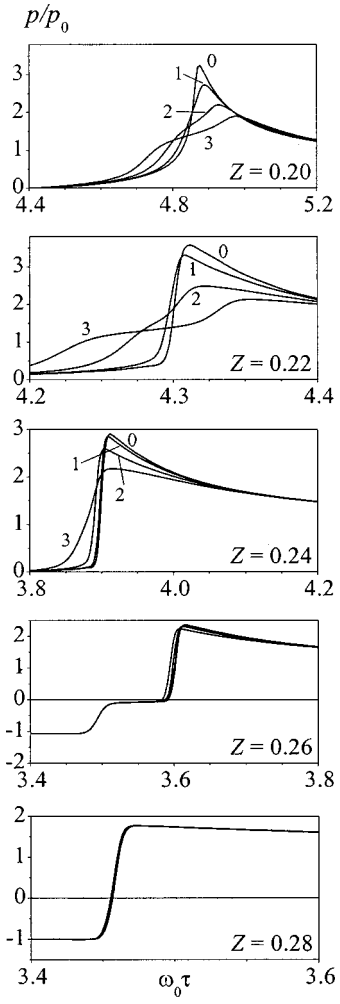


FIG. 6. Close-up view of the waveforms within the area around the shock shown by dashed lines in Fig. 5 at various distances $Z=z/z_d$ from the source for $N=10.8$. Numbers by the curves correspond to various lateral coordinates: $r=0$ (0), $r=0.15$ mm (1), $r=0.3$ mm (2), and $r=0.45$ mm (3).

the sharp positive peak pressure is observed very close to the axis only, then the peak value goes down rapidly off the axis. When the peak is smoothed due to strong absorption ($Z=0.24$ and 0.26) the axial and off-axis waveforms become closer, and almost coincide after shock collision, $Z=0.28$. Strong excess of peak positive pressure value thus can be observed at the very small paraxial distances only.

The effect of consequent formation of two shocks on the beam axis corresponds to very strong nonlinearity, $N=10.8$ or shock formation distance $z_n=10.7$ cm. In this case the first shock forms within the next to the last axial lobe of the fundamental frequency [Fig. 1(a)]. Then, after passing through the minimum, the first harmonic grows again with the different phase and the second shock forms. Evolution of the axial waveform for weaker nonlinearity, $N=4.8$, and shock formation distance $z_n=24$ cm, is shown in Fig. 7. It is seen that at the distance beyond the very last minimum, $Z=0.24$ and 0.26, the waveform also consists of two distorted fronts, but, opposite to the case $N=10.8$, both the fronts are not yet shocked. Then the fronts steepen and two shocks develop simultaneously, $Z=0.28$ and 0.30. The phase delay between shocks is rather small and they collide very soon at $Z=0.32$.

Axial distributions of the normalized peak positive P_+, peak negative P_-, and peak-to-peak P_{pp} pressure amplitudes for various nonlinearities, N, are shown in Fig. 8. In the case of linear propagation, $N=0$, the waveform is symmetrical and behavior of P_+, P_-, and P_{pp} is identical (dashed curves). The main difference between linear and nonlinear propagation is mostly pronounced in the last two diffraction lobes. The very last minimum of P_+, P_-, and P_{pp} does not reach zero and occurs at different distances closer to transducer for higher nonlinearities. Nonlinear propagation results in asymmetry of the waveform, so that the peak positive pressure increases and peak negative pressure decreases. Weak nonlinearity, $N=1.5$, results in weak enhancement of the peak positive and weak reduction of the peak negative pressure, while the peak-to-peak pressure distribution is almost not affected. Strong enhancement of the peak positive pressure is observed for strongly nonlinear propagation, $N=4.8$ and 10.8. About four times excess over initial pressure amplitude is observed compared to only two times excess for the linear case $N=0$. It is also seen from Fig. 8 that these high maximums of P_+ are reached at different distances, which correspond to formation of the sharpest second shock (see Fig. 7 for $N=4.8$ and Fig. 5 for $N=10.8$). The

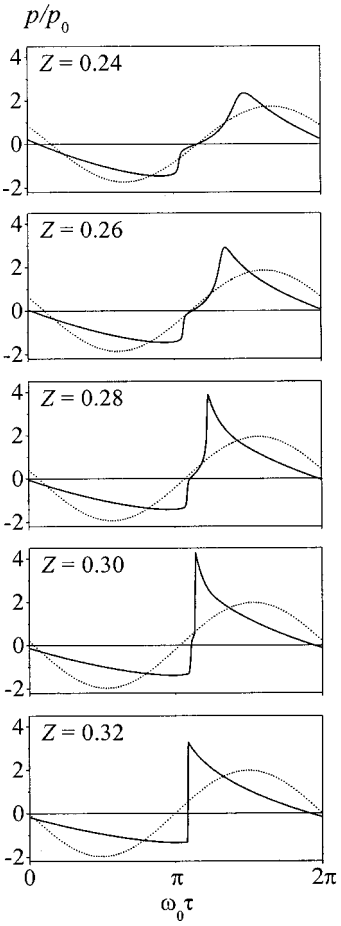


FIG. 7. Axial waveforms at various distances $Z=z/z_d$ from the source for $N=4.8$ (solid lines). Dotted lines are axial waveforms corresponding to the diffraction of the linear beam, $N=0$.

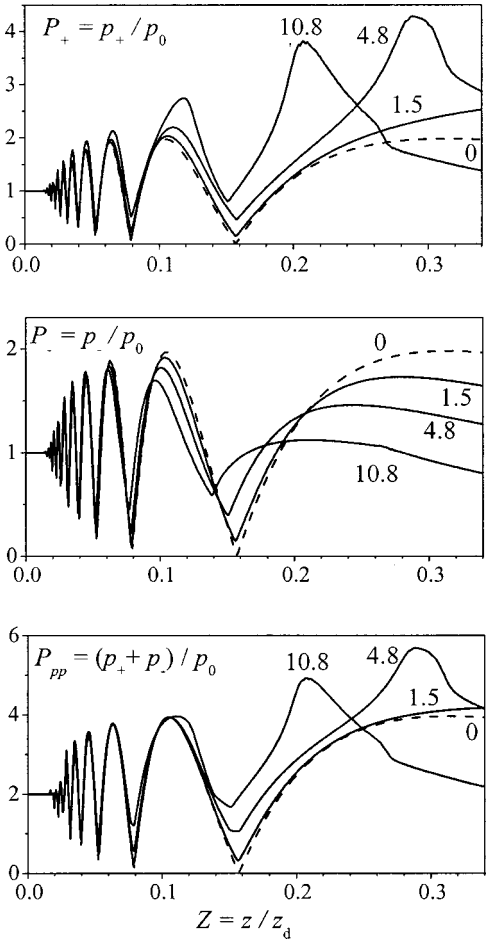


FIG. 8. Axial distributions of the normalized peak positive P_+, peak negative P_-, and peak-to-peak P_{pp} pressure for various values of nonlinear parameter N (numbers by the curves).

behavior of the peak negative pressure is more monotonic. Positions of the very last maximum and minimum shift towards the transducer and their values decrease for higher values of the nonlinearity N. The structure of the peak-to-peak pressure P_{pp} distribution is similar to that for P_+ but smoother, so that the excess over its linear distribution is less pronounced.

Distribution of the normalized mean intensity, Eq. (13), along the axis of the beam is shown in Fig. 9. Linear, $N=0$, and weakly nonlinear, $N=1.5$, distributions of intensity are almost the same, although the waveforms are already different. Formation of shocks and corresponding effective energy absorption in the case of strong nonlinearity leads to decrease of intensity value. The effect of collision of two shocks ($N=4.8$,10.8) manifests itself in distinctive phenomenon as a fracture or a sharp turn in intensity propagation curves. The positions of the fracture corresponds to the distances when collision of two shocks happens (see Figs. 5 and 7). As it is known,[14] the energy absorption at the shock is proportional to the third power of the shock amplitude A_s. When two steep fronts are presented in the waveform, energy absorption is thus proportional to the sum of the cubes of the two shock amplitudes, $A_{S,1}^3+A_{S,2}^3$. After shock collision, absorption is proportional to the cube of the sum of the shock amplitudes $(A_{S,1}+A_{S,2})^3$, which evidently exceeds the sum of cubes $A_{S,1}^3+A_{S,2}^3$. The wave intensity therefore turns to decrease much faster.

Two-dimensional spatial distributions of the normalized peak positive pressure are demonstrated in Fig. 10 for various values of nonlinearity N. It is seen that the structure of the field is different for different N. The combined effect of diffraction and nonlinearity results in excess of the peak positive pressure over the maximum $p/p_0=2$ reached in a linear beam, when $N=0$. This effect is pronounced even for the weak nonlinearity, $N=1.5$, close to the axis in a last diffraction lobe of the beam. The shape of the spots of 1.5 times excess, however, is still very similar as in the linear

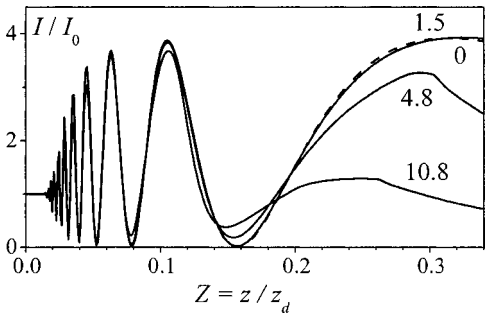


FIG. 9. Axial distributions of the normalized intensity I/I_0 for various values of nonlinear parameter N (numbers by the curves).

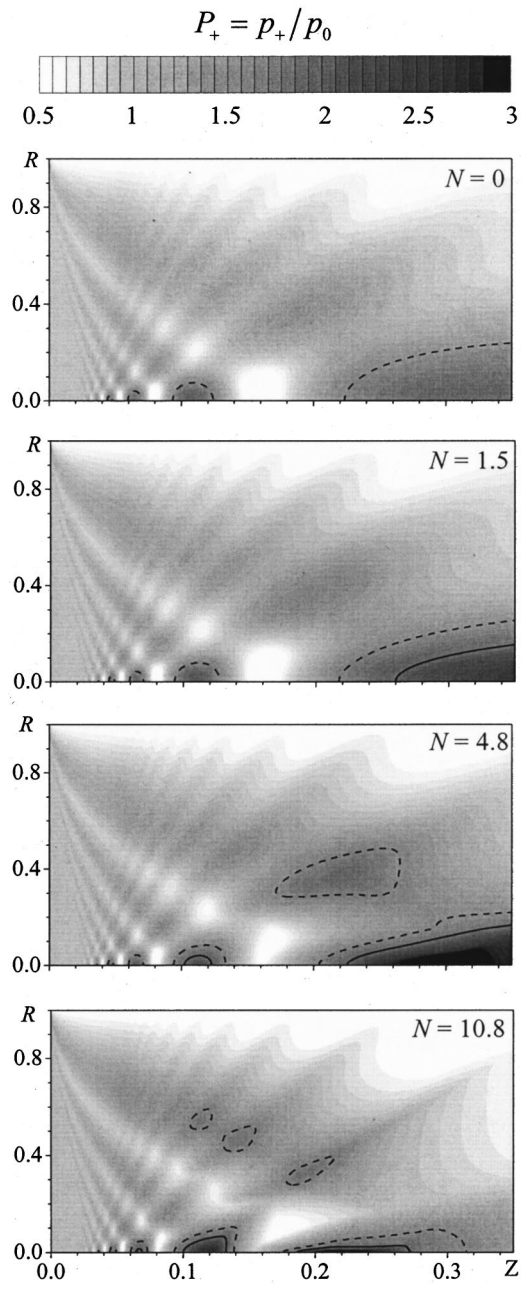


FIG. 10. Two-dimensional spatial distributions of the normalized peak positive pressure $P_+ = p_+/p_0$ for various values of nonlinear parameter $N = 0$, 1.5, 4.8, 10.8. Dashed and solid lines surround the areas of 1.5 and 2 times excess over p_0, respectively.

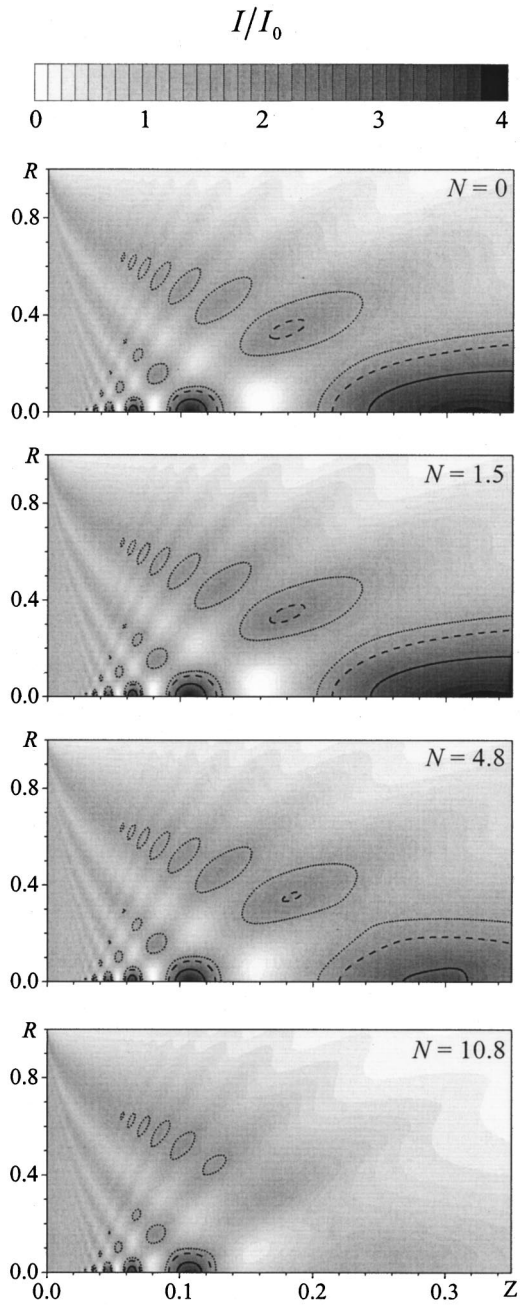


FIG. 11. Two-dimensional spatial distributions of the normalized intensity I/I_0 for various values of nonlinear parameter $N = 0$, 1.5, 4.8, 10.8. Dotted, dashed, and solid lines surround the areas of 1.5, 2, and 3 times excess over I_0, respectively.

beam. For higher nonlinearity, $N = 4.8$, the areas of 1.5 and 2 times excess in the last diffraction lobe are narrower off the axis and move towards the transducer. A small spot of double excess appears in the next to the last diffraction lobe close to the axis. An additional spot of 1.5 times excess forms off the axis, which is not observed in the linear and weakly nonlinear beam. Finally, for the strongest nonlinearity, $N = 10.8$, the excess axial spots are reduced in space, the formation of additional spots of 1.5 excess is observed off the axis. Strong nonlinear absorption results in decrease of the normalized peak pressure at longer distances $Z > 0.25$ compared to the cases of weaker nonlinearity.

Two-dimensional spatial distributions of the normalized intensity are shown in Fig. 11 for various values of nonlinearity N. It is seen that the near field of the linear beam, $N = 0$, has spotted spatial structure of different intensity less (up to zero value) and more (up to 4 times higher) than the constant value $I/I_0 = 1$ correspondent to the linear propagation of the plane wave. The spots of 1.5, 2, and 3 times excess are shown in the figure by the corresponding lines. It is seen that up to the distances before the shocks are developed, i.e., $Z < 0.28$ for $N = 4.8$ and $Z < 0.11$ for $N = 10.8$, the intensity distributions are quite similar to each other. Effective absorption at the shocks results in smoothening of the distributions for $N = 4.8$. For the highest nonlinearity, $N = 10.8$, the intensity does not reach even 1.5 excess level

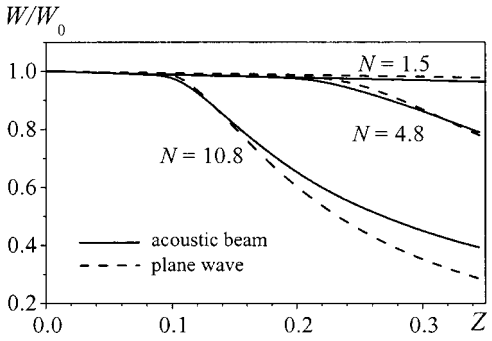


FIG. 12. Axial distributions of the normalized total power of the beam W/W_0 (solid lines) and the normalized intensity of the plane wave without diffraction I/I_0 (dashed lines). Various values of nonlinear parameter $N=1.5$, 4.8, 10.8 correspond to initial pressure amplitude $p_0=2$, 6.4, 14.3 bars.

within the last diffraction maximum as well as in the very last excess spot off the axis.

The nonlinear effect of the energy transfer to higher harmonics and formation of shocks are more pronounced within the areas of higher intensity which leads to more effective absorption. On the contrary, nonlinearity is less pronounced within the spots of intensity minimums. The average effect of nonlinearity on acoustic energy decay in a beam and in a plane wave with the same initial intensity is illustrated in Fig. 12. In this figure the axial distributions for the initial pressure amplitudes $p_0=2$, 6.4, and 14.3 bars are shown. Solid curves correspond to the normalized total power of the beam, Eq. (14), and dashed curves to the normalized intensity in the plane wave calculated along the propagation distance Z. It is seen that absorption rate is a little bit higher for the beam even at the distances before shock formation, which is all the propagation range for the weak nonlinearity, $N=1.5$, and smaller distances for higher nonlinearities, $N=4.8$ and 10.8. This small discrepancy is caused by the artificial viscosity introduced in the numerical algorithm in the layer close to the edge of the spatial window to avoid reflections from the edge. Formation of shocks, $N=4.8$ and 10.8, and corresponding rapid increase of absorption start earlier in a beam because of the amplitude fluctuation close to the axis. The propagation curve for the total power of the beam thus turns down earlier compared to the intensity of the plane wave. At longer distances, however, nonlinear absorption is more pronounced for the plane wave due to spatial inhomogeneities, diffraction smoothening in the beam, and delay of shock formation close to the edge of the beam. As it is seen from Fig. 4, even in the case of the strongest nonlinearity, $N=10.8$, the shock develops later and its amplitude is smaller at the radial distance $R=1$, compared to the plane wave propagation.

IV. CONCLUSION

Two theoretical models and the corresponding numerical codes for the description of nonlinear acoustic beams radiated from intense cw sources are presented. The main difference of the theoretical models contains diffraction effects, which are included using the Rayleigh integral in the first model, whereas the parabolic approximation is applied in the second model. The main differences of the numerical implementation are the following. The first code, originally developed for the pulsed regime, operates in the time domain and the exact analytic results are applied at each fractional step using the operator splitting procedure for all the physical effects included: absorption, nonlinearity, and diffraction. The second newly developed code operates in the frequency domain and the exact analytic solution is used for the absorption, while nonlinearity and diffraction are calculated using the finite difference method. To validate the codes, the simulations of the near field of an intense circular piston ultrasound source in water are performed and the results are compared with each other and with the available experimental data. The results of modeling are found to be in a good agreement both for the axial and lateral distributions of the harmonics amplitude (measured and simulated), as well as for the temporal acoustic waveforms (simulated). A strongly nonlinear regime of beam propagation, when shocks are developed in the waveform close to the source, is studied in details using the frequency-domain code. It is demonstrated that diffraction plays a major role in shock formation. A new nonlinear wave phenomenon of the development of two shocks in each cycle of an initially harmonic wave, followed by their motion towards each other and further collision, is predicted. The distance of shock collision depends on the initial pressure amplitude, which can be used for the calibration process. It is shown that in the case of nonlinear propagation, the maximum positive peak pressure in the near field strongly exceeds the linear predictions, which may have significant implications in medical applications of the intense ultrasound such as physiotherapy, HIFU surgery, and lithotripsy. The beam total power decay as a function of the propagation distance is compared with the intensity for the plane wave propagation without diffraction. It is found that the energy decay starts earlier for the beam, but decreases slower over longer distances. Formation of two shocks in a cycle does not allow us to employ a previously developed asymptotic spectral approach,[21] which enables us to effectively simulate nonlinear waves with shocks using a small number of harmonics.[17,18,22,23] The asymptotic method is based on the assumption that only one shock might be presented in one cycle of the periodic wave, which cannot be applied to the problem considered here. Direct calculations for a large number of harmonics are therefore performed, which makes the simulations rather time consuming.

Although the results of the simulations obtained by the two codes were presented here for the near field of a plane circular source in water, both the models developed can be applied also for the focused sources, and for propagation in tissue or layered media with arbitrary absorption and dispersion laws.

ACKNOWLEDGMENTS

This work was partly supported by PECO and RFBR grants.

[1] S. I. Aanonsen, ''Numerical computation of the nearfield of a finite amplitude sound beam,'' Tech. Rep. No. 73, Department of Mathematics, University of Bergen, Bergen, Norway (1983).

[2] S. I. Aanonsen, T. Barkve, J. Naze Tjotta, and S. Tjotta, ''Distortion and harmonic generation in the nearfield of a finite amplitude sound beam,'' J. Acoust. Soc. Am. **75**, 749–768 (1984).
[3] N. S. Bakhvalov, Y. M. Zhileikin, and E. A. Zabolotskaya, *Nonlinear Theory of Sound Beams* (American Institute of Physics, New York, 1987).
[4] T. Kamakura, N. Hamada, K. Aoki, and Yu. Kumamoto, ''Nonlinearly generated spectral components in the nearfield of a directive sound source,'' J. Acoust. Soc. Am. **85**, 2331–2337 (1989).
[5] F. M. Pestorius, ''Propagation of plane acoustic noise of finite amplitude,'' Technical report ARL-TR-7323 (Applied Research Laboratory, The University of Texas at Austin, 1973), AD 778 868.
[6] R. O. Cleveland, J. P. Chambers, H. E. Bass, R. Raspet, D. T. Blackstock, and M. F. Hamilton, ''Comparison of computer codes for the propagation of sonic boom waveforms through isothermal atmospheres,'' J. Acoust. Soc. Am. **100**, 3017–3027 (1996).
[7] Y. S. Lee and M. F. Hamilton, ''Time-domain modeling of pulsed finite amplitude sound beams,'' J. Acoust. Soc. Am. **97**, 906–917 (1995).
[8] M. A. Averkiou and M. F. Hamilton, ''Nonlinear distortion of short pulses radiated by plane and focused circular pistons,'' J. Acoust. Soc. Am. **102**, 2539–2548 (1997).
[9] J. Tavakkoli, D. Cathignol, R. Souchon, and O. A. Sapozhnikov, ''Modeling of pulsed finite-amplitude focused sound beams in time domain,'' J. Acoust. Soc. Am. **104**, 2061–2072 (1998).
[10] P. T. Christopher and K. J. Parker, ''New approaches to nonlinear diffractive field propagation,'' J. Acoust. Soc. Am. **90**, 488–499 (1991).
[11] F. P. Curra, P. D. Mourad, V. A. Khokhlova, and L. A. Crum, ''High intensity focused ultrasound and tissue heating: the effect of nonlinear sound propagation and vessel presence,'' in *Proc. of IEEE International Ultrasonics Symposium*, Sendai, Japan (1998), pp. 1419–1422.
[12] J. Naze Tjotta, S. Tjotta, and E. H. Vefring, ''Propagation and interaction of two collimated finite amplitude sound beams,'' J. Acoust. Soc. Am. **88**, 2859–2870 (1990).
[13] T. Christopher, ''A nonlinear plane-wave algorithm for diffractive propagation involving shock waves,'' J. Comput. Acoust. **1**, 371–393 (1993).
[14] M. F. Hamilton and D. T. Blackstock, *Nonlinear Acoustics* (Academic, San Diego, 1998).
[15] T. S. Hart and M. F. Hamilton, ''Nonlinear effects in focused sound beams,'' J. Acoust. Soc. Am. **84**, 1488–1496 (1988).
[16] B. Ystad and J. Bernsten, ''Numerical solution of the KZK equation for focusing sources,'' Acta Acust. **3**, 323–330 (1995).
[17] V. A. Khokhlova, M. A. Averkiou, S. J. Younghouse, M. F. Hamilton, and L. A. Crum, ''Fast Spectral Algorithm for Modeling of Focused Sound Beams in a Highly Nonlinear Regime,'' in *Proceedings of 16th International Congress on Acoustics and 135th Meeting of Acoustical Society of America*, Seattle, WA, 20–26 June 1998, Vol. 4, pp. 2875–2876.
[18] V. A. Khokhlova, S. S. Kashcheeva, M. A. Averkiou, and L. A. Crum, ''Effect of selective absorption on nonlinear interactions in high intensity acoustic beams,'' in Proceedings of 15th ISNA, Goettingen, Germany (1999).
[19] M. F. Hamilton, V. A. Khokhlova, and O. V. Rudenko, ''Analytical method for describing the paraxial region of finite amplitude sound beams,'' J. Acoust. Soc. Am. **101**, 1298–1308 (1997).
[20] A. N. Dubrovsky, V. A. Khokhlova, and O. A. Sapozhnikov, ''Nonlinear and diffraction effects in a beam of weak shocks,'' in *Advances in Nonlinear Acoustics*, Proceedings of 13th ISNA, Bergen, Norway (World Scientific, Singapore, 1993), pp. 227–232.
[21] Y. A. Pishchal'nikov, O. A. Sapozhnikov, and V. A. Khokhlova, ''A modification of the spectral description of nonlinear acoustic waves with discontinuities,'' Acoust. Phys. **42**, 362–367 (1996).
[22] V. G. Andreev, Y. A. Pischal'nikov, O. A. Sapozhnikov, V. A. Khokhlova, and R. O. Cleveland, ''Diagnostics of a relaxing medium by an acoustic pulse with shock front,'' Acoust. Phys. **45**(1), 8–13 (1999).
[23] S. S. Kasheeva, V. A. Khokhlova, O. A. Sapozhnikov, M. A. Averkiou, and L. A. Crum, ''Nonlinear Distortion and Absorption of Intense Acoustic Waves in Media with Power Frequency Law of Absorption,'' Acoust. Phys. **46**(2), 170–177 (2000).
[24] A. C. Baker, A. M. Berg, A. Sahin, and J. Naze Tjotta, ''The nonlinear pressure field of plane rectangular apertures: Experimental and theoretical results,'' J. Acoust. Soc. Am. **97**, 3510–3517 (1995).
[25] T. Kamakura, M. Tani, Y. Kumamoto, and K. Ueda, ''Harmonic generation in finite amplitude sound beams from a rectangular aperture source,'' J. Acoust. Soc. Am. **91**, 3144–3151 (1992).
[26] E. A. Zabolotskaya and R. V. Khokhlov, ''Quasi-plane waves in the nonlinear acoustics of confined beams,'' Sov. Phys. Acoust. **15**, 35–40 (1969).
[27] V. P. Kuznetsov, ''Equations of nonlinear acoustics,'' Sov. Phys. Acoust. **16**, 467–470 (1971).
[28] J. Bernsten, J. Naze Tjotta, and S. Tjotta, ''Nearfield of a large acoustic transducer. Part 4: Second harmonic and sum frequency radiation,'' J. Acoust. Soc. Am. **75**, 1383–1391 (1984).
[29] A. C. Baker, K. Anastasiadis, and V. F. Humphrey, ''The nonlinear pressure field of a plane circular piston: Theory and experiment,'' J. Acoust. Soc. Am. **84**, 1483–1487 (1988).
[30] S. Nachef, D. Cathignol, J. N. Tjotta, A. M. Berg, and S. Tjotta, ''Investigation of a high intensity sound beam from a plane transducer. Experimental and theoretical results,'' J. Acoust. Soc. Am. **98**, 2303–2323 (1995).
[31] O. V. Rudenko and S. I. Soluyan, *Theoretical Foundations of Nonlinear Acoustics* (Plenum, New York, 1977).
[32] J. Naze Tjotta, S. Tjotta, and E. H. Vefring, ''Propagation and interaction of two collimated finite amplitude sound beams,'' J. Acoust. Soc. Am. **88**, 2859–2870 (1990).
[33] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN*, 2nd ed. (Cambridge U. P., New York, 1992).

Linear and nonlinear frequency shifts in acoustical resonators with varying cross sections

Mark F. Hamilton, Yurii A. Ilinskii, and Evgenia A. Zabolotskaya
Department of Mechanical Engineering, The University of Texas at Austin, Austin, Texas 78712-1063

(Received 31 January 2000; accepted for publication 16 April 2001)

The frequency response of a nonlinear acoustical resonator is investigated analytically and numerically. The cross-sectional area is assumed to vary slowly but otherwise arbitrarily along the axis of the resonator, such that the Webster horn equation provides a reasonable one-dimensional model in the linear approximation. First, perturbation theory is used to derive an asymptotic formula for the natural frequencies as a function of resonator shape. The solution shows that each natural frequency can be shifted independently via appropriate spatial modulation of the resonator wall. Numerical calculations for resonators of different shapes establish the limits of the asymptotic formula. Second, the nonlinear interactions of modes in the resonator are investigated with Lagrangian mechanics. An analytical result is obtained for the amplitude-frequency response curve and nonlinear resonance frequency shift for the fundamental mode. For a resonator driven at its lowest natural frequency, it is found that whether hardening or softening behavior occurs depends primarily on whether the nonlinearly generated second-harmonic frequency is greater or less than the second natural frequency of the resonator. A fully nonlinear one-dimensional numerical code is used to verify the analytical result.
[DOI: 10.1121/1.1379080]

PACS numbers: 43.25.Gf, 43.20.Ks [ANN]

I. INTRODUCTION

Nonlinear resonance frequency shifts are well-known phenomena in the theory of nonlinear oscillations, but they have received little attention in acoustics.[1] For such frequency shifts to be significant in an acoustical resonator the sound level must be quite high, but in addition the natural frequencies of the resonator must not be integer multiples of the drive frequency. Only recently have such conditions been met in industrial applications of acoustics. Nonlinear amplitude-frequency response curves have been measured in acoustical resonators used as compressors.[2] The resonators are designed with varying cross sections to shift the natural frequencies and thus suppress harmonic generation. The response curves lean toward higher frequencies (hardening behavior) for some resonator shapes and toward lower frequencies (softening behavior) for others. Although these curves have been modeled numerically,[3,4] no explanation for the behavior has been proposed. Thermoacoustic engines and refrigerators represent another class of industrial applications that involves high sound levels in resonators having varying cross sections.[5] Here again, the shape of the resonator is recognized to influence significantly the nonlinear behavior of the devices.[6]

In general, it is clear that the nonlinear frequency response is determined by how the harmonics of the drive frequency, particularly the second harmonic, are related to the natural frequencies of the resonator. This relationship is determined in turn by the shape of the resonator. The purpose of the present article is to predict and explain the dependence of the frequency response on resonator shape. It is assumed that the cross-sectional area of the resonator varies slowly as a function of position along the axis, which permits the motion of the fluid inside the resonator to be treated as one-dimensional.

The article is divided into two main parts. In Sec. II, the Webster horn equation is solved by perturbation to obtain an asymptotic expression for the natural frequencies of a resonator with varying cross section. The solution is a refinement of a result obtained by Rayleigh using an energy approach. The asymptotic expression is compared with direct numerical solution of the Webster horn equation to ascertain the domain of validity. In Sec. III, Lagrangian mechanics is used to investigate nonlinear wave interactions in a resonator with varying cross section. Coupled equations are obtained for the drive frequency component, which include the dominant effects of second-harmonic generation, static deformation, and cubic nonlinearity. These equations are combined to obtain analytical expressions for the wave amplitude in the resonator due to an external drive, and for the nonlinear resonance frequency as a function of the wave amplitude. The derived expressions are shown to be in good agreement with direct numerical solution of the fully nonlinear equations for one-dimensional motion in a resonator with varying cross section.[3]

II. LINEAR THEORY

The Webster horn equation and perturbation theory are used in this section to derive an asymptotic result for the natural frequencies of resonators with slowly varying cross sections. The asymptotic result reveals how the natural frequency spectrum depends on the shape of the resonator. Comparisons with numerical solutions establish the validity of the analysis.

A. Natural frequency shifts

We consider a resonator with a cross-sectional area S that may be expressed in the form

$$S(x)=S_0+\delta S(x), \tag{1}$$

where x is the axial coordinate, S_0 is a reference area, and δS is a small perturbation that varies slowly with x. For simplicity we shall employ the terminology ''cylindrical resonator'' to describe the case $\delta S=0$, even though S_0 may be the area of a constant cross section with arbitrary, not necessarily circular, shape. An appropriate one-dimensional model for the sound pressure p inside the resonator is the Webster horn equation,[7]

$$\frac{1}{S}\frac{\partial}{\partial x}\left(S\frac{\partial p}{\partial x}\right)=\frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2}, \tag{2}$$

where c_0 is the small-signal sound speed. We consider the nth resonator mode with angular frequency ω_n and write

$$p(x,t)=A_n p_n(x)e^{i\omega_n t}, \tag{3}$$

where the coefficient A_n has dimensions of pressure, and therefore p_n is dimensionless. Substitution in Eq. (2) yields

$$\frac{1}{S}\frac{d}{dx}\left(S\frac{dp_n}{dx}\right)=-(\omega_n^2/c_0^2)p_n. \tag{4}$$

It is convenient to introduce the notation

$$\lambda_n=-\omega_n^2/c_0^2 \tag{5}$$

for the eigenvalues corresponding to the eigenfunctions (mode shapes) p_n of the differential operator

$$L=\frac{1}{S}\frac{d}{dx}\left(S\frac{d(\cdot)}{dx}\right)=\frac{d^2}{dx^2}+\frac{S'}{S}\frac{d}{dx}. \tag{6}$$

Primes indicate derivatives with respect to x. Equation (4) may now be rewritten as

$$Lp_n=\lambda_n p_n. \tag{7}$$

It is assumed that all walls of the resonator are rigid, and that the boundary conditions at the ends are

$$p_n'(0)=p_n'(l)=0, \tag{8}$$

where l is the length of the resonator.

As with the area perturbation in Eq. (1), we express the differential operator, its eigenfunctions, eigenvalues, and the natural frequencies as

$$L=L_0+\delta L, \tag{9}$$

$$p_n=p_{0n}+\delta p_n, \tag{10}$$

$$\lambda_n=\lambda_{0n}+\delta\lambda_n, \tag{11}$$

$$\omega_n=\omega_{0n}+\delta\omega_n, \tag{12}$$

where the leading terms correspond to $\delta S=0$ and therefore

$$L_0=d^2/dx^2, \tag{13}$$

$$p_{0n}=\cos(n\pi x/l), \tag{14}$$

$$\lambda_{0n}=-(n\pi/l)^2, \tag{15}$$

$$\omega_{0n}=n\pi c_0/l. \tag{16}$$

From Eq. (6) we have

$$\delta L=\frac{S'}{S}\frac{d}{dx}. \tag{17}$$

The objective is to determine the corresponding values of the natural frequency shifts $\delta\omega_n$. We do so by first determining the values of $\delta\lambda_n$.

To begin, substitute Eqs. (9)–(11) in (7) and keep only terms of first order in the perturbations to obtain

$$L_0\delta p_n+\delta L p_{0n}=\lambda_{0n}\delta p_n+\delta\lambda_n p_{0n}. \tag{18}$$

Now introduce the scalar product of two functions $\varphi_1(x)$ and $\varphi_2(x)$,

$$(\varphi_1,\varphi_2)=\int_0^l \varphi_1(x)\varphi_2(x)\,dx, \tag{19}$$

and take the scalar product of p_{0n} with each term in Eq. (18):

$$(p_{0n},L_0\delta p_n)+(p_{0n},\delta L p_{0n})=\lambda_{0n}(p_{0n},\delta p_n)+\delta\lambda_n(p_{0n},p_{0n}). \tag{20}$$

Since the Laplacian operator L_0 is self-adjoint, the first term on the left-hand side may be rewritten as

$$(p_{0n},L_0\delta p_n)=\lambda_{0n}(p_{0n},\delta p_n). \tag{21}$$

The corresponding terms in Eq. (20) cancel, leaving

$$\delta\lambda_n=\frac{(p_{0n},\delta L p_{0n})}{(p_{0n},p_{0n})}. \tag{22}$$

Defining

$$N_n=(p_{0n},p_{0n}) \tag{23}$$

and substituting Eq. (17) we have

$$\delta\lambda_n=\frac{1}{2N_n}\int_0^l\frac{S'}{S}\frac{dp_{0n}^2}{dx}dx. \tag{24}$$

With the relation

$$\frac{S'}{S}=\frac{d}{dx}\ln\frac{S}{S_0}, \tag{25}$$

Eq. (24) may be integrated by parts to obtain

$$\delta\lambda_n=-\frac{1}{2N_n}\int_0^l \ln(S/S_0)\frac{d^2p_{0n}^2}{dx^2}dx \tag{26}$$

after taking Eqs. (8) into account. Use of Eqs. (14), (15), and (23) yields

$$\frac{\delta\lambda_n}{\lambda_{0n}}=-\frac{2}{l}\int_0^l \ln(S/S_0)\cos(2n\pi x/l)\,dx, \tag{27}$$

which is an explicit expression at the first perturbation order for the shifts in the eigenvalues.

The desired relation for the shifts in the natural frequencies is obtained by substituting Eqs. (11) and (12) in (5) to obtain $\delta\lambda_n/\lambda_{0n}=2\,\delta\omega_n/\omega_{0n}$ at first order in the perturbation. Equation (27) thus becomes

$$\frac{\delta\omega_n}{\omega_{0n}} = -\frac{1}{l}\int_0^l \ln(S/S_0)\cos(2n\pi x/l)\,dx. \tag{28}$$

Substitution of Eqs. (13)–(17) in (18) provides the differential equation for the corresponding variation of the mode shapes,

$$\left[\frac{d^2}{dx^2} + \left(\frac{n\pi}{l}\right)^2\right]\delta p_n = \frac{n\pi}{l}\frac{S'}{S}\sin(n\pi x/l) + \delta\lambda_n \cos(n\pi x/l), \tag{29}$$

where $d(\delta p_n)/dx = 0$ at $x = 0,l$.

To interpret Eq. (28) it is helpful to consider the approximation

$$\ln\frac{S}{S_0} = \ln\left(1 + \frac{\delta S}{S_0}\right) \simeq \frac{\delta S}{S_0}. \tag{30}$$

At first order in δS Eq. (28) becomes

$$\frac{\delta\omega_n}{\omega_{0n}} = -\frac{1}{l}\int_0^l \frac{\delta S}{S_0}\cos(2n\pi x/l)\,dx, \tag{31}$$

which is the result obtained by Rayleigh[8] from energy considerations. Now expand the shape perturbation in harmonics of the lowest spatial frequency component of the resonator, $2\pi/l$:

$$\frac{\delta S}{S_0} = \sum_{n=1}^{\infty} [a_n \cos(2n\pi x/l) + b_n \sin(2n\pi x/l)]. \tag{32}$$

Substitution in Eq. (31) yields

$$\frac{\delta\omega_n}{\omega_{0n}} = -\frac{a_n}{2}. \tag{33}$$

Two observations follow immediately at this level of approximation. The first is that the nth natural frequency of the resonator is affected only by the nth spatial frequency of the shape perturbation. The second is that only the even (cosine) terms in the expansion of the shape perturbation affect the natural frequencies. Odd spatial modulations (sine terms) have no effect on the natural frequencies. To shift the nth natural frequency alone it is therefore sufficient to choose

$$\frac{\delta S}{S_0} = a_n \cos(2n\pi x/l), \tag{34}$$

provided δS is sufficiently small that higher-order terms in Eq. (30) are unimportant.

We now return to Eq. (28) and let

$$\ln\frac{S}{S_0} = \sum_{n=1}^{\infty} [a_n \cos(2n\pi x/l) + b_n \sin(2n\pi x/l)], \tag{35}$$

which yields

$$\frac{\delta\omega_n}{\omega_{0n}} = -\frac{a_n}{2}, \tag{36}$$

the same as Eq. (33). Likewise, to shift a single natural frequency it is sufficient to retain only one cosine term in Eq. (35), such that

$$\ln(S/S_0) = a_n \cos(2n\pi x/l), \tag{37}$$

and therefore

$$S = S_0 \exp[a_n \cos(2n\pi x/l)]. \tag{38}$$

As shown numerically in Sec. II B, this result is accurate for larger values of δS than is obtained with Eq. (34). In summary, each coefficient a_n individually is related directly to the shift in the nth natural frequency, and one can thus tailor the natural frequency spectrum of the resonator by choosing the values of a_n accordingly and the values of b_n arbitrarily.

We conclude by using Eq. (29) to calculate the corresponding variation of the nth mode shape. With S given by Eq. (38), Eq. (27) yields $\delta\lambda_n = a_n(n\pi/l)^2$, the solution of Eq. (29) is $\delta p_n = -\frac{1}{8}a_n \cos(3n\pi x/l)$, and the eigenfunction $p_n = p_{0n} + \delta p_n$ is thus

$$p_n = \cos(n\pi x/l) - \tfrac{1}{8}a_n \cos(3n\pi x/l). \tag{39}$$

Section II B is devoted to comparing predictions based on Eq. (36) with numerical solutions of Eq. (4), for the resonator shapes defined by Eqs. (34) and (38), to determine the accuracy of the perturbation theory developed above.

B. Comparison with numerical results

On the basis of Eqs. (34) and (38) we consider here the variations of the natural frequencies associated with the two shape functions

$$\frac{S}{S_0} = \exp[a_1 \cos(2\pi x/l)], \tag{40}$$

$$\frac{S}{S_0} = 1 + a_1 \cos(2\pi x/l), \tag{41}$$

for which to leading order the same result for the fundamental natural frequency is obtained from Eq. (28):

$$\frac{\omega_1}{\omega_{01}} = 1 - \frac{a_1}{2}. \tag{42}$$

Note that Eq. (41) is the expansion of Eq. (40) through first order in a_1, corresponding to the approximation in Eq. (30). The asymptotic result in Eq. (42) is compared with solutions obtained numerically from Eq. (4) for S/S_0 given by Eqs. (40) and (41), and over the parameter range $-1 < a_1 < 1$. The values $a_1 \simeq \pm 1$ introduce positions along the axis at which $S \ll S_0$, a situation that is clearly beyond the limits of the Webster horn equation. However, the purpose here is to examine the accuracy of Eq. (28) as a solution of Eq. (4).

In Fig. 1(a) are presented the fundamental natural frequencies ω_1, normalized by the corresponding natural frequency ω_{01} for a cylindrical resonator, predicted by numerical solutions of Eq. (4) for Eqs. (40) (dashed line) and (41) (dot–dash line), and compared with Eq. (42) (solid line). For $|a_1| < 0.5$ the frequency shifts produced by both Eqs. (40) and (41) are in close agreement with Eq. (42), but Eq. (40) is in closer agreement for larger values of $|a_1|$.

The effects of Eqs. (40) and (41) on the second natural frequency ω_2 are presented in Fig. 1(b). Asymptotic theory predicts no effect on this natural frequency, i.e., $\omega_2/\omega_{01} = 2$ for all values of a_1. For $|a_1| = 0.5$ the shift in the second natural frequency is seen to be 2% for Eq. (40) and 6% for

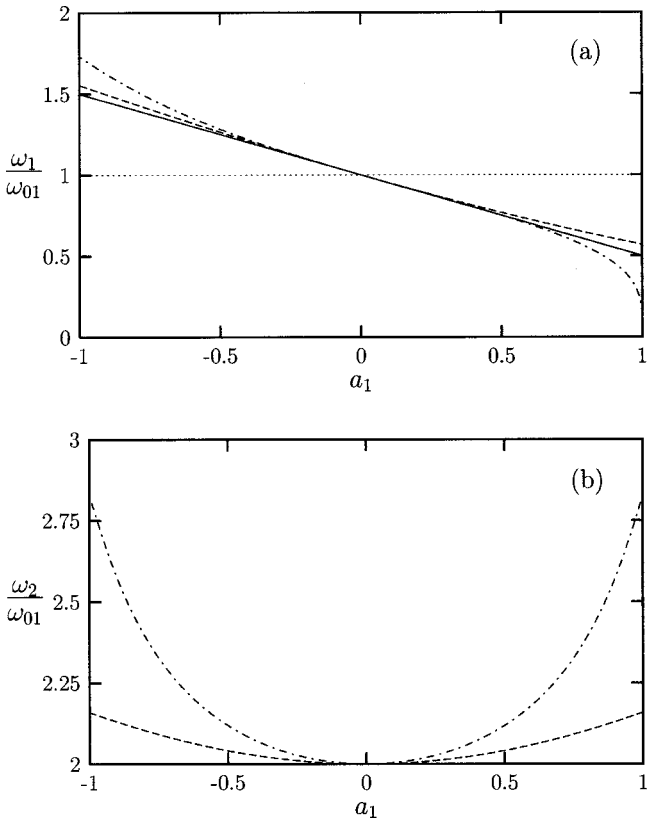


FIG. 1. (a) Fundamental and (b) second natural frequency shifts calculated numerically for Eqs. (40) (dashed line) and (41) (dot–dash line). The analytical result in (a) is Eq. (42) (solid line), and in (b) it is $\omega_2/\omega_{01}=2.0$.

Eq. (41), but this is to be compared with a 25% shift in the fundamental natural frequency in each case. As $|a_1|$ approaches unity, Eq. (41) introduces a substantial shift in the second natural frequency, whereas the shift introduced by Eq. (40) remains relatively small.

The asymptotic theory predicts that no natural frequency shifts result from the sine terms in Eqs. (32) and (35). We now investigate the accuracy of this prediction by considering the two shape functions

$$\frac{S}{S_0}=\exp[b_1\sin(2\pi x/l)], \tag{43}$$

$$\frac{S}{S_0}=1+b_1\sin(2\pi x/l). \tag{44}$$

For each of these functions Eq. (28) yields $\omega_1/\omega_{01}=1$, and this prediction is assessed in Fig. 2 via comparison with numerical solutions of Eq. (4) for $-1<b_1<1$. For $|b_1|<0.5$ the deviation from $\omega_1/\omega_{01}=1$ is only a few percent in each case. For $|b_1|>0.5$ the natural frequency shift introduced by Eq. (44) (dot–dash line) becomes substantial, whereas that due to Eq. (43) (dashed line) remains relatively small.

We conclude by investigating the shift in the second natural frequency due to the shape functions

$$\frac{S}{S_0}=\exp[a_2\cos(4\pi x/l)], \tag{45}$$

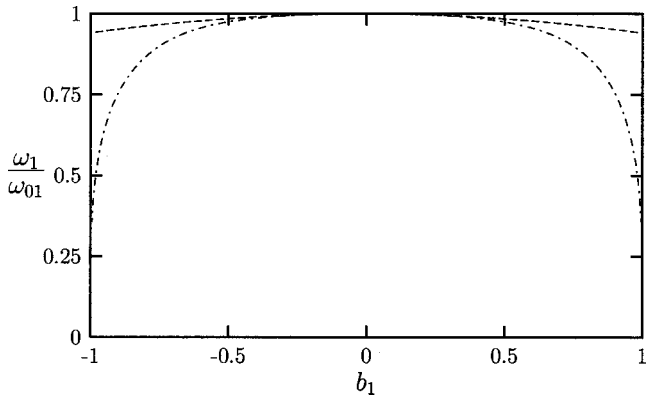


FIG. 2. Fundamental natural frequency shifts calculated numerically for Eqs. (43) (dashed line) and (44) (dot–dash line). The analytical result is $\omega_1/\omega_{01}=1.0$.

$$\frac{S}{S_0}=1+a_2\cos(4\pi x/l), \tag{46}$$

for which Eq. (28) predicts

$$\frac{\omega_2}{\omega_{01}}=2-a_2 \tag{47}$$

in each case. In Fig. 3(a) are shown the second natural frequencies ω_2 predicted by numerical solutions of Eq. (4) for Eqs. (45) and (46), and compared with Eq. (47). For $|a_2|<0.5$ the numerical solutions for both Eqs. (45) and (46) are in close agreement with Eq. (47), and for $|a_2|>0.5$ the shift due to Eq. (46) diverges more rapidly from Eq. (47), as was

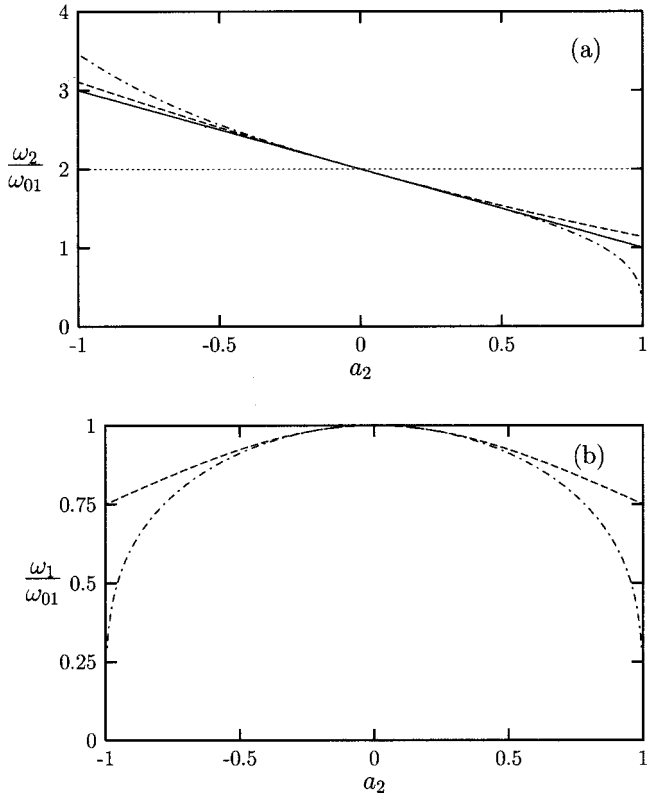


FIG. 3. (a) Second and (b) fundamental natural frequency shifts calculated numerically for Eqs. (45) (dashed line) and (46) (dot–dash line). The analytical result in (a) is Eq. (47) (solid line), and in (b) it is $\omega_1/\omega_{01}=1.0$.

the case for the fundamental natural frequency [recall Fig. 1(a)].

Equation (28) predicts no shift in the fundamental natural frequency for either Eq. (45) or (46), i.e., $\omega_1/\omega_{01}=1$. This prediction is evaluated in Fig. 3(b). Equations (45) and (46) introduce an approximately 10% shift in the fundamental natural frequency for $|a_2|=0.5$, compared with a 25% shift in the second natural frequency. For $|a_2|>0.5$ the shift produced by Eq. (46) (dot–dash line) becomes substantially greater than that due to Eq. (45) (dashed line).

In summary, Figs. 1–3 support the validity of Eq. (28) for predicting the natural frequency shifts due to variations in the shape function. Moreover, Eqs. (36) and (38) are established as simple yet accurate guidelines for specifying the shape functions that generate a desired distribution of the natural frequencies, provided these frequencies are reasonably close to those of a cylindrical resonator.

III. NONLINEAR THEORY

Lagrangian mechanics is used in the following to obtain coupled equations for the nonlinear interaction of modes in a resonator with varying cross section. The sound field is expanded in terms of normal modes predicted by linear theory, and in particular by the Webster horn equation. The time-varying displacement amplitudes of the modes are taken to be the generalized coordinates in the calculation of the Lagrangian. It is assumed that the natural frequencies differ significantly from those of a cylindrical resonator, in which case dynamical equations for just two coupled modes are sufficient to approximate the frequency response. An approximate solution is thus derived for the amplitude-frequency response and nonlinear resonance frequency shift of a resonator driven near its fundamental natural frequency. The result is compared with an appropriate numerical solution.

A. Methodology

We shall investigate nonlinear resonance frequency shifts by formulating the problem in terms of nonlinear interactions among the normal modes of the resonator. By ''normal modes'' we mean the modes of the associated linear problem. Although these are not normal modes of the nonlinear problem, we nevertheless retain this terminology. Specifically, we describe the sound field in the resonator by the particle displacement ξ expressed in terms of the following expansion:

$$\xi(x,t)=\sum_{n=1}^{\infty} q_n(t)\xi_n(x), \tag{48}$$

where q_n are displacement amplitudes. The dimensionless functions ξ_n are chosen to form a complete and orthogonal set that satisfies the boundary conditions

$$\xi_n(0)=\xi_n(l)=0 \tag{49}$$

at the rigid ends of the resonator. The most appropriate choice is the set that satisfies the Webster horn equation, expressed in terms of particle displacement:

$$\frac{d}{dx}\left(\frac{1}{S}\frac{d(S\xi_n)}{dx}\right)=-(\omega_n^2/c_0^2)\xi_n, \tag{50}$$

where ω_n is the natural frequency of the nth mode. Although the distinction between Lagrangian (material) and Eulerian (spatial) coordinates is irrelevant in the linear approximation, we note that x is henceforth considered to be the Lagrangian coordinate. A thorough discussion of relations between acoustical quantities expressed in terms of Lagrangian and Eulerian coordinates is provided by Hunt.[9]

Equation (48) permits identification of the functions q_n as generalized coordinates that satisfy Lagrange's equations,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_n}\right)-\frac{\partial L}{\partial q_n}=0, \tag{51}$$

where the dot indicates the time derivative of the coordinate, and

$$L=K-U \tag{52}$$

is the Lagrangian, with K and U the total kinetic and potential energies of the sound field, respectively. Once the normal modes ξ_n are obtained, K and U can be expressed in terms of q_n and $\dot{q}_n$, and Eq. (51) yields the dynamical equations for the system. The method just outlined is a common approach for describing vibrations of continuous systems.[10] A related approach has also been used to model the evolution of nonlinear Rayleigh waves.[11]

B. Normal modes

We may rewrite Eq. (50) in the form of the Liouville equation,[12]

$$\frac{d}{dx}\left(S\frac{d\xi_n}{dx}\right)+[R+(\omega_n^2/c_0^2)S]\xi_n=0, \tag{53}$$

where $R(x)=S''-(S')^2/S$. The normal modes ξ_n thus comprise a complete set of eigenfunctions that are orthogonal with weighting function S. We normalize them such that

$$\int_0^l \xi_m(x)\xi_n(x)S(x)\,dx=\frac{1}{2}S_0 l\,\delta_{mn}, \tag{54}$$

where δ_{mn} is the Kronecker delta. The normalization is chosen to ensure that the modal amplitudes for a cylindrical resonator ($S=S_0$) are unity, i.e.,

$$\xi_n^{\mathrm{cyl}}=\sin(n\pi x/l). \tag{55}$$

We note that when Eq. (4) is also rewritten in Liouville form, the eigenvalues given by Eq. (5) are found to be the same for both p_n and ξ_n, and the eigenfunctions p_n are orthogonal with weighting function S, as in Eq. (54). The two sets of normal modes may be related to one another through either of the following two expressions:

$$p_n=\frac{\alpha_n}{S}\frac{d(S\xi_n)}{dx}, \quad \xi_n=\beta_n\frac{dp_n}{dx}, \tag{56}$$

where the coefficients α_n and β_n are chosen to render the dimensions consistent and to account for the desired normalization. The first expression follows from the equations of

continuity and state, the second from the momentum equation.

C. Kinetic energy

Let $v(x,t)$ be the velocity distribution of the gas in the resonator. The kinetic energy of the entire volume of gas in the resonator is

$$K=\frac{\rho_0}{2}\int_0^l v^2(x,t)S(x)\,dx, \tag{57}$$

where ρ_0 is the Lagrangian density (which, because it is independent of space and time, is equivalent to the ambient Eulerian density). Now let the acoustic excitation result from shaking the resonator along its axis with velocity $v_0(t)$. We then have

$$v(x,t)=v_0(t)+u(x,t), \tag{58}$$

where

$$u=\frac{d\xi}{dt}=\dot{\xi} \tag{59}$$

is the acoustic particle velocity of the gas relative to the motion of the resonator. Equation (57) thus becomes

$$K=\frac{\rho_0}{2}\int_0^l(\dot{\xi}^2+2v_0\dot{\xi}+v_0^2)S(x)\,dx. \tag{60}$$

Substitution of Eq. (48), performing the integration, and taking Eq. (54) into account yields

$$K=\frac{m_0}{4}\sum_n\,(\dot{q}_n^2+2v_0e_n\dot{q}_n)+\frac{m}{2}v_0^2, \tag{61}$$

where

$$m=\rho_0\int_0^l S(x)\,dx,\quad m_0=\rho_0S_0l \tag{62}$$

are the mass of the gas in a noncylindrical and cylindrical resonator, respectively, and

$$e_n=\frac{2}{S_0l}\int_0^l\xi_n(x)S(x)\,dx \tag{63}$$

are the coefficients in the normal mode expansion of unity.

To interpret Eq. (61) it is helpful to introduce the normal mode expansion of the externally induced velocity field $v_0(t)$:

$$v_0(t)=v_0(t)\sum_n\,e_n\xi_n(x), \tag{64}$$

where the modal amplitudes are recognized to be $e_nv_0(t)$. In the case of a cylindrical resonator, for which the normal modes are given by Eq. (55), Eq. (63) yields

$$e_n^{\mathrm{cyl}}=\begin{cases}4/n\pi, & n\ \text{odd},\\ 0, & n\ \text{even}.\end{cases} \tag{65}$$

Furthermore, one may write $e_n=e_n^{\mathrm{cyl}}+O(\delta S)$.

Equations (65) indicate that shaking a cylindrical resonator at even multiples of the fundamental natural frequency has no effect on the sound field inside. The same is true for any shape function $S(x)$ that is symmetric about the midpoint $x=l/2$.

D. Potential energy

The potential (internal) energy per unit mass of an ideal gas is $U_m=c_vT$, where c_v is the specific heat at constant volume and T is temperature. Elimination of T using the equation of state gives $U_m=(\gamma-1)^{-1}P/\rho$, where P is pressure, ρ is density, and γ is the ratio of specific heats. The potential energy per unit Eulerian volume is $U_v=\rho U_m=P/(\gamma-1)$, and therefore the total potential energy stored in the sound field is

$$U=\frac{1}{\gamma-1}\int_0^l P(x^{\mathrm{E}},t)S(x^{\mathrm{E}})\,dx^{\mathrm{E}}, \tag{66}$$

where the integration is over the Eulerian coordinate x^{E} in a system attached to the resonator. We use the adiabatic relation

$$P/P_0=(\rho/\rho_0)^{\gamma} \tag{67}$$

to replace P by ρ in Eq. (66), where P_0 and ρ_0 are the ambient values of the pressure and density, respectively. The equation for mass conservation, relating the Lagrangian (x) and Eulerian (x^{E}) coordinates, is

$$\rho(x^{\mathrm{E}})S(x^{\mathrm{E}})dx^{\mathrm{E}}=\rho_0S(x)\,dx. \tag{68}$$

The coordinates themselves are related as follows:

$$x^{\mathrm{E}}=x+\xi,\quad dx^{\mathrm{E}}=(1+\partial\xi/\partial x)dx. \tag{69}$$

Substitution of Eqs. (67)–(69) in (66) yields, in terms of the Lagrangian coordinate x,

$$U=\frac{P_0}{\gamma-1}\int_0^l\frac{S^{\gamma}(x)\,dx}{[(1+\partial\xi/\partial x)S(x+\xi)]^{\gamma-1}}. \tag{70}$$

This relation is exact within the approximation of one-dimensional motion.

Now expand $S(x+\xi)$ in a Taylor series about $\xi=0$ to obtain

$$\left(1+\frac{\partial\xi}{\partial x}\right)S(x+\xi)=S(x)+\frac{\partial}{\partial x}\sum_{n=1}^{\infty}\frac{\xi^n}{n!}S^{(n-1)}(x), \tag{71}$$

where $S^{(n)}(x)=d^nS/dx^n$. Equation (70) thus becomes

$$U=\frac{P_0}{\gamma-1}\int_0^l\left[1+\frac{1}{S}\frac{\partial}{\partial x}\sum_{n=1}^{\infty}\frac{\xi^n}{n!}S^{(n-1)}(x)\right]^{-(\gamma-1)}S(x)\,dx, \tag{72}$$

which provides a means for expanding the potential energy as a power series in ξ^n.

We begin by considering the expansion through quadratic order:

$$U=P_0\int_0^l\left[\frac{S}{\gamma-1}-\frac{\partial}{\partial x}\left(S\xi+\frac{1}{2}S'\xi^2\right)+\frac{\gamma}{2}S\left(\frac{1}{S}\frac{\partial(S\xi)}{\partial x}\right)^2\right]dx+O(\xi^3). \tag{73}$$

The term containing $(S\xi+\frac{1}{2}S'\xi^2)$ integrates to zero because $\xi=0$ at $x=0,l$. The expansion of the potential energy is thus

$$U=U_0+U_2+U_3+U_4+\cdots, \tag{74}$$

where $U_n=O(\xi^n)$ and

$$U_0=\frac{\rho_0 c_0^2}{\gamma(\gamma-1)}\int_0^l S(x)\,dx=\frac{mc_0^2}{\gamma(\gamma-1)}, \tag{75}$$

$$U_2=\frac{\rho_0 c_0^2}{2}\int_0^l\left(\frac{1}{S}\frac{\partial(S\xi)}{\partial x}\right)^2 S(x)\,dx, \tag{76}$$

and the relation $c_0^2=\gamma P_0/\rho_0$ has been used. We now express U_2 in terms of the modal amplitudes q_n. Substitute Eq. (48) in (76) to obtain

$$U_2=\frac{\rho_0 c_0^2}{2}\sum_m\sum_n q_m q_n\int_0^l\frac{1}{S}\frac{d(S\xi_m)}{dx}\frac{d(S\xi_n)}{dx}dx \tag{77}$$

$$=-\frac{\rho_0 c_0^2}{2}\sum_m\sum_n q_m q_n\int_0^l\frac{d}{dx}\left(\frac{1}{S}\frac{d(S\xi_m)}{dx}\right)S\xi_n\,dx, \tag{78}$$

where the second expression results from integration by parts and use of Eqs. (49). Application of Eqs. (50) and (54) yields

$$U_2=\frac{m_0}{4}\sum_n \omega_n^2 q_n^2 \tag{79}$$

for the quadratic term in Eq. (74). Equation (79) accounts for the shape perturbation of the resonator through the shifts in the natural frequencies, i.e., via the difference between ω_n and ω_{0n}.

Higher-order terms in Eq. (74) account for nonlinear effects. Evaluation of the integrals in these terms requires explicit expressions for the normal modes ξ_n, which may be determined either numerically, or analytically with perturbation theory. Alternatively, since $U_{n+1}\ll U_n$ for small particle displacements, it is reasonable as a first approximation to evaluate U_3 and U_4 using the normal modes given by Eq. (55) for a cylindrical resonator, and to thus ignore the small correction due to the shape perturbation δS for a noncylindrical resonator. This is the approach we follow here. For a cylindrical resonator Eq. (70) reduces to

$$\frac{U^{\rm cyl}}{\rho_0 c_0^2 S_0}=\frac{1}{\gamma(\gamma-1)}\int_0^l\frac{dx}{(1+\partial\xi/\partial x)^{(\gamma-1)}} \tag{80}$$

and thus

$$\frac{U_3^{\rm cyl}}{\rho_0 c_0^2 S_0}=-\frac{1}{6}(\gamma+1)\int_0^l\left(\frac{\partial\xi}{\partial x}\right)^3 dx, \tag{81}$$

$$\frac{U_4^{\rm cyl}}{\rho_0 c_0^2 S_0}=\frac{1}{24}(\gamma+1)(\gamma+2)\int_0^l\left(\frac{\partial\xi}{\partial x}\right)^4 dx. \tag{82}$$

With Eq. (55), substitution of Eq. (48) in (81) and (82) yields, taking $U_3\simeq U_3^{\rm cyl}$ and $U_4\simeq U_4^{\rm cyl}$,

$$U_3=-(\gamma+1)\frac{\pi^3 m_0 c_0^2}{8l^3}\sum_{n_1+n_2=n_3} n_1 n_2 n_3 q_{n_1} q_{n_2} q_{n_3}, \tag{83}$$

$$U_4=(\gamma+1)(\gamma+2)\frac{\pi^4 m_0 c_0^2}{192l^4}\left(3\sum_{n_1+n_2=n_3+n_4}+4\sum_{n_1+n_2+n_3=n_4}\right)n_1 n_2 n_3 n_4 q_{n_1} q_{n_2} q_{n_3} q_{n_4}, \tag{84}$$

where all integers n_i are positive.

E. Dynamical equations

Here we develop dynamical equations for the amplitude of the sound field due to drive frequencies ω in the vicinity of the fundamental natural frequency ω_1. Since the natural frequencies of noncylindrical resonators are not integer multiples of each other, there is phase mismatch between the wave at the drive frequency and its harmonics. This mismatch impedes interaction between different modes, and it is assumed here to be sufficiently strong that we need only consider interaction between the first and second modes, $n=1,2$.

Coupled equations for the wave interaction are obtained from Eqs. (51) and (52), which may be rewritten

$$\frac{d}{dt}\left(\frac{\partial K}{\partial\dot{q}_n}\right)+\frac{\partial U}{\partial q_n}=0, \tag{85}$$

where through fourth order in q_n we may write

$$U=U_2+U_3+U_4. \tag{86}$$

With the interactions limited to modes $n=1,2$, Eqs. (83) and (84) reduce to

$$U_3=-(\gamma+1)\frac{\pi^3 m_0 c_0^2}{4l^3}q_1^2 q_2, \tag{87}$$

$$U_4=(\gamma+1)(\gamma+2)\frac{\pi^4 m_0 c_0^2}{64l^4}(q_1^4+4q_1^2 q_2^2), \tag{88}$$

and only the first two terms of the summations in Eqs. (61) for K and (79) for U_2 are retained. Equation (85) thus yields the coupled equations

$$\ddot{q}_1+2\delta_1\dot{q}_1+\omega_1^2 q_1+e_1\dot{v}_0=(\gamma+1)\frac{\pi^3 c_0^2}{l^3}\left[q_1 q_2-(\gamma+2)\frac{\pi}{8l}(q_1^3+2q_1 q_2^2)\right], \tag{89}$$

$$\ddot{q}_2+2\delta_2\dot{q}_2+\omega_2^2 q_2+e_2\dot{v}_0=(\gamma+1)\frac{\pi^3 c_0^2}{2l^3}\left[q_1^2-(\gamma+2)\frac{\pi}{2l}q_1^2 q_2\right]. \tag{90}$$

The loss factors

$$\delta_n=\frac{\omega_n}{2Q_n} \tag{91}$$

were introduced *ad hoc*, defined in terms of the quality factor Q_n for mode n, such that δ_n is the bandwidth of the given mode.

We now express the acceleration of the resonator $\dot{v}_0$ in the form

$$\dot{v}_0=\tfrac{1}{2}Ae^{i\omega t}+\text{c.c.}, \tag{92}$$

where A is its complex amplitude. Since $\omega \simeq \omega_1$ (drive frequencies near the fundamental natural frequency) we seek a steady-state solution of Eqs. (89) and (90) in the form

$$q_1 = \tfrac{1}{2} d_1 e^{i\omega t} + \text{c.c.}, \tag{93}$$

$$q_2 = d_0 + \tfrac{1}{2} d_2 e^{i2\omega t} + \text{c.c.} \tag{94}$$

The desired amplitude of the fundamental mode is d_1. The static component d_0 and the second-harmonic amplitude d_2, both generated in the second mode through q_1^2 on the right-hand side of Eq. (90), are $O(d_1^2)$. The terms $q_1 q_2$ and q_1^3 on the right-hand side of Eq. (89) thus produce $O(d_1^3)$ corrections to d_1. Because the two remaining nonlinear terms, $q_1 q_2^2$ and $q_1^2 q_2$, lead to higher-order corrections, they are ignored. Making the substitutions in Eqs. (89) and (90), and taking $e_n = e_n^{\text{cyl}}$ as defined in Eq. (65), one obtains

$$(\delta_1 + i\Delta_1) i\omega_1 d_1 + \frac{2}{\pi} A = \frac{(\gamma+1)\omega_{01}^3}{4c_0} \left[d_1^* d_2 + 2 d_0 d_1 - \frac{3(\gamma+2)\omega_{01}}{16 c_0} |d_1|^2 d_1 \right] \tag{95}$$

for the fundamental component, where the static and second-harmonic components are given in terms of d_1 by

$$d_0 = \frac{(\gamma+1)\omega_{01}^3 |d_1|^2}{4\omega_2^2 c_0}, \tag{96}$$

$$d_2 = \frac{(\gamma+1)\omega_{01}^3 d_1^2}{i16\omega_1 c_0 (\delta_2 + i\mu\Delta_2)}. \tag{97}$$

Dependence on the drive frequency ω enters through the small detuning parameters

$$\Delta_n(\omega) = n\omega - \omega_n \tag{98}$$

in Eqs. (95) and (97). The approximation $\omega \simeq \omega_1$ was used, except in the expressions for Δ_n; i.e., the relations $\omega^2 - \omega_1^2 \simeq 2\omega_1 \Delta_1$ and $(2\omega)^2 - \omega_2^2 \simeq 4\mu\omega_1\Delta_2$ were employed, where $\mu = \frac{1}{2}(1 + \omega_2/2\omega_1)$ is a coefficient that is close to unity.

F. Resonance frequency shift

It is convenient to introduce here the particle velocity

$$u_1 = i\omega d_1 \simeq i\omega_1 d_1. \tag{99}$$

Since the mode shape of the particle velocity is the same as that of the displacement, Eq. (55), we recognize $|u_1|$ as the peak particle velocity of the fundamental mode, which is obtained at $x = l/2$. Substitution of Eqs. (96) and (97) in (95) and converting to particle velocity yields

$$u_1 = -\frac{2A/\pi}{Z(\omega)}, \tag{100}$$

where

$$Z(\omega) = \delta_1 + i\Delta_1 + \frac{(\gamma+1)^2}{64} \frac{|u_1|^2}{c_0^2} \frac{\omega_{01}^6}{\omega_1^5} \left[\frac{\omega_1}{\delta_2 + i\mu\Delta_2} + i2\left(\frac{2\omega_1}{\omega_2}\right)^2 - i3\left(\frac{\gamma+2}{\gamma+1}\right)\frac{\omega_1^2}{\omega_{01}^2} \right]. \tag{101}$$

Equation (100) may be converted into a real cubic equation in $|u_1|^2$, and an analytical solution is thus available for the response $|u_1|$ as a function of drive frequency ω for a resonator with cross-sectional area variation $S(x)$. Recall that $\omega_{01} = \pi c_0 / l$ is the fundamental natural frequency of a cylindrical resonator, and $\omega_n = n\omega_{01} + \delta\omega_n$ $(n = 1,2)$ are the natural frequencies of the noncylindrical resonator under consideration, where $\delta\omega_n$ is given by Eq. (28).

The quantity $-2A/\pi$ in Eq. (100) may be regarded as the acceleration due to a body force (per unit mass) applied to the gas inside the resonator. Consequently, Z is an effective mechanical impedance (per unit mass). We define the nonlinear resonance frequency ω_{res} to be the value of the drive frequency ω at which the imaginary part of Z vanishes:

$$\text{Im}\, Z(\omega_{\text{res}}) \equiv 0, \tag{102}$$

and thus

$$|u_1|_{\max} \simeq \frac{2|A|/\pi}{Z(\omega_{\text{res}})}. \tag{103}$$

An approximate expression for ω_{res} may be obtained by considering the relation

$$\delta_2 + i\mu\Delta_2 = \omega_2/2Q_2 + i\mu[2(\omega - \omega_1) + (2\omega_1 - \omega_2)]. \tag{104}$$

For low losses $(Q_2 \gg 1)$, drive frequencies near the fundamental natural frequency $(\omega \simeq \omega_1)$, and strong phase mismatch (relatively large $|2\omega_1 - \omega_2|$), the terms $\omega_2/2Q_2$ and $2(\omega - \omega_1)$ may be neglected to obtain

$$\delta_2 + i\mu\Delta_2 \simeq i\omega_1[1 - (\omega_2/2\omega_1)^2]. \tag{105}$$

Substitution in Eq. (101) yields

$$Z(\omega) \simeq \delta_1 + i\left\{ \omega - \omega_1 - \frac{(\gamma+1)^2}{64} \frac{|u_1|^2}{c_0^2} \frac{\omega_{01}^6}{\omega_1^5} \times \left[\frac{1}{1 - (\omega_2/2\omega_1)^2} - 2\left(\frac{2\omega_1}{\omega_2}\right)^2 + 3\left(\frac{\gamma+2}{\gamma+1}\right)\frac{\omega_1^2}{\omega_{01}^2} \right] \right\}. \tag{106}$$

From Eqs. (102) and (103) we have

$$\frac{\omega_{\text{res}}}{\omega_1} \simeq 1 + \frac{(\gamma+1)^2}{64} \frac{|u_1|_{\max}^2}{c_0^2} \frac{\omega_{01}^6}{\omega_1^6} \left[\frac{1}{1 - (\omega_2/2\omega_1)^2} - 2\left(\frac{2\omega_1}{\omega_2}\right)^2 + 3\left(\frac{\gamma+2}{\gamma+1}\right)\frac{\omega_1^2}{\omega_{01}^2} \right] \tag{107}$$

and

$$|u_1|_{\max} \simeq 2|A|/\pi\delta_1. \tag{108}$$

Whether ω_{res} increases or decreases as the drive amplitude is increased (referred to as hardening or softening behavior, respectively) is thus determined by the natural frequencies of the resonator, which are in turn determined by the linear theory presented in Sec. II. The first term in the square brackets of Eq. (107) is associated with second-harmonic generation, the second with static deformation, and the third with cubic nonlinearity [see Eq. (95), where the contributions appear in the same sequence].

The singularity in Eq. (107) at $\omega_2/2\omega_1=1$ reveals that in the neighborhood of this transition point one obtains

$$\omega_{\rm res}>\omega_1 \quad \text{for } 2\omega_1>\omega_2, \tag{109}$$

$$\omega_{\rm res}<\omega_1 \quad \text{for } 2\omega_1<\omega_2. \tag{110}$$

The direction of the nonlinear resonance frequency shift is thus primarily due to interaction with the second-harmonic component. As seen by substituting Eq. (105) in (97), the conditions in Eqs. (109) and (110) determine the sign of the second harmonic, and therefore the sign of the second-harmonic contribution to the reactance in Eq. (101). For $2\omega_1=\omega_2$, which is outside the approximation leading to Eq. (105), the contribution of the second harmonic to Eq. (101) is resistive rather than reactive, representing the loss associated with resonant second-harmonic generation.

Equation (107) also reveals that there is a contribution to $\omega_{\rm res}$ due to static deformation and cubic nonlinearity, and these effects introduce an asymmetry in the amplitude-frequency response curves that is consistent with graphical results based on Eqs. (100) and (101), as shown below. Moreover, Eq. (107) indicates that there is a value of $\omega_2/2\omega_1$ at which the contributions due to second-harmonic generation, static deformation, and cubic nonlinearity counteract each other. To illustrate this point it is helpful to set $\omega_2=2\omega_{01}$, corresponding to a situation where the shape perturbation affects only the fundamental natural frequency ω_1, such that

$$\frac{\omega_{\rm res}}{\omega_1}\simeq 1+\frac{(\gamma+1)^2}{64}\frac{|u_1|^2_{\max}}{c_0^2}\frac{\omega_{01}^6}{\omega_1^6}\left[\frac{1}{1-\omega_{01}^2/\omega_1^2}+\left(\frac{\gamma+4}{\gamma+1}\right)\frac{\omega_1^2}{\omega_{01}^2}\right], \quad \omega_2=2\omega_{01}. \tag{111}$$

The term containing $\gamma+4$ is due to the combined effects of static deformation and cubic nonlinearity. From Eq. (107) one finds that the contribution due to cubic nonlinearity is approximately twice the magnitude of the contribution due to static deformation. The sign associated with cubic nonlinearity is positive, and with static deformation it is negative; therefore the net effect of both contributions is a positive frequency shift. For

$$\omega_1/\omega_{01}=\sqrt{3/(\gamma+4)} \tag{112}$$

the terms in the square brackets of Eq. (111) cancel, and $\omega_{\rm res}$ does not deviate from ω_1 as the drive amplitude is increased. With $\gamma=1.4$, Eq. (112) yields $\omega_1/\omega_{01}=\sqrt{5}/3$, which corresponds to detuning by 25%. Although this perturbation may exceed the approximations leading to Eq. (111), one may nevertheless anticipate the existence of a resonator shape for which there is no nonlinear resonance frequency shift at this level of approximation.

We conclude with a brief remark about losses. The terms $\delta_n\dot{q}_n$ in Eqs. (89) and (90) account for absorption but not dispersion, which is adequate for describing losses due to viscosity and heat conduction in the volume of the resonator. In contrast, boundary-layer losses are accompanied by dispersion, which introduces shifts in the natural frequencies. An estimate of boundary-layer effects may be deduced from the complex attenuation coefficient for propagation in a duct.[7] That is, boundary-layer effects may be taken into account by adding positive imaginary parts to the loss factors in Eq. (101), such that $\delta_n\rightarrow\delta_n+i\epsilon_n$ (with $\delta_n=\epsilon_n$ when only boundary-layer effects are considered). If we make the assumptions leading to Eq. (105), the only change in Eq. (107) is that ϵ_1/ω_1 is subtracted from the right-hand side. The new term merely indicates the slightly lower natural frequency due to the boundary layer, with the nonlinear response unaffected at this order of approximation.

G. Examples

As noted above, Eq. (100) may be recast as a real cubic polynomial in terms of $|u_1|^2$, which possesses either one or three roots. The latter case corresponds to multivalued solutions for the amplitude-frequency response curves, indicating that $|u_1|$ can jump from one value to another as the drive frequency ω is varied, and the system exhibits hysteresis. Since the analytical solution is cumbersome, the cubic equation was solved numerically to obtain the results presented below.

To assess the accuracy of Eq. (100), comparisons are made with finite difference solutions of the complete nonlinear wave equation for one-dimensional motion in a resonator of arbitrary shape. The linear terms in the wave equation are equivalent to the Webster horn equation with a viscous loss term. All nonlinear terms are retained in the equations of continuity and momentum and the adiabatic equation of state. The drive is an applied acceleration, as in Eq. (100). Derivation of the wave equation and description of the solution algorithm are presented in Ref. 3.

The resonator shape is chosen to be Eq. (38) with $n=1$:

$$S=S_0\exp[a_1\cos(2\pi x/l)]. \tag{113}$$

We let $a_1=\pm 0.2$, for which Eqs. (12) and (28) yield

$$\omega_1/\omega_{01}=1.0\mp 0.1 \quad \text{for } a_1=\pm 0.2, \tag{114}$$

$$\omega_2/\omega_{01}=2.0. \tag{115}$$

Next it is necessary to consider the loss factors defined by Eq. (91). It is assumed that losses are due only to viscosity and heat conduction in the volume of the resonator. In this case the loss factors increase quadratically with frequency, indicating that the quality factors are inversely proportional to frequency, and we may write

$$Q_1=(\omega_{01}/\omega_1)Q_{01}, \quad Q_2=\tfrac{1}{2}Q_{01}, \tag{116}$$

where Q_{01} is the quality factor at frequency ω_{01}. In the present examples we let $Q_{01}=100\pi=314$, which was a nominal value used in previous work[3] to obtain agreement with experiment. The value of the gas constant is $\gamma=1.4$. Finally, we choose

$$\hat{A}=|A|/l\omega_{01}^2 \tag{117}$$

as the dimensionless drive amplitude. Introducing the corresponding displacement amplitude $|D|=|A|/\omega^2$ and velocity amplitude $|V|=|A|/\omega$ of the resonator, one may interpret the

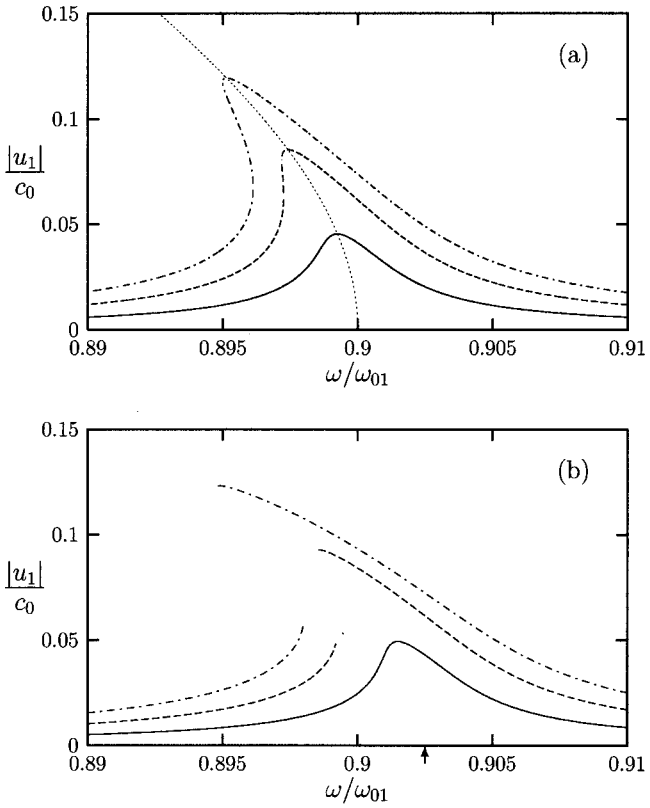


FIG. 4. (a) Analytical and (b) numerical results for amplitude-frequency response curves for $x=l/2$, $\omega_1/\omega_{01}=0.9$, $\omega_2/\omega_{01}=2.0$, and $10^4\hat{A}=0.3$ (solid lines), 0.6 (dashed lines), and 0.9 (dot–dash lines). The arrow in (b) identifies the natural frequency obtained numerically, $\omega_1/\omega_{01}=0.9025$.

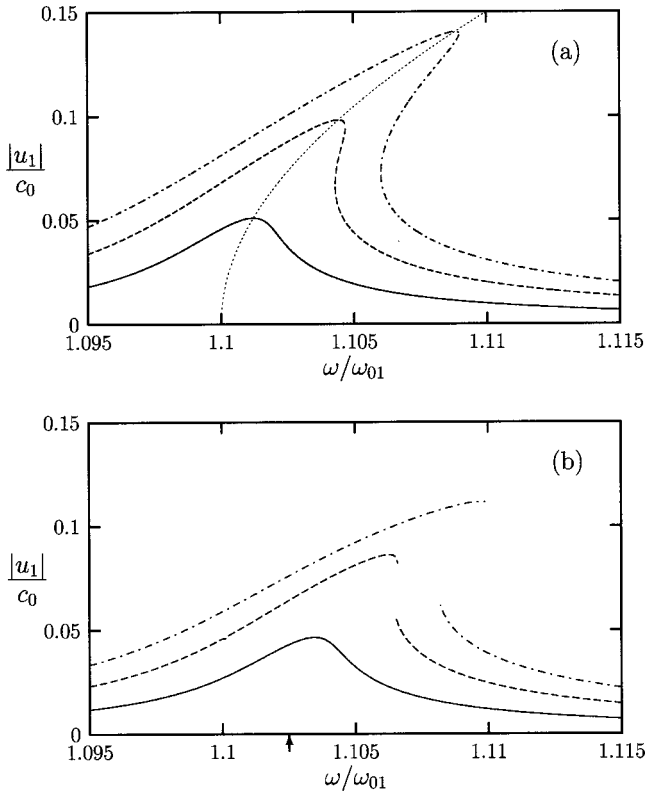


FIG. 5. (a) Analytical and (b) numerical results for amplitude-frequency response curves for $x=l/2$, $\omega_1/\omega_{01}=1.1$, $\omega_2/\omega_{01}=2.0$, and $10^4\hat{A}=0.5$ (solid lines), 1.0 (dashed lines), and 1.5 (dot–dash lines). The arrow in (b) identifies the natural frequency obtained numerically, $\omega_1/\omega_{01}=1.1025$.

drive amplitude used here, after taking $\omega\sim\omega_{01}$, as follows: $\hat{A}\sim|D|/l\sim|V|/\pi c_0$.

Figure 4 shows amplitude–frequency response curves for $a_1=0.2$ and three values of $\hat{A}$. Analytical results given by Eqs. (100) and (101) are presented in Fig. 4(a), numerical solutions[3] in Fig. 4(b). Since the quantity $|u_1|$ in the analytical solution is the peak amplitude of the fundamental mode, Fig. 4(b) shows the particle velocity amplitude calculated at $x=l/2$. The numerical solution yields $\omega_1/\omega_{01}=0.9025$ [identified by the arrow in Fig. 4(b)] from Eq. (4) for Eq. (113) with $a_1=0.2$, which is slightly greater than the asymptotic result $\omega_1/\omega_{01}=0.9$ used in Fig. 4(a). After this discrepancy is taken into account, which amounts to translating the curves in Fig. 4(a) to the right accordingly, the analytical results are seen to be in reasonable quantitative agreement with the numerical results. Gaps appearing in the curves in Fig. 4(b) correspond to unstable states that cannot be obtained from the numerical solution. The dotted line (backbone curve) tracking the resonance in Fig. 4(a) is obtained from Eqs. (101) and (102). Equation (111) provides a reasonable approximation of the dotted line up to $|u_1|_{\max}/c_0\sim0.1$.

Response curves for $a_1=-0.2$ are presented in Fig. 5. In general, the discussion of Fig. 4 applies to Fig. 5 as well. Here, the numerical solution yields $\omega_1/\omega_{01}=1.1025$ for the natural frequency, indicated by an arrow in Fig. 5(b), whereas the asymptotic result is $\omega_1/\omega_{01}=1.1$. Overall agreement of analytical predictions with numerical results is again reasonable.

IV. CONCLUSION

Asymptotic methods were used to obtain analytical expressions for the shifts in the natural frequencies ω_n and in the nonlinear resonance frequencies ω_{res} of acoustical resonators with slowly varying cross sections. Although approximate, the expressions reveal the main effects that influence these shifts. In particular, they permit the shifts to be determined completely by the function $S(x)$ defining the shape of the resonator.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research.

[1] A. H. Nayfeh and D. T. Mook, *Nonlinear Oscillations* (Wiley, New York, 1979).
[2] 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).
[3] Yu. 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).
[4] Y.-D. Chun and Y.-H. Kim, ''Numerical analysis for nonlinear resonant oscillations of gas in axisymmetric closed tubes,'' J. Acoust. Soc. Am. **108**, 2765–2774 (2000).
[5] G. W. Swift, ''Thermoacoustic engines,'' J. Acoust. Soc. Am. **84**, 1145–1180 (1988).
[6] S. Karpov and A. Prosperetti, ''Nonlinear saturation of the thermoacoustic

instability,'' J. Acoust. Soc. Am. **107**, 3130–3147 (2000).
[7] A. D. Pierce, *Acoustics* (Acoustical Society of America, New York, 1989).
[8] Lord Rayleigh (J. W. Strutt), *The Theory of Sound*, 2nd ed. (Dover, New York, 1945), Vol. 2, Sec. 265.
[9] F. V. Hunt, ''Notes on the exact equations governing the propagation of sound in fluids,'' J. Acoust. Soc. Am. **27**, 1019–1039 (1955), Sec. 4.
[10] L. Meirovitch, *Fundamentals of Vibrations* (McGraw-Hill, New York, 2001), Sec. 9.8.
[11] E. Yu. Knight, M. F. Hamilton, Yu. A. Il'inskii, and E. A. Zabolotskaya, ''General theory for the spectral evolution of nonlinear Rayleigh waves,'' J. Acoust. Soc. Am. **102**, 1402–1417 (1997).
[12] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Part 1, p. 719.

Acoustic chaos in a duct with two separate sound sources

W. G. Dong,[a] and X. Y. Huang,[b]
Center for Mechanics of Micro-Systems (CMMS), School of Mechanical and Production Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798

Q. L. Wu[c]
BSWA Technology Co., Ltd., 17 Zhongguancun Road, Beijing 10080, China

(Received 18 September 1998; revised 15 May 2000; accepted 12 April 2001)

Nonlinear interaction of two plane acoustic waves in the (0,0) mode of a square duct is investigated from the viewpoint of chaotic dynamics. Phase-space portraits are reconstructed from time series obtained in an experiment. It is demonstrated that limit sets formed by the phase space trajectories are ''attractors.'' The largest Lyapunov exponent and correlation dimension of these attractors are calculated, and the results indicate that these attractors are chaotic. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1379086]

PACS numbers: 43.25.Rq [MAB]

I. INTRODUCTION

Interest in nonlinear dynamics of physical systems has grown since the 1980s.[1,2] Acoustics has been among the first areas of physics to present examples of chaotic dynamics.[3] The origin of chaos in acoustics can be attributed to at least three sources: the nonlinearities of the fluid dynamics in the medium itself, the acoustic generator, and the reflection, impedance, or reception of the acoustic waves. A review of chaotic dynamics in some acoustics problems was given by Lauterborn and Holzfuss.[4] Most studies on the acoustic chaos have been focused on the chaotic noise from bubbles and cavitation in fluid (usually in water), in which the nonlinear behavior of the bubbles is believed to be the source of period doubling and chaotic acoustic phenomena in fluid. A very recent study on the subharmonic route to chaos in acoustic cavitation was reported by Cabeza.[5] Music chaos is another kind of acoustic chaos. Gibiat[6] studied period doubling and chaos in clarinet-like instruments. Embedding space or pseudo-phase-space techniques were used to look for qualitative behavior of the clarinet-like resonator.

So far, most research works on acoustic chaos have been carried out only in the acoustic cavitations, bubbles, and music acoustics; very little has been studied on chaos due to nonlinear interaction of acoustic waves in air. In the present paper, an experiment was performed to investigate the nonlinear interaction of plane acoustic waves in the (0,0) mode of an air-filled square duct with rigid walls. Commonly used nonlinear dynamical analysis methods, including phase portraits, Lyapunov exponents, and fractal dimensions, were applied to time series measured in the experiment. The goal of the present research is to provide an alternative insight into the nonlinear interaction of acoustic waves from the viewpoint of nonlinear dynamical systems.

II. BACKGROUND

A. Nonlinear interaction of acoustic waves

The nonlinear interaction of acoustic waves includes basically two cases: one is the nonlinear interaction of directional acoustic beams, the other is the nonlinear interaction of acoustic waves in ducts. The theory of the nonlinear interaction of plane waves was originally worked out by Westervelt,[7,8] which ultimately led to the practical applications of parametric arrays. Since Westervelt's work, the nonlinear interaction of acoustic beams has been well-examined theoretically and experimentally.[9,10] Many applications of parametric arrays involve strong directional acoustic beams, which are widely used in underwater acoustics.

The study of nonlinear interaction of acoustic waves in ducts is of a special importance, although the analysis of nonlinear interaction of acoustic waves in ducts is simpler than the case of directional acoustic beams. This is because acoustic waveguides are widely used in industries and laboratories. A better understanding of the propagation and interaction of acoustic waves in ducts is helpful to find solutions for practical applications. During the last three decades, nonlinear interaction of plane acoustic waves in ducts has received a great deal of attention. Both the collinear interactions and noncollinear interactions have been examined.[11–18] The latest report on the experiments about the nonlinear interaction of acoustic waves in ducts can be found in Refs. 17–20.

It has been shown[21] that nonlinearity of air accounts not only for the generation of sum and difference waves, but also for the generation of complete combination of $m\omega_1 \pm n\omega_2$ (m, n are integers), harmonics, and subharmonics. From the viewpoint of system dynamics, the acoustic field in a duct is a complex nonlinear system with a broadened frequency spectrum. One may not be able to make any long-term prediction to the behavior of this system, and chaos could occur in this system under certain conditions.

Most investigations of the nonlinear interaction of acoustic waves have been focused on the generation of the difference and sum frequency waves. Very little research has been directed to the nonlinear dynamic behavior of the systems. The main concern of the present paper is the collinear interaction of plane waves in a square duct. The time series measured from the experiment are studied from the viewpoint of nonlinear dynamics.

[a]Electronic mail: mwgdong@ntu.edu.sg
[b]Electronic mail: mxhuang@ntu.edu.sg
[c]Electronic mail: qlwu@bswa.com.cn

 0001-4966/2001/110(1)/120/7/$18.00

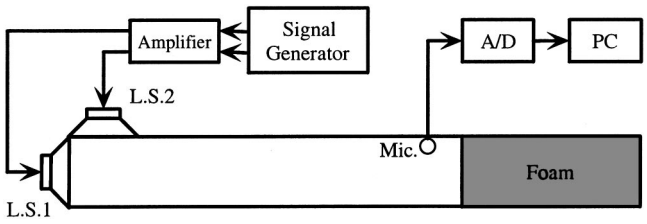


FIG. 1. Schematic of the experimental setup.

B. Chaos

An important feature of chaos is ''sensitive dependence on initial conditions,'' or SDIC. This means that small changes in the state of a system will grow at an exponential rate and will dominate the behavior of the system. SDIC can be quantified with Lyapunov exponents, which are the long-time average exponential rates of divergence of nearby states. If a system has at least one positive Lyapunov exponent, the system is chaotic. The larger the positive exponent, the more chaotic the system will be, that is, the shorter the time scale of system predictability. A system may possess any number of Lyapunov exponents; however, to confirm chaos one only needs to confirm that the largest Lyapunov exponent is positive. Hence, estimation of the dominant exponent is especially important.

Recent experiments on a wide variety of systems in engineering, chemistry, physics, and biology have demonstrated that chaos is common in nonequilibrium systems. Chaotic behavior in diverse systems is often mathematically similar, even for systems that have entirely different physical mechanisms and very different levels of complexity.[1,2] This means that one can apply common analysis techniques to time series obtained from different experiments.

In order to identify chaotic motions from experimentally measured time series, the most often used methods are: phase portraits, fractal dimensions, and Lyapunov exponents, which have also been used in the present study.

III. EXPERIMENT

The experimental setup is shown in Fig. 1. A square steel duct, which had a length of 5.4 m, a wall thickness of 1.5 mm, and inside cross section of 3.8×3.8 cm^2, was used in the experiment. The cutoff frequency of this duct was 5044 Hz. To simplify the analysis, an xoz coordinate system was established such that the z axis was along the duct and the x axis was perpendicular to the duct; the origin o was at bottom of one end of the duct where a loudspeaker is mounted (see Fig. 2).

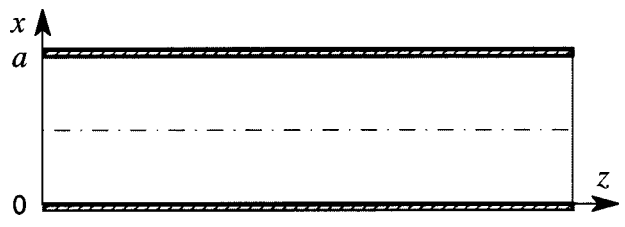


FIG. 2. Geometry of square duct.

Two MRD 1520 compression drivers were used to generate high-intensity sound waves. One driver (LS1) was mounted at the end of the duct, and the other (LS2) was mounted at the top wall near the end of the duct. A 1/4-in. microphone was used to measure the sound pressure in the duct. There were 18 holes for placing microphones with a space of 180 mm along the center line of the top wall. In order to assure a progressive wave motion in the z direction, the duct was terminated with 1.8-m-long foam. For a 300-Hz pure tone, with a measured standing wave ratio $s=1.65$, the corresponding absorption coefficient of the foam to 300-Hz pure tone was 94%. Thus, the reflected sound waves had sufficiently small amplitudes at frequencies above 300 Hz.

A data acquisition board was used to sample the sound pressure from the microphones. The commercial software LABVIEW was used as the interface between a PC and the data acquisition board to collect and save the data as ASCII text files for further analyses.

The primary waves were two tones generated by the drivers. The frequency of the signal applied to LS1 was $f_1=830$ Hz, and the frequency of the signal applied to LS2 was $f_2=1000$ Hz. Both the primary acoustic waves had the sound-pressure level of 135 dB (re: 20μ Pa). The sampling frequency was $f_s=12\,800$ Hz. The number of data in each saved file was $N=4096$. In the following analysis, time series $x(i)=\{x_i, i=1,2,...,N\}$ is used to denote the sampled sound pressure in each data file, and symbol x_i, $i=1,2,...,N$ is used to stand for the value of sampled sound pressure at discrete time i in each time series.

Time history, frequency spectrum, and autocorrelation function are three basic techniques often used in analysis of time series. The diagram of time history in the present study can be obtained by plotting $x(i)$ versus time $t(i)=i/f_s$, $i=1,2,...,N$. The frequency spectrum can be obtained by means of the fast Fourier transform (FFT) algorithm. The autocorrelation function $A(\tau)$ can be calculated as

$$A(\tau)=\frac{\frac{1}{N}\sum_{i=1}^{N}(x_{i+\tau}-\overline{x})(x_i-\overline{x})}{\frac{1}{N}\sum_{i=1}^{N}(x_i-\overline{x})^2},\quad \tau=0,1,...,N, \tag{1}$$

where $\overline{x}$ is the average of the time series.

Figures 3 and 4 show the typical time history and frequency spectrum of the sound pressure measured at $z=1.8$ m. Figure 5 shows the autocorrelation function of the data. It can be seen, from Fig. 4, that the frequency spectrum of the sound pressure contains broadened spectral lines. The frequency spectrum includes not only the two primary components at f_1 and f_2, but also components at combination frequencies of f_1 and f_2 [i.e., f_1-f_2, f_1+f_2, f_1+2f_2, $2f_1+f_2$, $2(f_1+f_2)$, etc.]. It can also be seen that there is a peak at 500 Hz which corresponds to the 1/2 subharmonics of $f_1=1000$ Hz. These properties indicate that there are complex nonlinear elements in the system and the data are not periodic. The autocorrelation function $A(\tau)$ shown in Fig. 5 also demonstrates that the data are not periodic, because $A(\tau)\to0$ as $\tau\to\infty$. The time lag where $A(\tau)$ drops to $1/e$ of its initial value is $\tau=0.0003$ s.

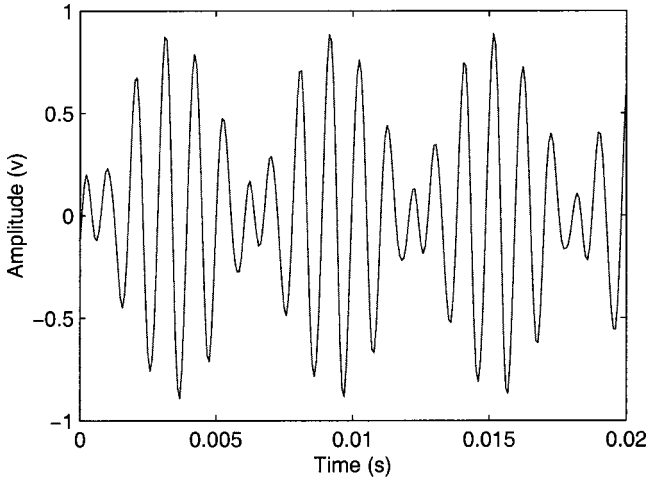


FIG. 3. Time history of sound pressure measured at $z=1.8$ m.

IV. PHASE PORTRAITS FROM TIME SERIES

The dynamical equations of a physical system can often be written in the following form:

$$\frac{d\mathbf{X}(t)}{dt}=G(\mathbf{X}(t)), \tag{2}$$

where $\mathbf{X}(t)$ is an n-component vector in the phase space of the system, and G stands for n nonlinear functions of these components. It is well-known that much could be learned about dynamical behavior of a system from analysis of trajectories in the phase space. For a practical physical system, the phase space often has a large number of components, i.e., the dimension of the phase space is very large. However, most dynamical systems involve only a small number of important degrees of freedom. One can often in practice restrict the study to a low-dimensional phase space which only includes those important degrees of freedom. Such a phase space is usually called embedded phase space.

In most experimental investigations, one can only get time series which are observations of a single variable of the system. In order to examine the dynamic properties of the system, the time-delay technique is usually used to reconstruct multidimensional phase portraits from the time

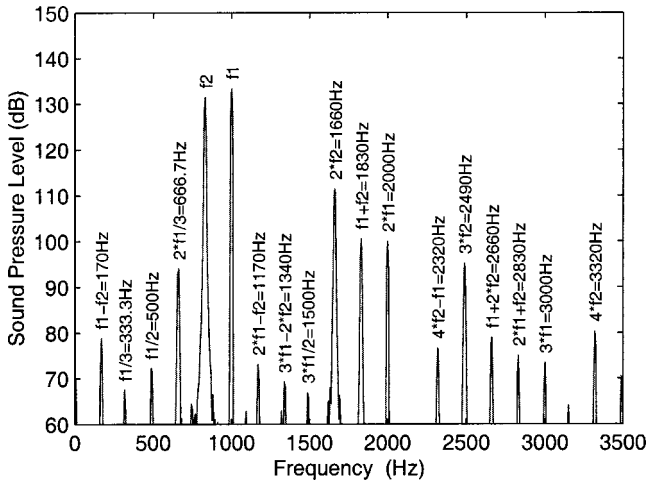


FIG. 4. Frequency spectrum of sound pressure measured at $z=1.8$ m.

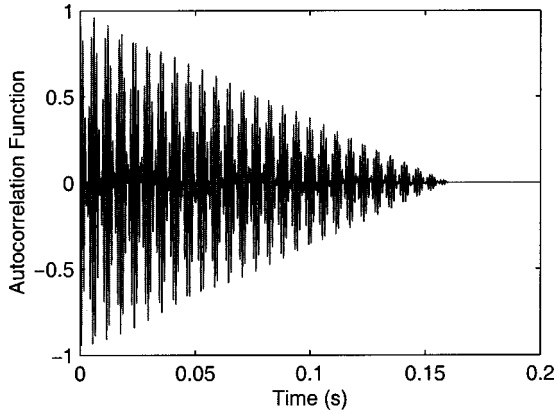


FIG. 5. Autocorrelation function of sound pressure measured at $z=1.8$ m.

series.[22–24] Take an N-point time series $x(i)$ as the example; the reconstructed trajectory, $\mathbf{X}$, can be expressed as a matrix

$$\mathbf{X}=(\mathbf{X}_1 \quad \mathbf{X}_2 \quad \cdots \quad \mathbf{X_M})^T, \tag{3}$$

where $\mathbf{X_k}$, $k=1,2,...,M$, is the state of the system at discrete time k, and is given by

$$\mathbf{X_k}=(x_k \quad x_{k+J} \quad \cdots \quad x_{k+(m-1)J}), \quad k=1,2,...,M \tag{4}$$

where J is the time delay on reconstruction delay, and m is the embedding dimension. The integers M, N, m, and J satisfy the following equation:

$$M=N-(m-1)J. \tag{5}$$

Thus, an m-dimension phase portrait is reconstructed from the time series $x(i)$. The embedding dimension m is usually estimated in accordance with Taken's theorem, i.e., $m>2n$, where n is the dimension of the attractor. Strictly speaking, the phase portrait obtained by this procedure gives an embedding of the original manifold and is called pseudo-phase portrait.[22,23] It has been shown that if the embedding dimension m is sufficiently large, the pseudo-phase portrait constructed by the method of time delays will in principle have the same properties as a phase portrait constructed from the independent variables.[22,23] For visualization of the reconstructed phase portrait, $m=2$ or $m=3$ is often used, while higher m is used for the further study to determine the largest Lyapunov exponent and fractal dimension of the attractor. The choice of the time delay J is almost but not completely arbitrary. A good approximation of J is equal to the discrete time lag where the autocorrelation function of the time series drops to $1/e$ of its initial value.

In the present study, the embedding dimension m is taken to be 2 and four different time delays $J=4$, 10, 20, and 40 are used to obtain the phase portraits, of which $J=4$ is equal to the discrete time lag where the autocorrelation function of the time series drops to $1/e$ of its initial value. Thus, two-dimensional vectors $\mathbf{X_k}=(x_k,x_{k+J})$, $k=1,2,...,M$, are constructed from the time series data $x(i)$. Figures 6, 7, 8, and 9 show the two-dimensional phase portraits obtained by plotting $x(i+J)$ versus $x(i)$ for different J. The time series data were measured at $z=1.8$ m.

In these phase portraits, it can be seen that the trajectories lie on a limit set in the embedded phase space for each time delay, which demonstrates that these phase portraits are

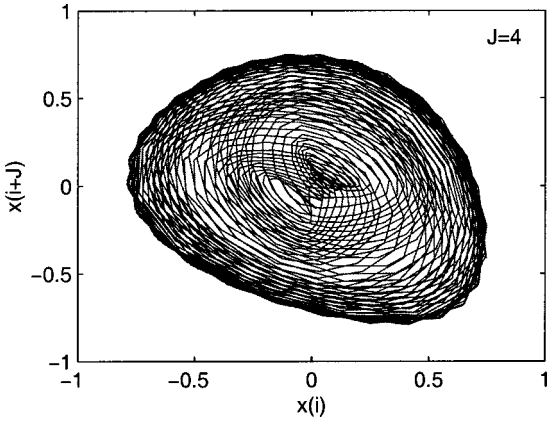


FIG. 6. Phase portrait reconstructed from the sound pressure measured at $z=1.8$ m (time lag $J=4$).

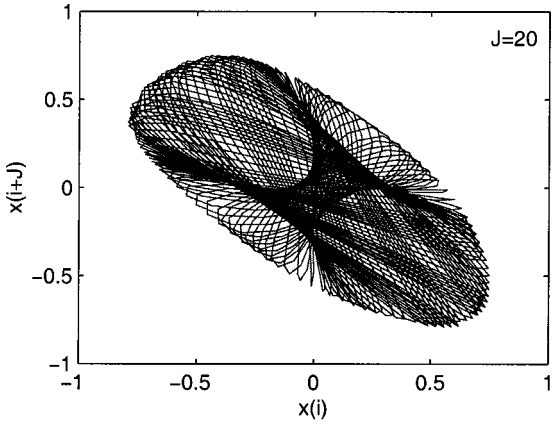


FIG. 8. Phase portrait constructed from the sound pressure measured at $z=1.8$ m (time lag $J=20$).

attractors.[1,2] In order to show these attractors to be strange or chaotic, Lyapunov exponents and fractal dimensions must be determined.

V. LYAPUNOV EXPONENTS AND FRACTAL DIMENSIONS

The most widely used criteria to quantify chaos in a system are the Lyapunov exponents and fractal dimensions. Positive Lyapunov exponents indicate the sensitive dependence on initial conditions. If a system has at least one positive Lyapunov exponent, the system is chaotic. A system may possess many Lyapunov exponents, but only the largest Lyapunov exponent needs to be checked in order to confirm chaos.

The fractal dimension is a noninteger dimension. A noninteger dimension is a hallmark of a chaotic attractor. There are many types of definitions and measures of fractal dimensions, of which the correlation dimension is the commonly used one.[25]

There are many algorithms that are applicable to estimate the largest Lyapunov exponent from time series.[26–34] The method used in the present paper was developed by Rosenstein *et al.*,[26] which is fast, easy to implement, and robust to changes in the parameters, such as the embedding dimension, the size of data set, the reconstruction delay, and the noise level. Furthermore, the correlation dimension can also be calculated simultaneously.

The largest Lyapunov exponent of a chaotic attractor is defined in the following equation:[26]

$$d(t)=Ce^{\lambda_1 t}, \tag{6}$$

where $d(t)$ is the average divergence of trajectories with nearby initial conditions on the attractor at time t, C is a constant that normalizes the initial separation, and λ_1 is the largest Lyapunov exponent.

The first step of Rosenstein's algorithm involves reconstructing the attractor dynamics from a single time series by using the method described in Eqs. (3)–(5). The algorithm then locates the nearest neighbor of each point on the trajectories. The nearest neighbor, $\mathbf{X}_{\hat{\mathbf{j}}}$, is found by searching for the point that minimizes the distance to the particular reference point $\mathbf{X}_\mathbf{j}$. This can be expressed as[26]

$$d_j(0)=\min_{\mathbf{x}_{\hat{\mathbf{j}}}}\|\mathbf{X}_\mathbf{j}-\mathbf{X}_{\hat{\mathbf{j}}}\| \quad j=1,2,...,M, \tag{7}$$

where $d_j(0)$ is the initial distance from the jth point to its nearest neighbor, $\|\ \|$ denotes the Euclidean norm, and M is the number of reconstructed vectors in the phase space. The

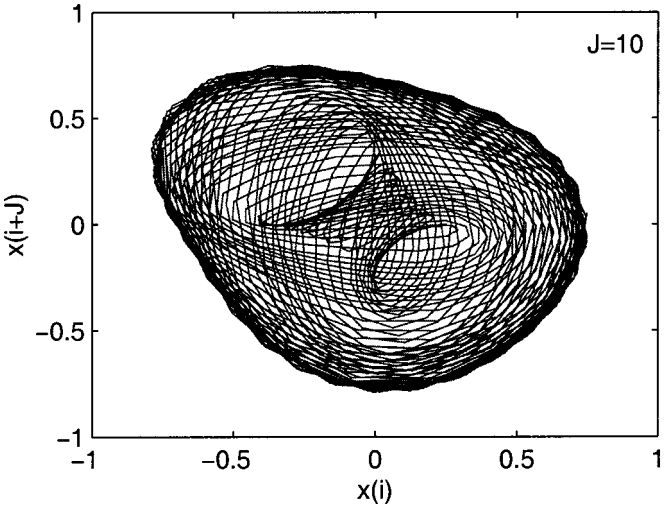


FIG. 7. Phase portrait constructed from the sound pressure measured at $z=1.8$ m (time lag $J=10$).

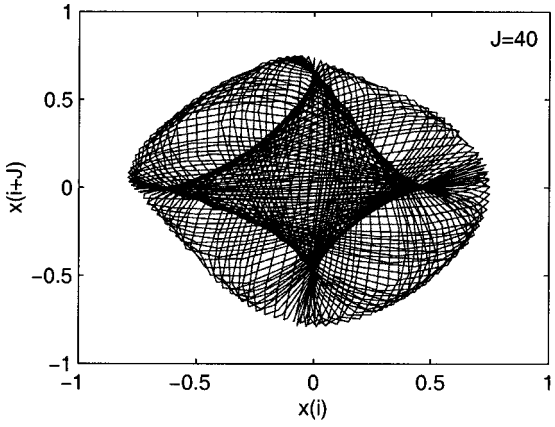


FIG. 9. Phase portrait constructed from the sound pressure measured at $z=1.8$ m (time lag $J=40$).

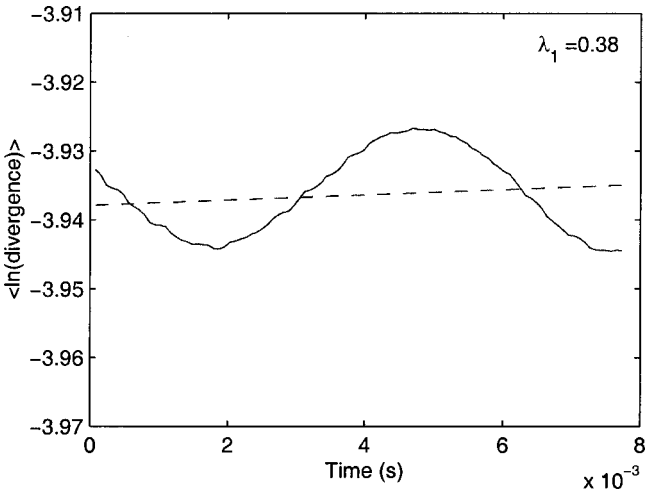


FIG. 10. ⟨ln(*divergence*)⟩ versus time for $m=10$, and $J=4$.

largest Lyapunov exponent is then estimated as the mean rate of separation of the nearest neighbors.

Let the nearest neighbors evolute with time along the trajectories. Assume that the jth pair of nearest neighbors diverges approximately at a rate given by the largest Lyapunov exponent;[26] that is,

$$d_j(i) \approx C_j e^{\lambda_1 (i\Delta t)} \quad j=1,2,...,M, \tag{8}$$

where C_j is the initial separation, i is the discrete time with which the nearest neighbors evolute, and Δt is the time interval between samples, $\Delta t = 1/f_s$. Taking the logarithm of both sides of Eq. (8) yields[26]

$$\ln d_j(i) \approx \ln C_j + \lambda_1 (i\Delta t) \quad j=1,2,...,M. \tag{9}$$

Equation (9) represents a set of approximately parallel lines (for $j=1,2,...,M$), with a slope proportional to λ_1. Taking the average of $\ln d_j(i)$ over all values of j, the largest Lyapunov exponent λ_1 then equals to the slope of a line defined by[26]

$$\langle \ln d_j(i) \rangle = \lambda_1 (i\Delta t), \tag{10}$$

where $\langle\ \rangle$ denotes the average over all values of j.

Figures 10 and 11 show the plots of $\langle \ln d_j(i) \rangle$ vs $(i\Delta t)$.

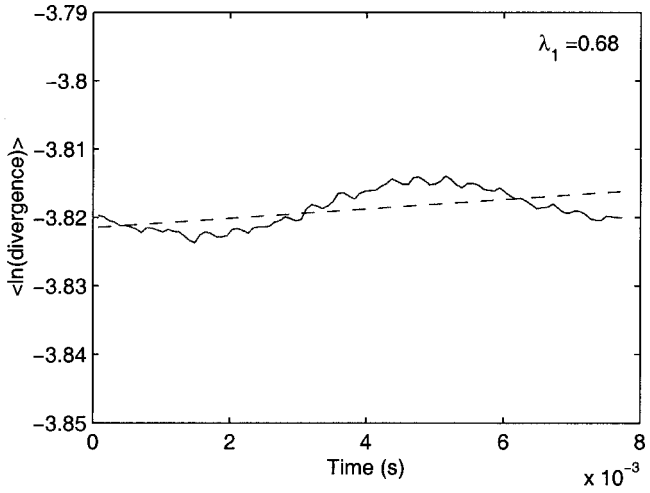


FIG. 11. ⟨ln(*divergence*)⟩ versus time for $m=10$, and $J=10$.

TABLE I. The calculated λ_1 and D_c for different J ($m=10$).

J	λ_1	D_c
10	0.89	2.19
15	0.61	2.18
20	0.39	2.04
25	0.24	1.98
30	0.39	2.28
35	0.31	2.02
40	0.25	2.28

The solid lines are the calculated results, and the dashed lines are the lines that fit the calculated results in a least-square sense. The largest Lyapunov exponent λ_1 is equal to the slope of the dashed lines. The analyzed time series $x(i)$ is the same as presented in the previous sections. Two different time delays $J=4$ and $J=10$ are used in the calculation. In each case, embedding dimension remains constant, $m=10$. The calculated largest Lyapunov exponents are $\lambda_1=0.38$ for $J=4$, and $\lambda_1=0.68$ for $J=10$, respectively. The positive largest Lyapunov exponents indicate that the attractors are chaotic.

Tables I and II list the calculated λ_1 for different m and J. It can be seen from these results that the largest Lyapunov exponents are all positive, which indicates the existence of chaos.

Another important quantitative measure of an attractor is the correlation dimension. The correlation dimension estimates the complexity of a system, which can be determined from the correlation function $C(R)$.[25] Let R denote a small positive value; for a reconstructed attractor $\mathbf{X}$, the two-point correlation function is defined as[25]

$$C(R) = \frac{1}{M(M-1)} \sum_{i=1}^{M} \sum_{j \neq i}^{M} \theta[R - d(\mathbf{X_i} - \mathbf{X_j})], \tag{11}$$

where M is the number of reconstructed vectors, $\theta(\ )$ is the Heaviside function, and $d(\mathbf{X_i}-\mathbf{X_j})$ represents the distance between two m-dimensional vectors. The Euclidean norm can be used to measure the distance between two vectors. $C(R)$ is interpreted as the fraction of pairs of points that is separated by a distance less than or equal to R.

The correlation dimension of an attractor is defined in the following equation:[2]

$$C(R) = \lim_{R \to 0} a R^{D_c}, \tag{12}$$

where a is a constant value, D_c is the correlation dimension, and J is the time delay. The correlation dimension can now

TABLE II. The calculated λ_1 and D_c for different m ($J=10$).

m	λ_1	D_c
5	0.89	2.19
6	0.81	2.19
7	0.78	2.25
8	0.75	2.29
9	0.72	2.39
10	0.68	2.22

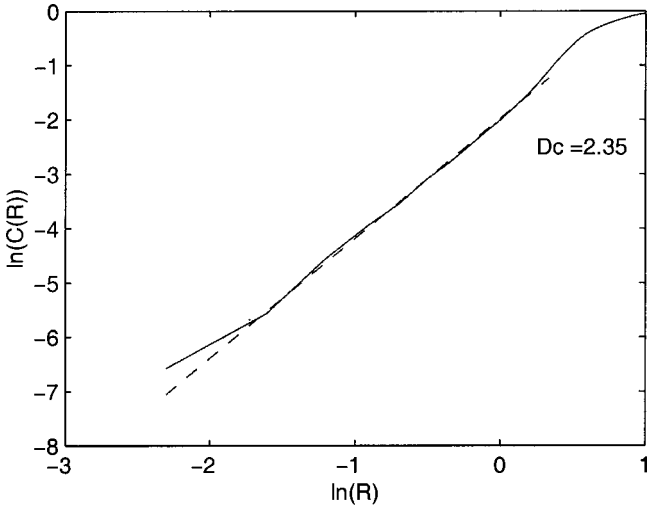


FIG. 12. $\ln C(R)$ vs $\ln R$ for $m=10$, and $J=4$.

be estimated from the slope of the linear part of the $\log[C(R)]$ vs $\log(R)$ curve according to

$$D_c = \lim_{R\to 0} \frac{\partial \log[C(R)]}{\partial \log(R)}. \tag{13}$$

Figures 12 and 13 show the curves of $\log[C(R)]$ versus $\log(R)$ for two different time delays $J=4$ and $J=10$. The solid lines are the calculated results. The dashed lines are the lines that fit the calculated results in a least-square sense. In each figure, the correlation dimension is determined by the slope of the dashed line. Embedding dimension m is taken equal to 10. The calculated correlation dimensions are $D_c=2.15$ for $J=4$ and $D_c=2.5$ for $J=10$.

Tables I and II list the calculated D_c for different m and J. It can be seen from these results that the system has noninteger dimensions, which means that the attractors shown in Figs. 6, 7, 8, and 9 are ''strange,'' or chaotic.

VI. CONCLUSIONS

Experiments on the nonlinear interaction of plane acoustic waves in a square duct have been performed. The measured time series data of the sound pressure have been analyzed by constructing the phase-space portraits, evaluating the largest Lyapunov exponent and calculating the fractal dimension. The results show that the phase-space trajectories define a limit set in the phase space that is an ''attractor.'' This attracting set is shown to be strange or chaotic—nearby trajectories separate exponentially on the average. Moreover, the fractal dimension of the attractor has also been calculated, which gives the degree of the system complexity.

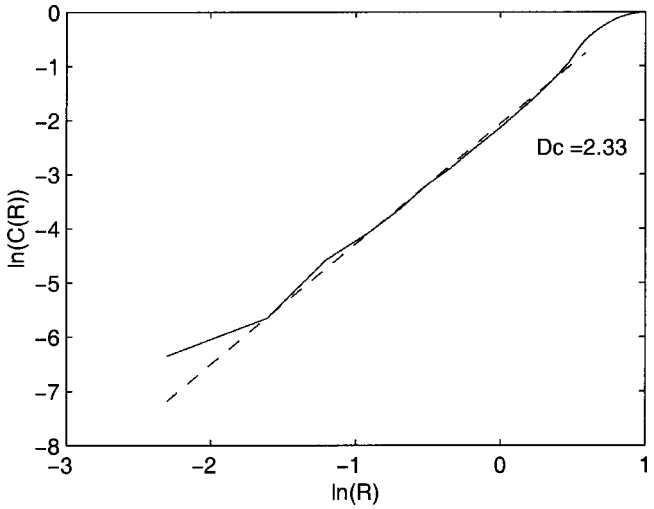


FIG. 13. $\ln C(R)$ vs $\ln R$ for $m=10$, and $J=10$.

[1] S. H. Strogatz, *Nonlinear Dynamics and Chaos* (Addison-Wesley, Singapore, 1996).
[2] F. C. Moon, *Chaotic and Fractal Dynamics* (Wiley, New York, 1992).
[3] W. Lauterborn and E. Cramer, ''Subharmonic route to chaos observed in acoustics,'' Phys. Rev. Lett. **47**, 1445–1447 (1981).
[4] W. Lauterborn and J. Holzfuss, ''Acoustic chaos,'' Int. J. Bifurcation Chaos Appl. Sci. Eng. **1**(1), 13–26 (1991).
[5] C. Cabeza, A. C. Sicardi-Schifino, C. Negreira, and G. Montaldo, ''Experimental detection of a subharmonic route to chaos in acoustic cavitation through the tuning of a piezoelectric cavity,'' J. Acoust. Soc. Am. **103**, 3227–3229 (1998).
[6] V. Gibiat, ''Phase space representations of acoustical musical signals,'' J. Sound Vib. **123**, 529–536 (1988).
[7] P. J. Westervelt, ''Parametric end-fire array,'' J. Acoust. Soc. Am. **32**, 934–935 (1960).
[8] P. J. Westervelt, ''Parametric acoustic array,'' J. Acoust. Soc. Am. **35**, 535–537 (1963).
[9] B. K. Novikov, O. V. Rudenko, and V. I. Timoshenko, *Nonlinear Underwater Acoustics*, translated by Robert T. Beyer and Mark F. Hamilton (The Acoustical Society of America, New York, 1987).
[10] M. F. Hamilton and D. T. Blackstock, *Nonlinear Acoustics* (Academic, New York, 1998).
[11] A. H. Nayfeh and M. S. Tsai, ''Nonlinear wave propagation in acoustically lind ducts,'' J. Sound Vib. **35**, 77–89 (1974).
[12] A. H. Nayfeh and M. S. Tsai, ''Nonlinear acoustic propagation in two-dimensional ducts,'' J. Acoust. Soc. Am. **55**, 1166–1172 (1974).
[13] A. H. Nayfeh, ''Nonlinear propagation of a wave packet in a hard-walled circular duct,'' J. Acoust. Soc. Am. **57**, 803–809 (1975).
[14] P. G. Vaidya and K. S. Wang, ''Nonlinear propagation of complex sound fields in rectangular ducts. Part I. The self-excitation phenomenon,'' J. Sound Vib. **50**, 29–42 (1977).
[15] J. H. Ginsberg, ''Finite amplitude two-dimensional waves in a rectangular duct induced by arbitrary periodic excitation,'' J. Acoust. Soc. Am. **65**, 1127–1133 (1979).
[16] J. H. Ginsberg and H. C. Miao, ''Finite amplitude distortion and dispersion of a nonplanar mode in a waveguide,'' J. Acoust. Soc. Am. **80**, 911–920 (1986).
[17] M. F. Hamilton and J. A. TenCate, ''Sum and difference frequency generation due to noncollinear wave interaction in a rectangular duct,'' J. Acoust. Soc. Am. **81**, 1703–1712 (1987).
[18] M. F. Hamilton and J. A. TenCate, ''Finite amplitude sound near cutoff in higher-order modes of a rectangular duct,'' J. Acoust. Soc. Am. **84**, 327–334 (1988).
[19] W. Dong, Q. Wu, and S. F. Ling, ''Application of nonlinear interaction of sound waves in sound reproduction,'' Proceedings of the 16th International Congress on Acoustics/135th meeting of the Acoustical Society of America **I**, 549–550 (1998).
[20] W. Dong and Q. Wu, ''Audio sound reproduction based on nonlinear interaction of acoustic waves,'' J. Audio Eng. Soc. **47**, 602–606 (1999).
[21] F. H. Fenlon, ''Extension of the Bessel–Fubini series for a multiple frequency CW acoustic source of finite amplitude,'' J. Acoust. Soc. Am. **51**, 284–289 (1972).
[22] F. Takens, in *Lecture Notes in Mathematics*, edited by D. A. Rand and L. S. Young (Springer, Berlin, 1981), pp. 366–381.
[23] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, ''Geometry from time series,'' Phys. Rev. Lett. **45**(9), 712–716 (1980).
[24] J. C. Roux, R. H. Simoyi, and H. L. Swinney, ''Observation of a strange attractor,'' Physica D **8**, 257–266 (1983).
[25] P. Grassberger and I. Procaccia, ''On the characterization of strange attractors,'' Phys. Rev. Lett. **50**, 346–349 (1983).
[26] M. T. Rosenstein, J. J. Collins, and C. J. D. Luca, ''A practical method for

calculating largest Lyapunov exponents from small data sets,'' Physica D **65**, 117–134 (1993).
[27] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, ''Determining Lyapunov exponents from a time series,'' Physica D **16**, 285–317 (1985).
[28] S. Sato, M. Sano, and Y. Sawada, ''Practical methods of measuring the generalized dimension and the largest Lyapunov exponent in high dimensional chaotic systems,'' Prog. Theor. Phys. **77**, 1–5 (1987).
[29] J. D. Farmer and J. J. Sidorowich, ''Predicting chaotic time series,'' Phys. Rev. Lett. **59**, 845–848 (1987).
[30] M. Casdagli, ''Nonlinear prediction of chaotic time series,'' Physica D **35**, 335–356 (1989).
[31] H. D. I. Abarbanel, R. Brown, and J. B. Kadtke, ''Prediction in chaotic nonlinear systems: Methods for time series with broadband Fourier spectra,'' Phys. Rev. A **41**(4), 1782–1807 (1990).
[32] J. Deppich, H. U. Bauer, and T. Geisel, ''Hierarchical training of neural networks and prediction of chaotic time series,'' Phys. Lett. A **158**, 57–63 (1991).
[33] S. Ellner, A. R. Gallant, D. F. McCaffrey, and D. W. Nychka, ''Convergence rates and data requirements for jacobian based estimates of Lyapunov exponents from data,'' Phys. Lett. A **153**, 357–363 (1991).
[34] D. J. Wales, ''Calculating the rate of loss information from chaotic time series by forecasting,'' Nature (London) **350**, 485–488 (1991).

Acoustical nonlinearity of an electroacoustic cell containing electrolyte

N. Tankovsky[a)]
Physical Department, Sofia University, Blvd. James Bourchier 5, Sofia 1126, Bulgaria

(Received 4 August 2000; accepted for publication 27 March 2001)

The nonlinear mechanical properies of an electrolytic, electroacoustic cell have been analyzed. Three main nonlinear mechanisms are present. (1) The nonlinear electric current in a circuit with a time-varying capacitance of the electrical double layer, modulated by the applied electric voltage. (2) The nonlinear time dependencies of the ion charges and of the local conductance at the interface membrane solution, governed by a Boltzmann distribution of the ion concentration. (3) The quadratic electrostatic forces of attraction, exciting a second acoustic harmonic. The interplay of these mechanisms is complicated. An analytical expression for the time dependence of the membrane displacements of the electroacoustic cell is obtained and the harmonics amplitudes are analyzed by numerical evaluations. © *2001 Acoustical Society of America.*
[DOI: 10.1121/1.1375844]

PACS numbers: 43.25.Ts, 43.38.Ar [MAB]

I. INTRODUCTION

A new type of capacitive ultrasound transducer, based on the electrostatic forces occurring at the interface electrolyte–dielectric, when an external ac electric field is applied has been recently reported.[1] The construction of the electroacoustic cell is presented schematically in Fig. 1(a), where 1 is a thin, metal-foil electrode, 2 is a thin elastic membrane, 5 is a fixed solid metal electrode, 4 is a thin isolating dielectric layer, 3 and 3′ are the electrical double layers at the two interfaces, correspondingly. The space distribution of the electric field in the cell is shown in Fig. 1(b). Obviously the geometry of the cell allows one to use a one-dimensional approximation, where the coordinate x is normal to the electrode plates. The electroacoustic cell can be treated as a capacitive, electrostatic transducer.[2] However, experimental evidence has been obtained,[3] which shows that besides the second harmonic, which is characteristic for the standard electrostatic transducers, the electrolyte cell also generates the fundamental (even when no dc electric field is applied) and higher acoustic harmonics. The purpose of the present work is to analyze the nonlinear properties of the electroacoustic cell and to explain the excitation of acoustic harmonics.

The time varying capacitor $C_i(t)$ comprises a metal-foil electrode, isolated from the electrolyte by a thin elastic membrane, and a second virtual electrode disposed at a distance X, defined by the coordinate of the ''center of mass'' of all ionic charges $Q(x,t)$ in the solution as follows:

$$X(t)=\frac{\int_0^{D/2} xQ(x,t)dx}{\int_0^{D/2} Q(x,t)dx}. \tag{1}$$

As seen from Fig. 1(b) the electric field distribution passes a zero point in the middle of the cell, i.e.,

$$E\left(\frac{D}{2},t\right)=0. \tag{2}$$

Hence the integration is in the range from the interface membrane–electrolyte at $x=0$ to the middle of the cell $x=D/2$, where the electric field drops to zero.

The two electrodes of the ionic capacitor are attracted to one another by the electrostatic force, whereas the displacement of the elastic electrode is balanced by the elastic force. The friction and the inertia forces acting on the elastic electrode are smaller and can be neglected. The mobility of the second virtual electrode is much higher and is limited solely by the friction of the charged particles in the solution. The inertia forces acting upon the virtual ionic electrode should be taken into account as well since its velocity is varying in time. Physically it is clear that the velocity is much higher in the initial moment when the electric field is applied and displacement currents are dominant compared to a later moment when slower ion migration currents dominate. We can assume that the electrostatic force is balanced by the elastic deformation of the membrane electrode, on the one side, and by the frictional and inertia forces of the ionic electrode, on the other side. So we can write down the following simple equation:

$$l\xi(t)=r\frac{dX(t)}{dt}+\rho\frac{d^2X(t)}{dt^2}. \tag{3}$$

Here ξ is the displacement of the membrane along the x coordinate, l is the elastic constant of the membrane, r is an effective ion friction coefficient, and ρ is the density of the solution.

By differentiating (1) we find that

$$\frac{dX(t)}{dt}=\frac{\int_0^{d/2} x\frac{dQ(x,t)}{dt}dx}{q(t)}-\frac{X(t)}{q(t)}\int_0^{D/2}\frac{dQ(x,t)dx}{dt}. \tag{4}$$

Here $q(t)$ is the net charge, defined as follows:

$$q(t)=\int_0^{D/2} Q(x,t)dx. \tag{5}$$

a)Electronic mail: tank@phys.uni-sofia.bg

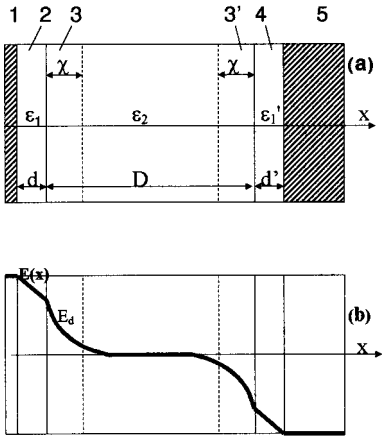


FIG. 1. (a) Construction of the electroacoustic cell. (b) Spatial distribution of the electric field in the cell.

II. ION CHARGE NONLINEAR DYNAMICS

First we shall find out the time dependence of the charge variations $q(t)$ driven by the external electric field. As seen from Eq. (5) the total charge on the electrode of the capacitor C_i equals the integral of the ion charges distribution in the region, where the electric field varies from maximum at $x=0$ to zero at $x=D/2$. The ion charge spatial distribution $Q(x,t)$ is defined by the sum of the concentration distributions of all types of charged particles in the solution:

$$Q(x,t)=\sum_{\alpha} Z_{\alpha} e \int_0^{D/2} n_{\alpha}(x,t)\,dx, \tag{6}$$

where Z_{α} is the valency of the ions of type α, e is the elementary electrical charge, and the ion concentration distribution obeys the Boltzmann law:[4]

$$n_{\alpha}(x,t)=n_{\alpha}\exp\left[-\frac{Z_{\alpha} e E(x,t)}{kT}\right]. \tag{7}$$

Here k is the Boltzmann constant and T is temperature.

The electron charges in the metal electrode, when in equilibrium, are equal to the sum of all ion charges in the solution as given in Eq. (5), but obviously the charge dynamics is limited by the much slower ionic charges. Thus, substituting (7) in (6) and (6) in (5) we obtain

$$q(t)=\sum_{\alpha} Z_{\alpha} e \int_0^{D/2} n_{\alpha}(x,t)\,dx. \tag{8}$$

The integral in (8) is difficult to approach directly, because of the complicated function $E(x,t)$ in the exponent of (7), but it can be essentially simplified if we take the time derivative:

$$\frac{dq(t)}{dt}=\sum_{\alpha} Z_{\alpha} e \int_0^{D/2} \frac{dn_{\alpha}(x,t)\,dx}{dt}. \tag{9}$$

Now we can use the continuity, one-dimensional equations, expressing that the gain in number of ions in a volume element can only be due to ions entering or leaving through its surface:[5]

$$\frac{dn_{\alpha}(x,t)}{dt}+\frac{d}{dx}[n_{\alpha}(x,t)v_{\alpha}(x,t)]=0. \tag{10}$$

Here v_{α} is the velocity of the ions of type α.

After substituting (10) in (9) we obtain

$$\frac{dq(t)}{dt}=-\sum_{\alpha} Z_{\alpha} e n_{\alpha}(x,t)v_{\alpha}(x,t)_0^{D/2}+K_1. \tag{11}$$

The integration constant K_1 will be specified later. The ion mobility μ_{α} is defined by

$$v_{\alpha}(x,t)=\mu_{\alpha}E(x,t). \tag{12}$$

Taking into account (12) and (2) we can simplify (11) as follows:

$$\frac{dq(t)}{dt}=E(0,t)\sum_{\alpha} Z_{\alpha} e \mu_{\alpha} n_{\alpha}(0,t)+K_1, \tag{13}$$

where $E(0,t)$ is the time dependence of the external electric field at the interface $x=0$:

$$E(0,t)=\frac{E_0}{\varepsilon_1}\sin\omega t. \tag{14}$$

Taking into account (14) and (7) we can integrate (13) over time to obtain $q(t)$:

$$q(t)=\frac{E_0}{\varepsilon_1}\sum_{\alpha} Z_{\alpha} e \mu_{\alpha}\int\left[\sin\omega t \times\exp\left(-\frac{Z_{\alpha} e E_0\sin\omega t}{\varepsilon_1 kT}\right)+K_1\right]dt+K_2. \tag{15}$$

The integration constant K_2 will be chosen to express the constant charge q_0, occurring when charged particles in the solution are adsorbed at the interface, i.e., $K_2=q_0$. It can be seen from (15) that the ion charge dynamics is defined by the nonlinear time dependence of the ion concentration at the interface membrane–electrolyte.

The argument of the exponential function in (15) is proportional to the ratio of the electric energy to the thermal energy of the ions and is usually much smaller than unity. This allows one to expand the exponent in (15) into series and take only the first few terms. The integration constant K_1 is chosen to eliminate the constant terms in the series expansion, thus preventing the appearance of terms growing linearly with time. The integration can be carried easily for the expansion series after transforming the powers of the corresponding trigonometric functions to trigonometric functions of a multiple argument, e.g.,

$$\sin^2(\omega t)=\frac{1}{2}-\frac{\cos(2\omega t)}{2},\qquad \sin^3(\omega t)=\frac{3\sin(\omega t)}{4}-\frac{\sin(3\omega t)}{4}. \tag{16}$$

The result of the integration, limited up to the third power of the exponential function is given in the following:

$$q(t)=q_0-\frac{F}{\omega}(q_1 e^{i\omega t}+iq_2 e^{2i\omega t}+q_3 e^{3i\omega t})+\text{conjugate}, \tag{17}$$

$$q_1=\frac{\mu_1-\mu_2}{2},\quad q_2=\frac{\alpha(\mu_1+\mu_2)}{8},$$
$$q_3=-\frac{\alpha^2(\mu_1-\mu_2)}{48},\quad F=\frac{E_0Zne}{2\varepsilon_1},\quad \alpha=\frac{ZeE_0}{\varepsilon_1 kT}\ll 1. \tag{18}$$

Here, for simplicity we have assumed a binary, symmetric electrolyte, i.e.,

$$Z_+=Z_-=Z,\quad n_+=n_-=n/2,\quad n=n_++n_-. \tag{19}$$

It should be noted that the obtained result can easily be generalized for asymmetrical electrolytes and for higher orders of accuracy, i.e., for higher acoustic harmonics in (17). From (17) and (18) it can be seen that the ion charge dynamics is nonlinear, whereas the nth-order harmonic amplitudes are proportional to α^{n-1}. The nonlinearity parameter α is proportional to the electric field amplitude E_0 and to ion valency Z, and is inversely proportional to temperature T.

III. NONLINEAR TIME DEPENDENCE OF THE LOCAL ION CONDUCTANCE

In the expression for the time derivative of the interelectrode gap $dX(t)/dt$ in Eq. (4) we have defined the charge time-dependence $q(t)$ [see Eq. (17)]. Next we shall evaluate the integral in the first term of the right-hand side of Eq. (4). We shall use once again the continuity equations (10) and relation (12) to obtain

$$\int_0^{D/2} x\frac{dQ(x,t)}{dt}dx=-\sum_\alpha Z_\alpha e\mu_\alpha\left[xn_\alpha(x,t)E(x,t)_0^{D/2}-\int_0^{D/2}n_\alpha(x,t)E(x,t)dx\right]. \tag{20}$$

Taking into account (2) we can ommit the first term on the right-hand side of (20):

$$\int_0^{D/2} x\frac{dQ(x,t)}{dt}dx=\sum_\alpha Z_\alpha e\mu_\alpha\int_0^{D/2}n_\alpha(x,t)E(x,t)dx. \tag{21}$$

Now, keeping in mind that $n_\alpha(x,t)$ is a non-negative function of x in the integration interval we can apply the mean-value theorem, i.e.,

$$\int_0^{D/2}E(x,t)n_\alpha(x,t)dx=E(x_0t)\int_0^{D/2}n_\alpha(x,t)dx, \tag{22}$$

where

$$E(x_0,t)=\frac{2}{D}\int_0^{D/2}E(x,t)dx. \tag{23}$$

To define the mean value of the electric field $E(x_0,t)$ we have to model the dynamics of the electric field spatial distribution $E(x,t)$ in the electrolyte. The following conditions must be satisfied.

(1) In the initial moment $t=0$ when the electric field is switched on, and the ions have not yet responded, the space distribution should be linear as in a dielectric:

$$E(x,0)=E(0,0)\frac{(D/2-x)}{D/2}. \tag{24}$$

Here $E(0,t)$ is defined by Eq. (14).

(2) At the moment when the external electric field amplitude is maximal, i.e., at $t=\pi/2\omega$, the distribution should achieve an exponential dependence, characteristic for an *edl* of width χ:

$$E\left(x,\frac{\pi}{2\omega}\right)=E\left(0,\frac{\pi}{2\omega}\right)\exp\left(-\frac{x}{\chi}\right). \tag{25}$$

(3) The mean-value electric field distribution should oscillate at a double frequency between the two extreme cases—a linear dependence and an exponential distribution.

These conditions can be satisfied by the following formula:

$$E(x,t)=\frac{E_0}{\varepsilon_1}\sin\omega t\left[\left(\frac{D/2-x}{D/2}\right)\cos\omega t+\exp\left(-\frac{x}{\chi}\right)\sin\omega t\right]. \tag{26}$$

Now, we substitute (26) in (23) and after integration over x we obtain

$$E(x_0,t)=\frac{E_0}{\varepsilon_1}\sin\omega t\left(\frac{\cos\omega t}{2}+\frac{2\chi\sin\omega t}{D}\right)=\frac{E_0}{\varepsilon_1}\left[\frac{\sin 2\omega t}{4}+\frac{\chi}{D}(1-\cos 2\omega t)\right]. \tag{27}$$

Substituting (27) in (21) we obtain

$$\int_0^{D/2} x\frac{dQ(x,t)}{dt}dx=\frac{E_0}{\varepsilon_1}\left[\frac{\sin 2\omega t}{4}+\frac{\chi}{D}(1-\cos 2\omega t)\right]\times\sum_\alpha Z_\alpha e\mu_\alpha\int_0^{D/2}n_\alpha(x,t)dx. \tag{28}$$

It can be noticed that the integral on the right-hand side of Eq. (28) is the same as the integral in Eq. (8), except for the extra multiplying term μ_α under the summation sign in Eq. (28). So we can write down

$$\int_0^{D/2} x\frac{dQ(x,t)}{dt}dx=\frac{E_0}{\varepsilon_1}\left[\frac{\sin 2\omega t}{4}+\frac{\chi}{D}(1-\cos 2\omega t)\right]K(t). \tag{29}$$

Here $K(t)$ has a conductance dimension and can be considered as the local conductance of the solution at the interface $x=0$, varying in time through the modulation of the local ion concentration. Obviously $K(t)$ can be defined, in the same approximation as $q(t)$, by the series

$$K(t)=k_0+\frac{F}{\omega}(k_1e^{i\omega t}+ik_2e^{2i\omega t}+k_3e^{3i\omega t})+\text{conjugate}, \tag{30}$$

where k_0 is the steady-state conductance at the interface, the coefficients k_1 are given in the following, while F and α are defined in Eq. (18),

$$k_1=\frac{\mu_1^2-\mu_2^2}{2},\quad k_2=\frac{\alpha(\mu_1^2+\mu_2^2)}{8},\quad k_3=-\frac{\alpha^2(\mu_1^2-\mu_2^2)}{48}. \tag{31}$$

We can now substitute Eq. (29) in (4) and obtain

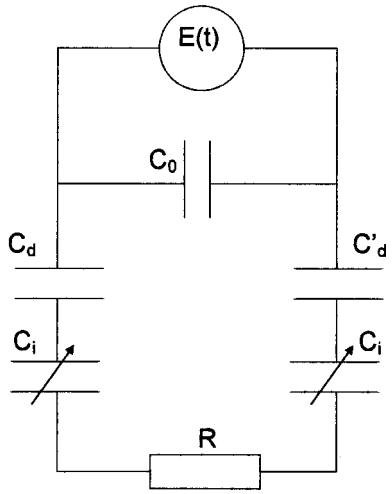


FIG. 2. Equivalent electric circuit of the electroacoustic cell.

$$\frac{dX(t)}{dt}=\frac{E_0}{\varepsilon_1}\left[\frac{\sin 2\omega t}{4}+\frac{\chi}{D}(1-\cos 2\omega t)\right]\frac{K(t)}{q(t)}-\frac{i(t)}{q(t)}X(t). \tag{32}$$

Note that here we have expressed the time derivative of the charges $dq(t)/dt$ by the electric current $i(t)$ in the circuit. This can be justified since the electron mobility in the wires of the circuit is much higher than the ion mobility in the solution. Thus, the time derivative of the electric charges is well approximated by its electron component $i(t)$, while the charge dynamics is limited by the slower ion charges $q(t)$. However, the time dependence of the current in the circuit $i(t)$ should be nonlinear as the ionic capacitor $C_i(t)$, driven by the external electric field, is time dependent.

IV. EVALUATION OF THE NONLINEAR ELECTRIC CURRENT IN THE CIRCUIT

To calculate $i(t)$ we shall follow the approach reported in Ref. 6, where an exact solution has been found for the nonlinear current in a circuit with time varying capacitance.

The equivalent electric circuit of the electroacoustic cell is presented in Fig. 2, where R is the resistance of the electrolyte, C_0 is the capacitance of the metal electrodes of the cell, $C_d=\varepsilon_1 S/d$ is the static capacitance of the dielectric membrane, $C_d'=\varepsilon_1 S/d'$ is the capacitance of the isolating layer d', and $C_i(t)=\varepsilon_2 S/X(t)$ is the time varying capacitance of the *edl* at the interface dielectric–electrolyte.

The equation governing the electric current in the circuit is the following:

$$\frac{dq(t)}{dt}+q(t)\left[\frac{1}{RC_d}+\frac{2}{RC_i(t)}\right]=\frac{E(t)}{R}. \tag{33}$$

Here C_0 is ignored in defining the effective capacitance of the circuit, because C_0 is much smaller than C_i and C_d. The capacitance C_d' is omitted as well, because it can be eliminated in the construction of the cell, or its value may be absorbed by C_d.

We want to evaluate the influence of the time-varying capacitance on the electron current in the circuit, so we consider the interelectrode gap oscillations between a minimum value χ and a maximum value $D/2$, driven by the electrosatatic force at a double frequency:

$$X(t)=\frac{D}{2}-\left(\frac{D}{2}-\chi\right)\sin 2\omega t. \tag{34}$$

Preserving the notations from Ref. 6 we can write the general solution of the differential equation (33) as follows:

$$q(\varphi)=(y-2b\sin\varphi)e^{-y\varphi-b\cos 2\varphi}\int \sin\varphi e^{y\varphi+b\cos 2\varphi}d\varphi, \tag{35}$$

where $\varphi=\omega t$ is a dimensionless time, y is the reciprocal of a dimensionless frequency given by

$$y=\frac{1}{\omega RC_d}+\frac{D}{\omega R\varepsilon_2 S}, \tag{36}$$

and b is defined as follows:

$$b=\frac{(D/2)-\chi}{\omega R\varepsilon_2 S}. \tag{37}$$

Substituting (35) in (33) we can obtain the following expression for the current:

$$\frac{i(\varphi)}{i_0}=\sin\varphi-(y-2b\sin 2\varphi)e^{-y\varphi-b\cos 2\varphi}\times\int \sin\varphi e^{y\varphi+b\cos 2\varphi}d\varphi, \tag{38}$$

where $i_0=E_0/R$.

Since our purpose is to find out the Fourier harmonics in the steady-state part of (38), the indefinite integral is transformed to a definite integral with the help of the technique described in Appendix II of Ref. 6,

$$\frac{i(\varphi)}{i_0}=\sin\varphi-\alpha(y-2b\sin 2\varphi)e^{-y\varphi-b\cos 2\varphi}\times\int_0^{2\pi}\sin(\mu+\varphi)e^{y(\mu+\varphi)+b\cos 2(\mu+\varphi)}d\mu, \tag{39}$$

where

$$a=(e^{2x\varphi}-1)^{-1}. \tag{40}$$

Next, we want to expand $i(\varphi)/i_0$ in the complex Fourier series:

$$\frac{i(\varphi)}{i_0}=\sum_{n=-\infty}^{\infty}C_n\exp(in\varphi), \tag{41}$$

where

$$C_n=\frac{1}{2\pi}\int_0^{2\pi}\frac{i(\varphi)}{i_0}\exp(-in\varphi)d\varphi. \tag{42}$$

Substituting (39) in (42), we find that

$$f_0 - C_n = \frac{a}{2\pi}\int_0^{2\pi}(y - 2b\sin 2\varphi)e^{-(in+y)\varphi - b\cos 2\varphi}d\varphi \times \int_0^{2\pi}\sin(\mu+\varphi)e^{y(\mu+\varphi)+b\cos 2(\mu+\varphi)}d\mu, \tag{43}$$

where

$$f_0 = \frac{1}{2\pi}\int_0^{2\pi}\sin(\varphi)e^{-in\varphi}d\varphi. \tag{44}$$

Following Ref. 6 we introduce two auxilary functions:

$$g(\varphi) = e^{-(in+y)\varphi - b\cos 2\varphi}, \qquad G(\varphi) = \int_0^{2\pi}\sin(\mu+\varphi)e^{y(\mu+\varphi)+b\cos 2(\mu+\varphi)}d\mu. \tag{45}$$

Now, the integral relative to φ in (43) can be expressed through $dg(\varphi)/d\varphi$ and integrated by parts as follows:

$$f_0 - C_n = -\frac{a}{2\pi}g(\varphi)G(\varphi)_0^{2\pi} + \frac{a}{2\pi}\int_0^{2\pi}g(\varphi)\frac{dG(\varphi)}{d\varphi}d\varphi - \frac{ina}{2\pi}\int_0^{2\pi}g(\varphi)G(\varphi)d\varphi. \tag{46}$$

It can be shown that the first term on the right-hand side of Eq. (46) equals zero, the second term is f_0, and (46) can be simplified as follows:

$$C_n = \frac{ina}{2\pi}\int_0^{2\pi}g(\varphi)G(\varphi)d\varphi. \tag{47}$$

The further calculations of the complex coefficients C_n are carried out in the Appendix and the final result is

$$C_n = \begin{cases} 0 & \text{for } n = 0,2,4,6,... \\ -\dfrac{n\pi e^{in\pi/4}}{2(1+e^{-y\pi})}J_{(n-iy)/2}(ib)[e^{i3\pi/4}J_{(-iy+1)/2}(ib) - e^{-i3\pi/4}J_{(-iy-1)/2}(ib)] & \text{for } n = 1,3,5,7,.... \end{cases} \tag{48}$$

Here $J_\nu(z)$ are the Bessel functions of order ν and complex argument z.

It can be seen that all even harmonics of the electric current are zero and only odd harmonics are present. This result reflects the symmetry of the two edls, adjacent to the electrodes.

V. RESULTS, NUMERICAL EVALUATIONS, AND DISCUSSION

Finally, we have defined all quantities in the relation for the derivative $dX(t)/dt$ in Eq. (32). The time dependence of the interelectrode gap $X(t)$ can be approximated by its second-harmonic component as defined in Eq. (34). After substitution of (34) in (32) and differentiating we can obtain the second derivative $d^2X(t)/dt^2$. Then we substitute the first and second derivatives in (3) to obtain the following expression for the time dependence of the membrane displacement $\xi(t)$:

$$l\xi(t)q^2(t) = \frac{E_0}{\varepsilon_1}\left[\frac{\sin 2\omega t}{4} + \frac{\chi}{D}(1-\cos 2\omega t)\right] \times \left[rq(t)K(t) + \rho q(t)\frac{dK(t)}{dt} - \rho K(t)i(t)\right] + i(t)D\left\{2\omega\left(0.5 - \frac{\chi}{D}\right)q(t)\cos 2\omega t + \left[0.5 - \left(0.5 - \frac{\chi}{D}\right)\sin 2\omega t\right][\rho i(t) - rq(t)] + \frac{\rho E_0\omega}{\varepsilon_1}K(t)q(t)\left(\frac{\cos 2\omega t}{2} + \frac{2\chi}{D}\sin 2\omega t\right)\right\}. \tag{49}$$

To evaluate the acoustical harmonics amplitudes we shall expand all quantities in Eq. (49) in series of exponent powers. The time dependence of the ionic charge $q(t)$ is given by Eq. (17) and the dynamics of the solution conductivity $K(t)$ at the interface is given by Eq. (30). The electric current $i(t)$ is given by the series

$$i(t)/i_0 = C_1e^{i\omega t} + C_3e^{i3\omega t} + C_5e^{i5\omega t} + \cdots + \text{conjugate}, \tag{50}$$

where the complex coefficients C_n are given by (48).

The time dependence of the membrane displacement $\xi(t)$ can be expanded in an analogous way, as follows:

$$\xi(t) = \xi_0 + \xi_1e^{i\omega t} + \xi_2e^{i2\omega t} + \xi_3e^{i3\omega t} + \cdots + \text{conjugate}, \tag{51}$$

where ξ_0, ξ_1, ξ_2, ξ_3 are unknown complex coefficients to be defined from Eq. (49).

We can truncate all infinite series $q(t)$, $K(t)$, $i(t)$, and $\xi(t)$ above a given term number and by substituting in (49) we shall define the amplitudes of the first few harmonics. For this purpose we make equal the coefficients standing before the exponent terms of the same power and we obtain a linear system of equations of order defined by the chosen accuracy of approximation.

For simplicity we have evaluated numerically only the first three harmonics. The amplitudes S_1, S_2, and S_3 of the first three harmonics are shown in Fig. 3 as a function of the dimensionless frequency y^{-1}. Although some experimental

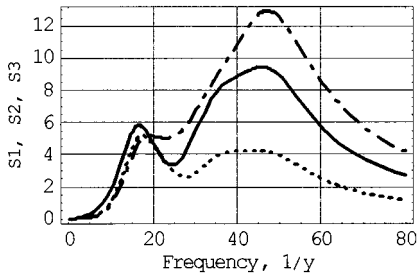


FIG. 3. Dependence of the acoustic harmonics amplitudes S_1, S_2, and S_3 of the dimensionless frequency y^{-1}. S_1 is denoted by a dashed line, S_2 is presented by a continuous line, and the dotted line corresponds to S_3.

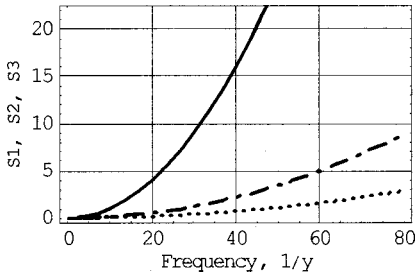


FIG. 4. Limiting case when no charges are adsorbed at the interface ($q_0=0$). S_1—dashed line, S_2—continuous line, and S_3—dotted line.

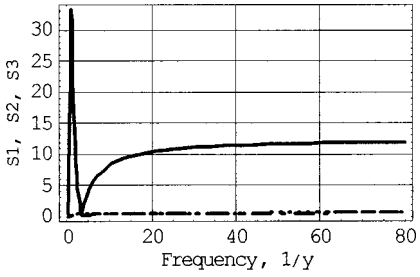


FIG. 5. Limiting case when the mobilities of the two types of ions are equal. S_1—dashed line, S_2—continuous line, and S_3—dotted line.

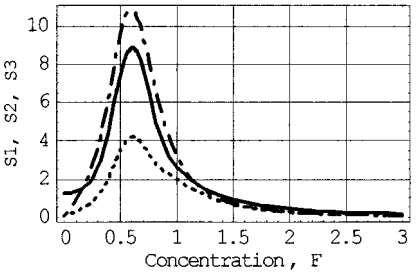


FIG. 6. Dependence of the acoustic harmonics amplitudes S_1, S_2, and S_3 of the ion concentration F. S_1 corresponds to a dashed line, S_2 corresponds to a continuous line, and S_3 corresponds to a dotted line.

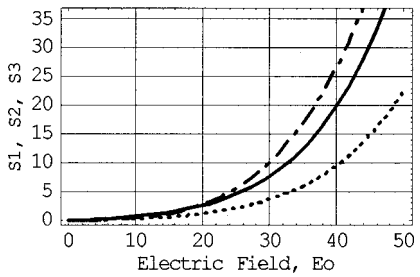


FIG. 7. Dependence of the acoustic harmonics amplitudes S_1, S_2, and S_3 of the electric field E_0 applied at the electrodes of the cell. S_1—dashed line, S_2—continuous line, and S_3—dotted line.

spectra have been reported[7] a detailed comparison between the experimental curves and Fig. 3 should be avoided because the experimental results are influenced by many factors: acoustic wave diffraction and reflection effects, frequency characteristics of the piezosensor and of the electric amplifier, etc.

Next, we can examine two limiting cases. The same spectra as in Fig. 3, but when no charges are adsorbed at the interface ($q_0=0$), are presented in Fig. 4. It can be noticed that S_2 is dominant and grows steeper with frequency than S_1 and S_3. In Fig. 5 spectra are presented in the case when the mobilities of the two types of ions are equal ($\mu_1=\mu_2$). In this case the second harmonic is dominant, as well, with a sharp maximum in the low-frequency range. However it should be noted that both limiting cases are theoretical idealizations and cannot be achieved in a real system. Neither charge adsorption can be completely avoided at the interface, nor the mobilities of two different types of ions can be perfectly equal.

The dependence of the three harmonics amplitudes of the parameter F are presented in Fig. 6. The parameter F as seen from Eq. (18) is proportional to the ion concentration so varying F and keeping all other parameters constant we obtain the ion concentration dependencies of the three harmonics. The concentration dependencies of the second acoustic harmonic have been examined experimentally[8] in different solutions. The curves in Fig. 6 agree qualitatively well with the experimental curves obtained in the last section of Ref. 8 for the case when the influence of the so-called ''memory effect,'' expressing the adsorption of ions from previous measurements in the micropores of the membrane, is avoided.

The nonlinear dependencies of the three acoustic harmonics amplitudes of the exciting electric field are shown in Fig. 7. The calculated curves agree with the experimental curves reported in Ref. 3, where polynomial expansion provides the best fitting with the experimental results.

Finally, we can conclude that the presented theoretical model successfully explains the nonlinear behavior of the electroacoustic electrolytic cell. The theoretical analysis provides a better insight into the phenomena and mechanisms of excitation of acoustic harmonics in electrolyte solutions. The electroacoustic cell can be used as a tool to study the properties of electrolyte solutions, as an ultrasound transducer or as a nonlinear signal processor. In the first case the theoretical analysis can be useful for better interpretation of the experimental results, obtained in complex solutions, while in the other cases the theoretical model may allow optimization of the cell parameters for a specific implementation.

APPENDIX: FOURIER COEFFICIENTS OF THE ELECTRIC CURRENT $i(t)$

On substitution of Eq. (45) in Eq. (47) the Fourier coefficients can be expressed as follows:

$$C_n=\frac{ina}{2\pi}\int_0^{2\pi} e^{y\mu+in\mu}d\mu\int_0^{2\pi}\sin(\mu+\varphi)$$
$$\times\exp-in(\mu+\varphi)-2b\sin\mu\sin(2\varphi+\mu)d\varphi. \qquad \text{(A1)}$$

Then we can make the substitution $\xi=\varphi+\mu/2$ and obtain the following result:

$$C_n=\frac{ina}{2\pi}\int_0^{2\pi}e^{y\mu+in\mu}d\mu\int_{\mu/2}^{\mu/2+2\pi}\sin\left(\xi+\frac{\mu}{2}\right)\times\exp\left[-in\left(\xi+\frac{\mu}{2}\right)-2k\sin\mu\sin 2\xi\right]d\xi. \quad \text{(A2)}$$

Since the second integral in (A2) is periodic, we may transform the boundaries in the range from 0 to 2π. After expanding the function $\sin(\xi+\mu/2)$ to exponents and ordering we obtain

$$C_n=\frac{na}{4\pi}(I_1-I_2), \quad \text{(A3)}$$

where

$$\begin{aligned} I_1&=\int_0^{2\pi}\exp\left[i\left(\frac{n+1}{2}iy\right)\mu\right]d\mu\\ &\quad\times\int_0^{2\pi}\exp[-i(n-1)\xi-2b\sin\mu\sin 2\xi]d\xi,\\ I_2&=\int_0^{2\pi}\exp\left[i\left(\frac{n-1}{2}iy\right)\mu\right]d\mu\\ &\quad\times\int_0^{2\pi}\exp[-i(n+1)\xi-2b\sin\mu\sin 2\xi]d\xi. \end{aligned} \quad \text{(A4)}$$

We shall describe the calculation of the integral I_1 in detail. The integral I_2 can be treated in an analogous way.

On making the substitution $\mu=\lambda+\pi$ we obtain for I_1:

$$I_1=2\pi\exp\left[i\left(\frac{n+1}{2}iy\right)\pi\right]\int_0^{2\pi}\exp[-i(n-1)\xi]\times J_{[(n+1)/2]iy}(2ib\sin 2\xi)d\xi. \quad \text{(A5)}$$

Here we have used the following presentation of the Bessel function:[9]

$$J_v(z)=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-iz\sin\theta+iv\theta}d\theta. \quad \text{(A6)}$$

Then we make the substitution $2\xi=\theta+\pi/2$ and obtain

$$I_1=\pi e^{y\pi}\exp\left[i\left(\frac{n+3}{4}\right)\pi\right]\int_{-\pi/2}^{7\pi/2}\exp\left[-i\left(\frac{n-1}{2}\right)\theta\right]\times J_{[(n+1)/2]iy}(2ib\cos\theta)d\theta. \quad \text{(A7)}$$

We present the integral in Eq. (A7) as a sum of four integrals by splitting the integration range as follows:

$$I_1=\pi e^{y\pi}\exp\left[i\left(\frac{n+3}{4}\right)\pi\right]\left[\int_{-\pi/2}^{\pi/2}f(\theta)d\theta+\int_{\pi/2}^{3\pi/2}f(\theta)d\theta+\int_{3\pi/2}^{5\pi/2}f(\theta)d\theta+\int_{5\pi/2}^{7\pi/2}f(\theta)d\theta\right]. \quad \text{(A8)}$$

Here $f(\theta)$ denotes the function under integral in Eq. (A7). We shall consequently calculate the four integrals.

The first integral can be transformed as follows:

$$\int_{-\pi/2}^{\pi/2}f(\theta)d\theta=2\int_0^{\pi/2}J_{[(n+1)/2]ty}\times(2ib\cos\theta)\cos\left[\frac{(n-1)\theta}{2}\right]d\theta. \quad \text{(A9)}$$

Using the relation[7]

$$\int_0^{\pi/2}J_{\nu+\mu}(2z\cos x)\cos[(\nu-\mu)x]dx=\frac{\pi}{2}J_\nu(z)J_\mu(z) \quad \text{(A10)}$$

we can obtain

$$\int_{-\pi/2}^{\pi/2}f(\theta)d\theta=\pi J_{(n-iy)/2}(ib)J_{(1-iy)/2}(ib). \quad \text{(A11)}$$

Next, we shall transform the boundaries of the second integral in (A8), in the proper range $(-\pi/2,\pi/2)$, by the substitution $\theta=\lambda+\pi$. After similar operations as with the first integral and using the property

$$J_v(-z)=e^{iv\pi}J_v(z), \quad \text{(A12)}$$

we can find

$$\int_{\pi/2}^{3\pi/2}f(\theta)d\theta=-\pi e^{y\pi}J_{(n-iy)/2}(ib)J_{(1-iy)/2}(ib). \quad \text{(A13)}$$

In the third integral we substitute $\theta=\lambda+2\pi$ and again using (A12) we obtain

$$\int_{3\pi/2}^{5\pi/2}f(\theta)d\theta=-\pi e^{-in\pi}J_{(n-iy)/2}(ib)J_{(1-iy)/2}(ib). \quad \text{(A14)}$$

In the fourth integral we substitute $\theta=\lambda+3\pi$ and find in an analogous way that

$$\int_{5\pi/2}^{7\pi/2}f(\theta)d\theta=\pi e^{y\pi}e^{-in\pi}J_{(n-iy)/2}(ib)J_{(1-iy)/2}(ib). \quad \text{(A15)}$$

After summation of the four components in Eq. (A8) we obtain for I_1:

$$I_1=\begin{cases}0 & \text{for } n=\text{even}\\ 2\pi^2e^{y\pi}(1-e^{y\pi})\exp\left[j\left(\frac{n+3}{4}\right)\pi\right] & \\ \quad\times J_{(n-iy)/2}(ib)J_{(1-iy)/2}(ib) & \text{for } n=\text{odd.}\end{cases} \quad \text{(A16)}$$

In absolutely the same way we calculate I_2:

$$I_2=\begin{cases}0 & \text{for } n=\text{even}\\ 2\pi^2e^{y\pi}(1-e^{y\pi})\exp\left[i\left(\frac{n-3}{4}\right)\pi\right] & \\ \quad\times J_{(n-iy)/2}(ib)J_{(-1-iy)/2}(ib) & \text{for } n=\text{odd.}\end{cases} \quad \text{(A17)}$$

Finally, substituting (A17) and (A16) into Eq. (A3) we obtain the result in Eq. (48).

[1] N. Tankovsky, ''Capacitive ultrasound transducer, based on the electrical double layer in electrolytes,'' J. Appl. Phys. **87**, 538–542 (2000).

[2] M. J. Anderson, J. A. Hill, C. M. Fortunko, N. S. Dogan, and R. D. Moore, ''Broadband electrostatic transducers: Modeling and experiments,'' J. Acoust. Soc. Am. **97**, 262–272 (1995).

[3] N. Tankovsky, ''Electroacoustic effect in electrolytes,'' J. Appl. Phys. **75**, 1239–1241 (1994).

[4] G. A. Martinov and R. R. Salem, *Lecture Notes in Chemistry* (Springer, Berlin, 1983), Vol. 33.
[5] P. Debye, ''A method for the determination of the mass of electrolyte ions,'' J. Chem. Phys. **1**, 13–16 (1933).
[6] J. R. MacDonald and D. E. Edmondson, ''Exact solution of a time-varying capacitance problem,'' Proc. IRE **49**(2), 453–466 (1961).
[7] N. Tankovsky and J. Burov, ''Ionic solution electroacoustic transducer,'' IEEE Ultrasonic Symposium, 1993, p. 435.
[8] N. Tankovsky and K. Baerner, ''Concentration-dependent electroacoustic spectra of some simple alkali-halide aqueous electrolytes,'' Ber. Bunsenges. Phys. Chem. **101**, 1480 (1997).
[9] I. S. Gradstein and I. M. Ryzhik, in *Tables of Integrals, Sums, Series and Products* (Moscow, State Publishing House for Physics and Mathematics, 1963).

Turning point filters: Analysis of sound propagation on a gyre-scale

Matthew Dzieciuch,[a] Peter Worcester, and Walter Munk
Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093

(Received 27 January 2001; revised 2 February 2001; accepted 26 March 2001)

Acoustic transmissions from Pioneer Seamount off California to a vertical array near Hawaii are analyzed using a technique which we call a *turning-point filter*. The observables, travel time and axial inclination, are interpreted in terms of the ocean sound–speed field. The method permits a uniform treatment of the arrival pattern, from the early ray-like arrivals to the late mode-like arrivals, including peak arrivals which cannot be identified as either rays or modes. An adiabatic range dependence is assumed, and Wentzel–Kramers–Brillouin–Jeffreys formalism is applied. [DOI: 10.1121/1.1377869]

PACS numbers: 43.30.Pc, 43.60.Rw [DLB]

I. INTRODUCTION

In July 1996, a 28 Hz signal was transmitted from a source near Pioneer Seamount off California to a vertical receiving array off Hawaii at a range of 3500 km. It was widely believed that the internal wave field in the ocean would lead to short vertical coherence lengths and thus prevent phase-coherent beamforming at very large ranges. At this low frequency however, the transmissions were strongly coherent across a broad vertical aperture, which came as a pleasant surprise. In this article we describe an attempt to take advantage of the favorable coherence environment.

One of the purposes of the experiment was to evaluate ultralow-frequency transmissions for ocean acoustic tomography. An account of tomographic methods is given by Munk, Worcester, and Wunsch (1995), to be referred to as MWW. The underlying principles are simple: the speed of sound increases with increasing temperature (by roughly 5 m/s per °C), so the travel time between a fixed source and receiver is predominantly a measure of the temperature of the intervening water. Sound in the ocean is trapped in the water column and travels along distinct ray paths. The resolution and identification of separate ray arrivals, from early steep rays which span the entire water column, to late flat rays which remain near the sound–speed axis, permit an inversion of the ray travel times $\tau(t)$ to the range-averaged sound–speed profile $C(z)$.

Most of the travel–time record is remarkably well represented by the ray archetype and thus tomography has relied nearly exclusively on the ray model. A large portion of the arrival energy is in the final crescendo though where the observed pattern of arrivals is more easily interpreted by a mode construction. The modes produce an interference pattern which lacks easily identifiable observable features (such as a travel time) with which to compare data and models. Furthermore the interference pattern is very sensitive to scattering induced by ocean internal gravity waves. The ineffectual use of the late arriving energy in an inverse solution for ocean properties is the motivatation for this article.

It has been customary to separate ray-based and mode-based analyses; here we introduce the *turning-point filter* in an attempt to find a unified procedure which makes effective use of all the received energy. The method can be viewed either as an extension to linear beamforming, by accounting for ray curvature, or as modal horizontal wave number filtering with a vertical array.

As mentioned above, the recent experiment near Pioneer Seamount at low frequency (28 Hz) revealed that scattering from internal waves was relatively small compared to previous higher frequency experiments at similar ranges (for example, see Worcester *et al.* (1999)]. There is no attempt to quantify the scattering here, instead the acoustic propagation is assumed to be coherent. Similarly, since the acoustic path crosses no strong fronts and does not interact with any bottom features the propagation is assumed to be adiabatic. Any extension of this work to higher frequency will need to consider the effects of increased scattering, such as biased arrival times or broadened late arriving energy, but these aspects are beyond the scope of this article.

The presentation is as follows: first the experimental setup is described. Then a heuristic derivation of the turning-point filter is given. A more formal derivation of the sound pressure field by a mode expansion follows. Experimental results detailing the application of the turning-point filter are shown. Finally an analysis of the variation of the observable quantities is presented which permits the application of inverse methods to deduce ocean properties.

Previous work has greatly influenced the ideas presented here. Guthrie and Tindle (1976) gave an early account of ray/mode equivalence. In their work, rays are represented as the sum of modes of the equivalent angle. Cox, Zeskind, and Myers (1990) analyzed a similar ocean acoustics problem, that of source localization. Their analysis was based on a ray angle-depth diagram, and they approximated the turning-point filter by a collection of linear beamformers. Since the turning-point filter can viewed as a generalization of the Cox *et al.* approach, there may be application to the source localization problem.

[a]Electronic mail: mad@ucsd.edu

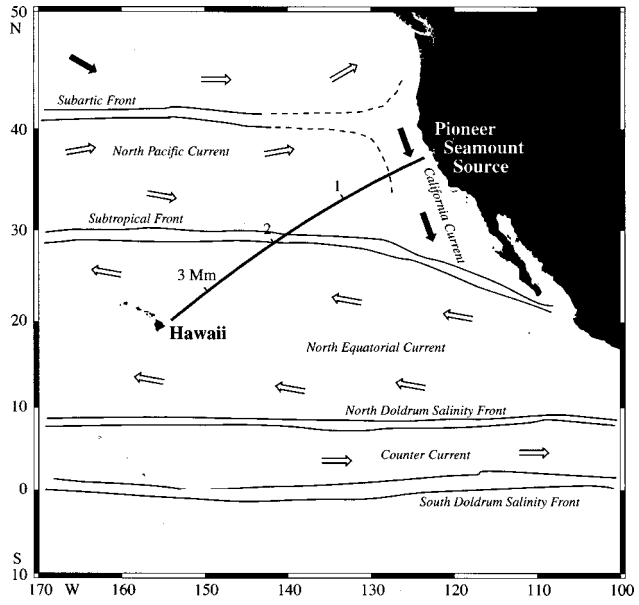


FIG. 1. Geodesic path from a source near Pioneer Seamount to a vertical receiving array near Hawaii, at a range of 3502 km. Range in megameters (1 Mm=1000 km) is indicated along the path. Mean flow and major ocean boundaries are adapted from the work by Roden (1975). Cold currents are indicated by solid arrows; warm currents by outlined arrows.

II. EXPERIMENT

The experiment took place in the eastern North Pacific Ocean (Fig. 1). A low-frequency, broadband acoustic source (HLF-6A) was deployed by the M/V INDEPENDENCE at 35° 17.508′ N, 123° 35.000′ W, 7 nm west–southwest of the Acoustic Thermometry of Ocean Climate (ATOC) source on Pioneer Seamount [see ATOC Consortium (1998)]. The source depth was 652 m.

The source transmitted a phase-modulated signal at a center frequency of 28 29/682 Hz. The phase modulation was derived from a linear maximal-length shift-register sequence containing 255 digits. The signal was modulated at a rate of three carrier cycles per sequence digit. Thus each digit was 106.980 ms long, and each sequence period was 27.28 s in duration. Forty transmissions of 20 min duration were separated by a minimum of 4 h. The source power level was 130 W (192 dB re 1 μPa at 1 m).

A moored vertical line array (VLA) near Hawaii (20° 10.592′ N,154° 00.917′ W) acted as a receiver. Source and receiver locations were determined with the global positioning system (GPS) in differential mode. The unrefracted geodesic range could then be determined to within a few meters. The range in WGS-84 coordinates to the Hawaii VLA was 3 501 582 m.

The Hawaiian VLA consisted of 20 hydrophones, spaced at 35 m, spanning the depths from 300 to 1000 m. The sound pressure signals were amplified, bandpass filtered, and sampled at 300 Hz by 16-bit analog-to-digital converters. The data quantity was reduced by forming four-period (27.28 s) block averages according to

$$p_{\text{average}}(n)=\sum_{m=0}^{3} p(n+mN), \tag{1}$$

where $N=f_sT$, $f_s=300\,\text{Hz}$, and $T=27.28\,\text{s}$.

Details of the geodesic slice from the source to the Hawaiian VLA are shown in Fig. 2. The eastern North Pacific Ocean is not as oceanographically complex as western boundary current areas. The ocean sound channel deepens from the source to the VLA.

Intensities at adjacent hydrophones of the VLAs are closely correlated, so that an intensity contour map in time, depth space can be drawn (Fig. 3). The lower panel shows the computed arrival pattern for the range-dependent environment in Fig. 2 using adiabatic mode theory (see Sec. IV C for details). The measured pattern lags the computed ray pattern by about 0.25 s. This suggests colder than climatological mean temperatures, but part of the difference between the measured and computed patterns is likely due to an error in the sound–speed equation at high pressures [Worcester *et al.* (1999)].

A series of clearly defined ridges sweep across the receiver array. The simplest interpretation is in terms of ray

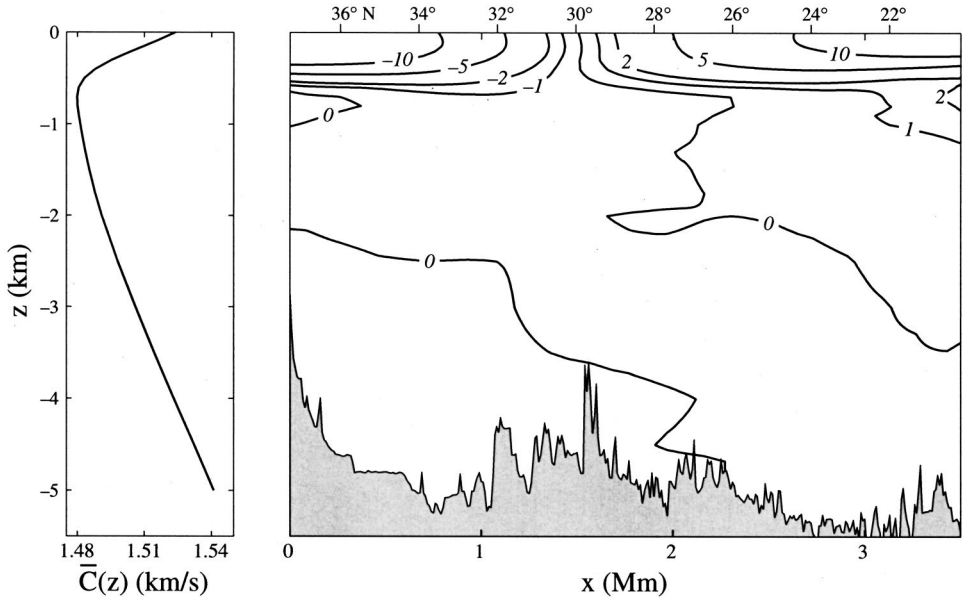


FIG. 2. July field of sound–speed anomaly $C(x,z)-\bar{C}(z)$ along the geodesic from Pioneer Seamount to Hawaii according to Levitus (1994) and to Levitus *et al.* (1994). The range-averaged sound–speed profile, $\bar{C}(z)$, is shown at the left. The bathymetry, shown in gray, is according to Smith and Sandwell (1997).

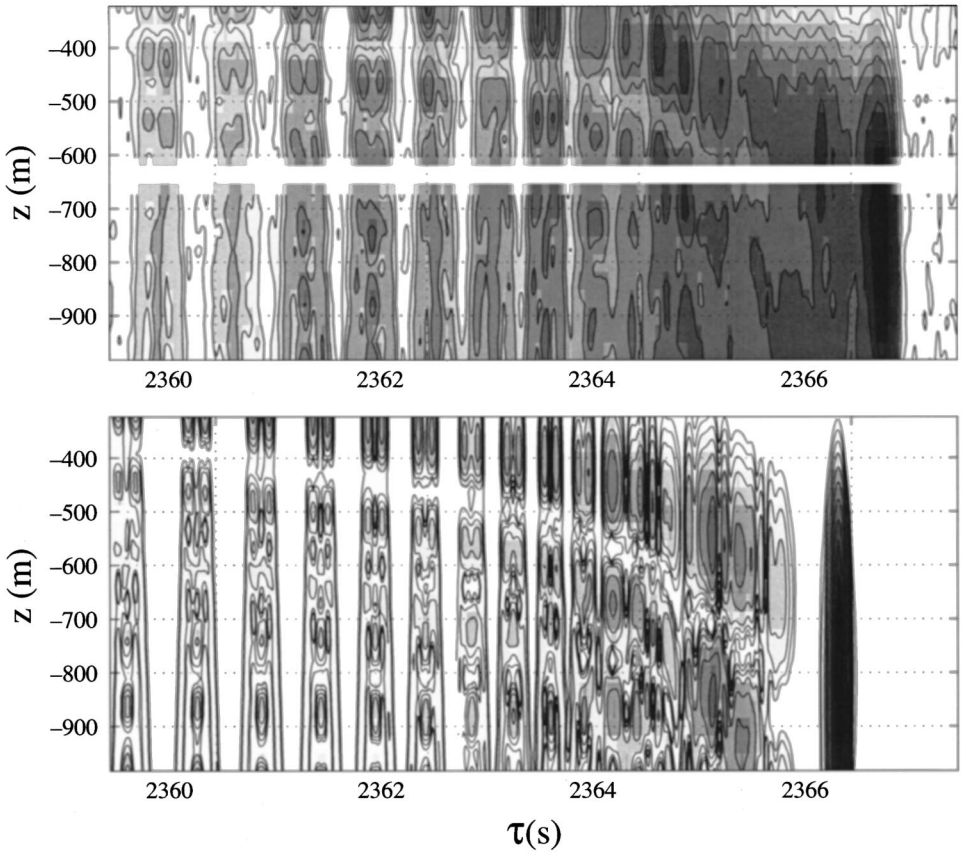


FIG. 3. Intensity of the Pioneer Seamount to Hawaii transmissions at 28 Hz, contoured at 3 dB intervals, with white designating an intensity null. The measured pattern (upper panel) is the average over 40 transmissions during 1996, yeardays 183–191. The hydrophone at −620 m did not work. The computed pattern (lower panel) is based on adiabatic normal mode theory.

theory. Each of the recorded ridges can be identified with a ray front and the observables, travel time and axial inclination, can be measured. For the later arrivals, the mode structure is demonstrated by the null in the measured reception. The null crosses 500 m depth at travel time $\tau = 2364$ s, and deepens with increasing τ, closely corresponding to the adiabatic mode prediction. This is a remarkably clean data set. There is very little of the scattering from internal waves that has been seen in data sets at higher frequency at this range. For example, compare the time front shown here with that of Colosi *et al.* (1999).

The question of how to extract useful observables from the later portion of the arrival pattern is the main focus of this article. This would provide important information about the vertical sound–speed structure near the sound–channel axis.

III. TURNING-POINT FILTER

Ocean acoustic tomography experiments have relied almost exclusively on the measurement of ray travel times, which are robust, stable, and identifiable features of the recorded data that may easily be compared to predictions derived from climatology. Furthermore, differences between measurements and predictions of ray travel times can be interpreted to first order in terms of the integrated sound–speed anomaly along the unperturbed ray path (MWW, pp. 52–56).

Not all of the received energy can be identified with ray arrivals (Fig. 3). Near the final crescendo, the data are better modeled by acoustic modes with modal group delays (or, equivalently, modal travel times), a potential observable. The experiment design can be difficult since the signals must simultaneously have a wide bandwidth to produce sharp ray travel–time peaks and a narrow bandwidth so that modal frequency dispersion is not a factor.

Here we describe an alternate approach based on the Wentzel–Kramers–Brillouin–Jeffreys (WKBJ) approximation for extracting observable from the final crescendo. We have called this approach a *turning-point filter*. The turning-point filter offers a unified approach to underwater acoustics, joining what are often separate procedures for ray and mode analysis. Implicit in our analysis is that the acoustic propagation is adiabatic and that internal wave scattering is much weaker than in higher frequency experiments [see Colosi and the ATOC Group (1999)] leading to strong vertical coherence.

As motivation for what follows, consider the vertical angle, depth diagram shown in Fig. 4. The diagram describes the *state* of rays that have propagated from a point source at a depth of 652 m to a range of 100 km through a range-independent ocean acoustic channel corresponding to $\overline{C}(z)$ in Fig. 2. At the initial range, all rays leaving a point sound source are located at the source depth but span all vertical angles (represented by the dashed horizontal line). As the rays propagate, the horizontal line twists itself into the spiral, shown by the heavy black line representing the state of the rays at 100 km. Any particular ray is represented by a point on the line, for illustration the −12, 0, +10, and +12 degree rays have been indicated. As a particular ray propagates, it traces an orbit in the angle, depth diagram. For example, the +10 degree ray traces the orbit shown by the thin line

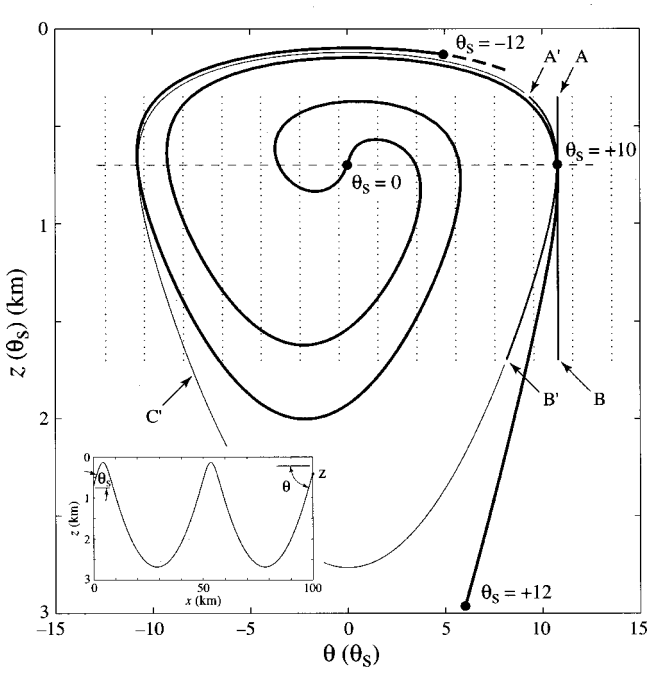


FIG. 4. Ray inclination, depth diagram corresponding to $\bar{C}(z)$ in Fig. 2. The heavy line shows $\theta(\theta_s), z(\theta_s)$ at a fixed range of 100 km, for a source depth of 652 m, as the source launch angle, θ_s, sweeps from $-12°$ to $+12°$. Multiple intersections at a fixed depth correspond to a different number of ray loops arriving at different times and inclinations. The thin line shows the inclination, depth orbit for a fixed source launch angle, $+10°$, as a function of range. A 40 element vertical array extending from -0.3 to -1.7 km is represented by columns of dots, horizontally spaced at the angular resolution of the array ($\lambda/L = 0.38$ rad$= 2.2°$). Linear beamforming corresponds to fitting the ray state by the vertical $A-B$ line as shown. Turning-point filtering uses the more accurate ray orbit approximation $A'-B'$.

$A'-B'-C'-A'$, as it propagates from the source to a receiver at the 100 km range.

A single hydrophone receives arrivals from all angles, corresponding to the multiple intersections of the ray state with any fixed depth. An array of hydrophones has a vertical extent and thus has angular resolution. Linear beamforming assumes a plane wave arrival for each ray path, and thus amounts to fitting the ray state by a vertical line in angle, depth space. In Fig. 4, linear beamforming by a vertical line array is represented by the heavy black vertical line $A-B$. The length of the line is the length of the array and the horizontal spacing between the parallel vertical dotted lines represents the array angular resolution. Thus for this particular example, most of the energy in the ray arrivals tangent to the line $A-B$ easily fits within the angular resolution. This is not the case for rays with small inclination.

A better approximation than linear beamforming accounts for wave front curvature across the array. We define the *turning-point filter* as using the ray orbit $A'-B'$ as an approximation to the ray state. The improvement over linear beamforming is especially pronounced for long arrays (narrow angular resolution) and small ray inclinations (strong curvatures). Our reason for calling this approximation a turning-point filter will become clearer in Sec. IV.

Matched field processing attempts to account for the *entire* measured ray state. The turning-point filter is an intermediate step between linear beamforming and the full matched field processor. The successive improvements are at the cost of increased requirements.

The mechanics of the turning-point filter are quite simple. It consists of beamforming with delays that allow for the curvature of the ray time front. Although we have been motivated by ray models, we interpret the turning point filter in mode language using the WKBJ approximation. In order to keep the discussion focused on essentials, the following is limited to a range-independent sound–speed profile.

Ocean sound waves produce a complicated spatial pattern that can be decomposed into the sum of horizontally propagating modes. A mode can be thought of as a (vertically) standing wave resulting from the interference pattern of up- and down-going waves,

$$u_m(z,\omega) = A_m(z,\omega)e^{j\phi_m(z,\omega)} + A_m(z,\omega)e^{-j\phi_m(z,\omega)}, \quad (2)$$

where m is the mode index number. Typically the amplitude A_m is a slowly varying function of depth, whereas the phase ϕ_m varies rapidly with depth. In the WKBJ approximation, A_m is taken as a constant between depths, $\tilde{z}^- < z < \tilde{z}^+$ (called turning-point depths), and zero outside of those depths.

Many modes m can have the same turning depths provided they have the appropriate frequency ω_m. Modes with the same turning depths ensonify the same ocean layer $\tilde{z}^- < z < \tilde{z}^+$, and thus, as shown later, sample the ocean in a similar manner [see Brekhovskikh and Lysanov (1991)]. They can be combined to give a robust observable, hence the name *turning-point filter*. A vertical array is required to accomplish this.

Let $C(z)$ designate the sound speed, and $S(z) = 1/C(z)$ the sound slowness. The up- and down-going waves in Eq. (2) have a wave number $k(z) = \omega S(z)$ tilted in a direction $\theta(z)$ with respect to the horizontal. The horizontal projection of the wave number is then

$$k_H(z) = k(z)\cos\theta(z) = k_0 \cos\theta_0, \quad (3)$$

with the right-hand side following from Snell's law. Here $k_0 = \omega S_0$ is a reference wave number (usually taken at the sound axis where the slowness is a maximum), so θ_0 is the axial tilt of the modal wave number.

It is convenient to define a local vertical wave number,

$$k_V(z) = \sqrt{k^2(z) - k_H^2}, \quad = \omega\sqrt{S^2(z) - S_0^2\cos^2\theta_0}. \quad (4)$$

The vertical wave number changes from a real to an imaginary value at depths $\tilde{z}$ (the turning-point depths) defined by

$$S(\tilde{z}) = S_0 \cos\theta_0. \quad (5)$$

The phase difference between a hydrophone at depth z_i and depth z_0,

$$\phi_i = \int_{z_0}^{z_i} k_V(z)\,dz, \quad (6)$$

is associated with a time delay

$$\Delta\tau_i = \frac{\phi_i}{\omega} = \int_{z_0}^{z_i} \sqrt{S^2(z) - S_0^2\cos^2\theta_0}\,dz. \quad (7)$$

For any chosen axial tilt θ_0, (7) defines a time-delay beamformer for hydrophones between depths $\tilde{z}^-$ and $\tilde{z}^+$. An important consideration is that the time delays are not a func-

tion of frequency ω. So adding the delayed hydrophone outputs selects *all* modes m and frequencies ω_m associated with turning points at $\tilde{z}^{\pm}$. There is no need for mode separation and frequency filtering. For a constant sound slowness $S=S_0$, the previous expression reduces to a linear beamformer with

$$\Delta\tau_i=(z_i-z_0)S_0\sin\theta_0. \tag{8}$$

We represent the source signal by

$$s(t)=a(t)e^{i\omega_c t}, \tag{9}$$

where ω_c is the carrier frequency. [Here we have ignored all *m*-sequence processing and demodulations; see Metzger (1983) or MWW, pp. 183–197 for details.] The signal received at time t at the hydrophone i at depth z_i can be written

$$p_i(t)=\beta_i(t)s(t-\tau_i), \tag{10}$$

where $\beta_i(t)$ allows for the propagation loss and random phase changes, and

$$\tau_i=\tau_0+\Delta\tau_i \tag{11}$$

is the travel time to hydrophone i, with τ_0 designating the travel time to the axial hydrophone at z_0. A beamformed signal is obtained by summing the recorded pressures each lagged by an amount $\Delta\tau_i$:

$$b(t)=\sum_i p_i(t+\Delta\tau_i). \tag{12}$$

IV. MODE ANALYSIS

In Sec. III we made the assertion that modes with the same turning-point depths sample the ocean in the same way. It is necessary to put this statement on firmer ground. Although Sec. IV is long and mathematically tedious, the reward is a set of formulas relating observable acoustic quantities such as travel time and inclination to ocean parameters.

We omit all features not crucial to the present application. The notation follows MWW. The wave equation,

$$\left(\nabla^2-S^2(z)\frac{\partial^2}{\partial t^2}\right)p=0, \tag{13}$$

with sound slowness $S(z)$, has a separable solution for the pressure,

$$p(x,z,t)=Q(x)P(z)e^{i\omega t}. \tag{14}$$

At long ranges

$$Q(x)=\sqrt{\frac{2}{\pi k_H x}}e^{-i(k_H x-\pi/4)}. \tag{15}$$

The vertical wave function $P(z)$ must satisfy

$$\frac{d^2P}{dz^2}+[\omega^2S^2(z)-k_H^2]P=0. \tag{16}$$

Equation (16) together with boundary conditions then determines the values of the integration constant $k_H(m,f)$. We recognize $\omega S(z)=k(z)$ as the scalar wave number, and the separation constant k_H as a horizontal wave number, so that

$$k_V(z)=[\omega^2S^2(z)-k_H^2]^{1/2} \tag{17}$$

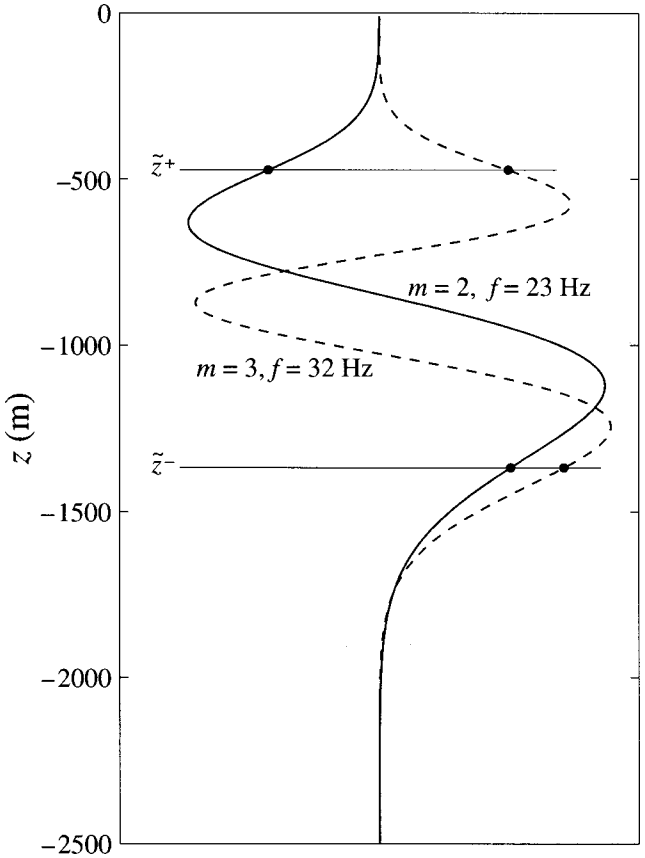


FIG. 5. Wave functions $P_m(z,f)$ for mode $m=2$ (2 extrema) at $f=23$ Hz and mode 3 (3 extrema) at 32 Hz. The frequencies have been so chosen that the two modes have the same upper ($\tilde{z}^+=-460$ m) and lower ($\tilde{z}^-=-1430$ m) turning points. Modes with equal turning points interfere constructively to form ray-like arrivals.

is the *local* vertical wave number. As the slowness varies with depth, $k_V(z;m,f)$ changes from a real to an imaginary value at depths $\tilde{z}$, called the *turning depths*, where the phase slowness is given by

$$s_p\equiv k_H/\omega=S(\tilde{z})\equiv\tilde{S}. \tag{18}$$

An important modal parameter (corresponding to the axial ray inclination) is the inclination, θ_0, of the wave number at the axis:

$$\tan[\theta_0(m,f)]=\frac{k_V(z_0;m,f)}{k_H(m,f)}. \tag{19}$$

We consider only ''trapped'' modes, with turning points well beneath the surface and above the bottom. Solutions $P_m(z;f)$ to (16) are the wave functions at frequency $f=\omega/2\pi$ for modes $m=1,2,...$, with m extrema between the turning depths. Figure 5 shows $P_2(z;23\,\text{Hz})$ and $P_3(z;32\,\text{Hz})$, which have identical turning depths at 460 and 1430 m. In general, modes with the same value of $(m-1/2)/f$ have the same turning-point depths (MWW, p. 64). For large m, the many overlapping modes interfere constructively to form rays with ray turning-points at the modal turning-points.

A. Dispersion

It can be shown (MWW, Eq. 2.10.8) that the group slowness is given by

$$s_g(m,f)=dk/d\omega=\overline{S^2}/s_p, \tag{20}$$

where

$$\overline{S^2}=\frac{\int_{-\infty}^{\infty}dz S^2(z)P^2(z)}{\int_{-\infty}^{\infty}dz P^2(z)} \tag{21}$$

is a mode-weighted sound slowness. The WKBJ approximation [e.g., see Brekhovskikh and Lysanov (1991), Sec. 6.7] leads to an important simplification. The group slowness is still given by Eq. (20), but with a new formula for $\overline{S^2}$:

$$\overline{S^2}=\frac{\int_{\tilde{z}^-}^{\tilde{z}^+}dz S^2(S^2-\tilde{S}^2)^{-1/2}}{\int_{\tilde{z}^-}^{\tilde{z}^+}dz(S^2-\tilde{S}^2)^{-1/2}}. \tag{22}$$

The result is

$$s_g(\theta_0)=\tilde{S}+\frac{\int_{\tilde{z}^-}^{\tilde{z}^+}dz(S^2-\tilde{S}^2)^{1/2}}{\tilde{S}\int_{\tilde{z}^-}^{\tilde{z}^+}dz(S^2-\tilde{S}^2)^{-1/2}}, \tag{23}$$

$$\tilde{S}=S_0\cos\theta_0. \tag{24}$$

With $S(z)$ given and θ_0 specified, Eq. (24) determines $\tilde{S}=S(\tilde{z})$ and hence $\tilde{z}^{\pm}$, and the integrations can be performed. Thus the travel time, $\tau=s_g x$, depends only on a single parameter $\theta_0(m,f)$, rather than on m and f separately. This implies that modes with the same turning points will have a travel time that is dependent on the same function of sound slowness.

We now go back to the ''exact'' solution (20). For a range-independent environment, the travel time at range x is given by $\tau=s_g x$. In an adiabatically range-dependent environment, the sound–speed profile is assumed to vary slowly with range, i.e., the energy within each mode is conserved. The group slowness is then interpreted locally, and the travel time is given by

$$\tau_m(x,f)=\int_0^x s_g(x';m,f)dx'. \tag{25}$$

The relation holds for the modal case $s_g(x;m,f)$ as given by Eq. (20), or in the WKBJ approximation $s_g(x;\theta_0)$ as given by Eq. (23).

In the adiabatic approximation the energy within each mode is conserved (see Brekhovskikh and Lysanov, (1991), p. 132 for the WKBJ assumptions). In fact, propagation across ocean fronts and over bottom ridges leads to very significant mode-to-mode scattering and coupling as demonstrated by McDonald *et al.* (1994) and by Shang *et al.* (1994) (see also MWW, p. 335, Figs. 8.10–8.13). The transmission path was chosen to avoid sharp frontal and bathymetric features (Figs. 1 and 2); still the neglect of range-dependent scattering especially from internal waves [see Colosi *et al.* (1994)] is a serious shortcoming of our analysis.

Figure 6 shows the dispersion relations computed according to Eq. (25) for the transmission to Hawaii (Fig. 1), with $C(x,z)$ shown in Fig. 2. It will be shown that computed and observed arrival patterns are in excellent agreement, and we consider the measured dispersion conditions to be fairly represented by Fig. 6. The panels show the ''exact'' modal solutions $\tau_m(f)$ for given mode numbers m and frequencies f. In the upper panel, the group velocity c_g and associated travel time τ_m is plotted over the transmitted bandwidth of the 28 Hz signal, $f=18-38$ Hz, for modes $m=1,2,...,30$. In the middle panel, the corresponding dispersion curves are shown as a function of the axial ray inclination $\theta_0=\theta(z_0)$ at the receiver. In the lower panel, the curves are shown in τ_m,θ_0 space. The important point is that all nonsurface interacting modes (except $m=1$) and all frequencies coalesce onto a single curve $\tau(\theta_0)$ in accordance with WKBJ theory to within 25 ms.

There is no point extracting separate modes from the array records. All combinations of modes and frequencies having the same value of $\theta_0(m,f)$ also have the same values of

$$\tilde{S},\ \tilde{z}^{\pm},\ s_g,\ s_p,\ \tau,\ A, \tag{26}$$

and sample the ocean column in a similar manner (the action, A, will be defined later). Although the main goal of the turning-point filter is to obtain robust observables for the entire arrival pattern, another aim can be fulfilled: gaining statistical precision by combining all hydrophone records associated with a fixed θ_0, or equivalently with fixed turning-point depths $\tilde{z}^{\pm}(\theta_0)$, or fixed turning-point slownesses $\tilde{S}^{\pm}(\theta_0)$. For the early, steep arrivals many modes interact and this ''constructive interference'' is equivalent to the ray beamforming previously discussed. For late, near-axial arrivals only a few modes interact constructively (see Fig. 7), and the ray representation is no longer useful. Still there is a moderate statistical gain from combining late modal arrivals, and there is a further advantage in a uniform procedure for interpreting the early ray-like and late mode-like arrivals.

B. Ray/mode duality

In order to derive expressions for observable quantities and to perform the perturbation analysis in Sec. VI, it is necessary to endure a bit of algebra concerning the duality of ray and mode expressions. By observable quantities, we mean travel time and inclination for any of three types of observations: rays, modes, and turning-point filters. The method will be to use stationary phase arguments to find the locations of peaks in travel-time, inclination space.

The complete solution to Eq. (13) can be written as a mode superposition integrated over the source bandwidth,

$$p(x,z,t)=\int d\omega\sum_m p_m(x,z,\omega)e^{i\omega t}. \tag{27}$$

Each mode is excited in proportion to the value $P_m(z_S)$ of the mode function at the source depth; the receiver amplitude is proportional to $P_m(z_R)$, so (writing x,z for x_R,z_R)

$$p_m(x,z,\omega)=\sqrt{\frac{2}{\pi k_H x}}P_m(z_S)P_m(z)e^{-ik_H x}. \tag{28}$$

We are concerned with solutions that are trapped in an interior sound channel, i.e., solutions that decay exponen-

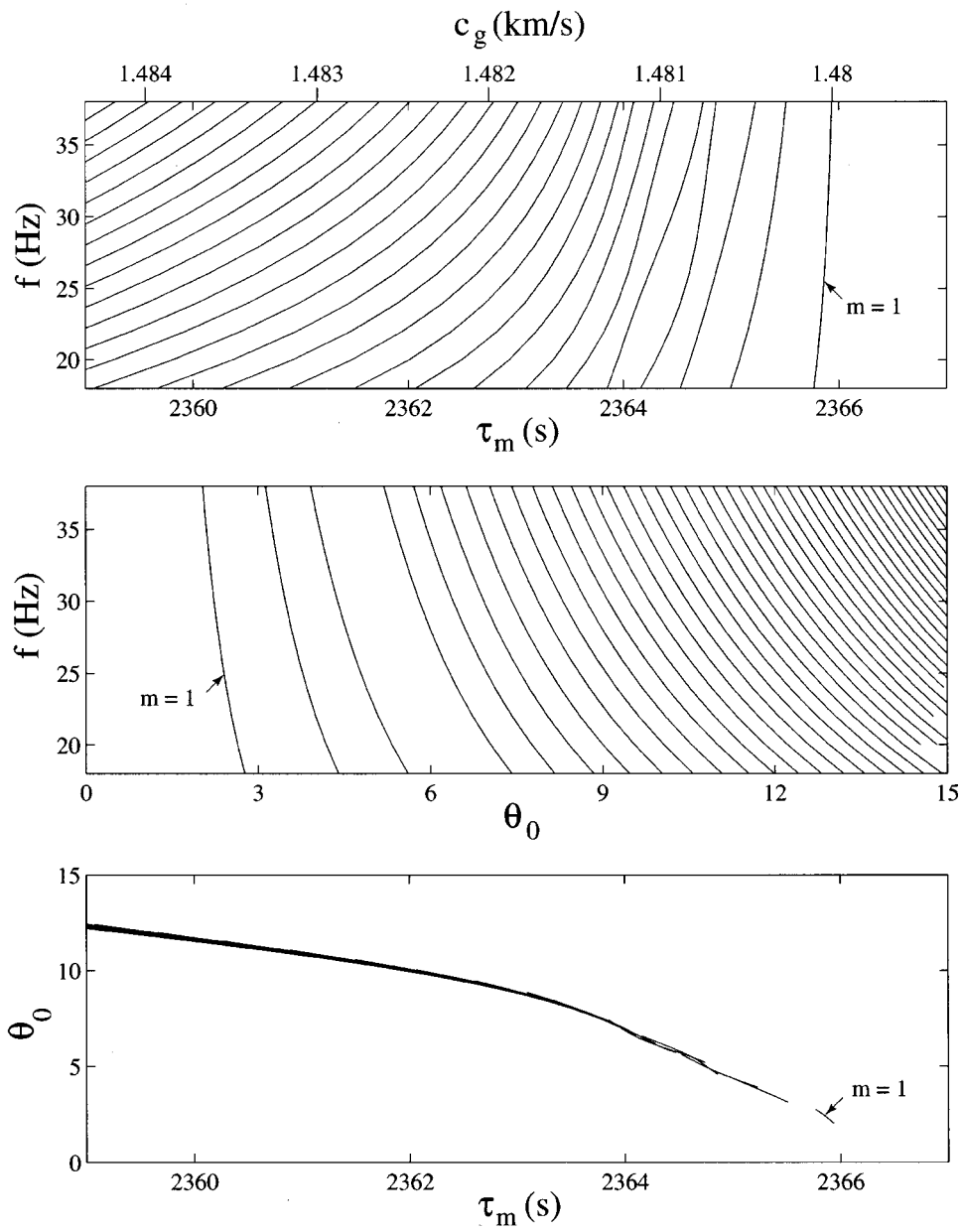


FIG. 6. Computed modal dispersion curves at a range of 3200 km. Upper panel: The group velocity c_g and associated travel time τ_m vs frequency. Middle panel: Corresponding dispersion curves as a function of the axial ray inclination θ_0 at the receiver. Bottom panel: Modal dispersion in τ_m, θ_0 space.

tially beneath the lower turning depth $\tilde{z}^-$ and above $\tilde{z}^+$, and are oscillatory for $\tilde{z}^- < z < \tilde{z}^+$. Using WKBJ theory, solutions can be written as

$$P_m(z) \sim \cos[\psi_m(z) - \tfrac{1}{4}\pi], \quad \psi_m(z) = \int_{\tilde{z}^-}^{z} dz' k_v(z'). \tag{29}$$

The $\frac{1}{4}\pi$ term is required to match solutions and derivatives at the turning points. The phase integral $\psi_m(z)$, written previously as (16), measures the phase changes from $\tilde{z}^-$ upward. The total phase change between the two turning points,

$$\psi_m(\tilde{z}^+) = \omega \int_{\tilde{z}^-}^{\tilde{z}^+} dz [S^2(z) - \tilde{S}^2]^{1/2} \equiv \tfrac{1}{2}\omega A, \tag{30}$$

defines the ''action'' A. It can be shown [Brekhovskikh (1980), Sec. 24] that the patching of solutions at the turning points requires

$$\psi_m(\tilde{z}^+) = \pi(m - \tfrac{1}{2}), \quad m = 1,2,\ldots. \tag{31}$$

The mode number m equals the number of extrema in $P_m(z)$.

Using (29) we can write the triple product in (28) as the sum of four terms

$$p_m = \sqrt{\frac{2}{\pi k_H x}} (e^{i\psi_a} + e^{i\psi_b} + e^{i\psi_c} + e^{i\psi_d}), \tag{32}$$

with

$$\psi_a = k_H x + \psi_S - \psi_R, \quad \psi_c = k_H x + \psi_S + \psi_R - \frac{\pi}{2},$$
$$\psi_b = k_H x - \psi_S + \psi_R, \quad \psi_d = k_H x - \psi_S - \psi_R + \frac{\pi}{2}. \tag{33}$$

These represent a group of four ray arrivals, as we shall see, with

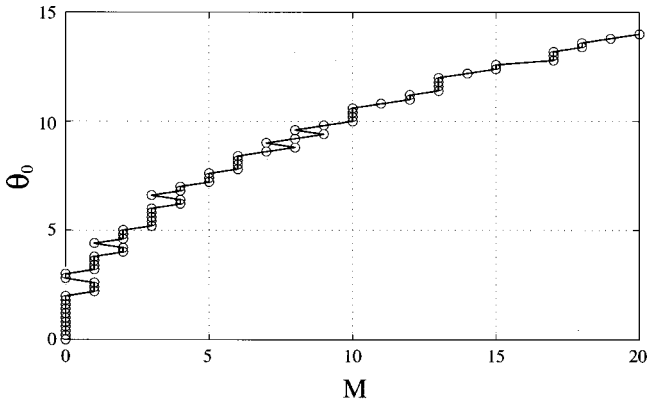


FIG. 7. Number of modes, M, summed to form the arrival pattern as a function of arrival axial inclination. Note that at low angles there are no modes available in the sum.

$$p_a(x,z,t)=\int d\omega \sum_m \sqrt{\frac{2}{\pi\omega\tilde{S}x}}\, e^{i\omega(t-\tau_x-\tau_S+\tau_R)} \tag{34}$$

and similarly for p_b, p_c, and p_d. In the above, we have used $k_H=\omega\tilde{S}$, and the delays

$$\tau_x=\tilde{S}_x, \tag{35a}$$

$$\tau_S=\int_{\tilde{z}^-}^{z_S} dz'(S^2-\tilde{S}^2)^{1/2}, \tag{35b}$$

$$\tau_R=\int_{\tilde{z}^-}^{z} dz'(S^2-\tilde{S}^2)^{1/2}. \tag{35c}$$

From (30) and (31)

$$\omega=\frac{2\pi}{A}\left(m-\frac{1}{2}\right), \tag{36a}$$

$$\frac{d\omega}{dA}=-\frac{2\pi}{A^2}\left(m-\frac{1}{2}\right), \tag{36b}$$

and we can rewrite Eq. (34) as an integral over the action (dropping subscripts *a, b, c, d*),

$$p(x,z,t)=-2\int dA\,\tau_x^{-1/2}A^{-3/2}\sum_m \left(m-\frac{1}{2}\right)^{1/2} \times\exp\left[i\frac{2\pi}{A}\left(m-\frac{1}{2}\right)(t-\tau_x-\tau_S+\tau_R)\right]. \tag{37}$$

The m summation is between limits m_L and m_U that depend on the action and the source bandwidth from $\omega_L=2\pi f_L$ to $\omega_U=2\pi f_U$ according to (36a):

$$m_L=f_LA+\tfrac{1}{2},\quad m_U=f_UA+\tfrac{1}{2}. \tag{38}$$

Equation (37) serves as a basis for demonstrating ray/mode duality. For a fixed receiver location at x,z, the pressure time series $p(x,z,t)$ is written as a sum of contributions from all excited modes m, and from all possible actions $A(\theta_0)$. Examination of Eq. (37) reveals that, in order for constructive interference to occur, the phasor exponent,

$$\phi=\frac{2\pi}{A}\left(m-\frac{1}{2}\right)(t-\tau_x-\tau_S+\tau_R), \tag{39}$$

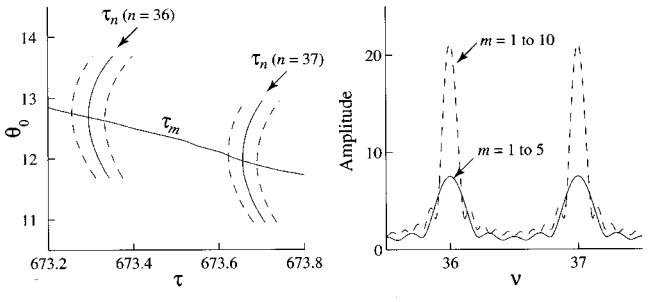


FIG. 8. (Left): The Chapman function (42) in a sound channel corresponding to $\bar{C}(z)$ in Fig. 2 for $n=36\pm M^{-1}$ and $n=37\pm M^{-1}$ at a range of 1 Mm. These are constructive interference peaks for $M=10$ modes. The minima correspond to the eigenrays with 36 and 37 downward loops and occur along the mode travel–time curve $\tau_m(\theta_0)$. (Right): Summation of modes 1–5 (solid) and 1–10 (dashed) according to the absolute value of the inner sum of (37). For fixed x and θ_0, the ordinate is a time axis proportional to $t/A(\theta_0)$. Constructive interference pulses occur at intervals of the action integral A (corresponding to integer values of ν). These pulses become increasingly sharper with an increasing number M of interacting modes.

must fulfill two conditions. The phase must be stationary for some value of A and must point in the same direction for different values of m. Thus

$$\frac{d\phi}{dA}=0, \tag{40a}$$

$$\frac{\Delta\phi}{\Delta m}=2\pi n, \tag{40b}$$

where n is an integer. We will examine these conditions in more detail.

Let us assume that the source depth and receiver depth are the same so that τ_S and τ_R are eliminated from (39). (This will result in considerably easier algebra; more general geometries are dealt with in MWW, p. 46.) From (40a) and (35a), interference can only occur when

$$t=\left(\tilde{S}-A\frac{d\tilde{S}}{dA}\right)x=\tau_m, \tag{41}$$

which can be recognized as the modal group velocity (23) times range. Similarly, (40b) leads to

$$t=x\tilde{S}+nA=\tau_n, \tag{42}$$

which is known as the Chapman equation. Since both conditions are functions of $\tilde{S}$ and remembering that $\tilde{S}=S_0\cos(\theta_0)$, we plot (41) and (42) as a function of θ_0 in Fig. 8 (left panel). Intersection points mark locations of interference peaks for various values of n. The travel–time minima correspond to eigenrays and occur along the mode travel–time curve $\tau_m(\theta_0)$. This sketch can be seen as an illustration of the Fermat principle that eigenrays are a minimum travel–time solution. The locus defined by (41) is just that of the bottom panel of Fig. 6 with the WKBJ approximation.

The sharpness of the peaks in Fig. 8 (left) is a function of the number of constructively interacting modes $M=m_U-m_L=A(\tilde{S})(f_U-f_L)$ in the inner sum of (37). To illustrate this point, the inner summation of (37) is plotted in Fig. 8 (right) against $\nu=(t-\tau_x)/A$ which, for fixed x and $\tilde{S}$, can be regarded as a plot against time. As the number of modes

increases the sum becomes higher and narrower. Either increased source bandwidth or increased inclination can result in a larger M. In the limit of infinite bandwidth the signal consists of delta functions corresponding to ray arrivals. Early arrivals correspond to large θ_0 and large A, have a large number of constructively interfering modes, and give sharp ray-like arrivals. For later arrivals, θ_0 and A both diminish and in the limit the arrivals lose their ray-like character. The factor $A^{-3/2}$ in the integrand (37) is associated with the high intensity of the final cutoff.

To complete the comparison with rays, define

$$R(S,\tilde{S}) = \frac{-dA}{d\tilde{S}} = 2\int_{\tilde{z}^-}^{\tilde{z}^+} \frac{\tilde{S}dz}{(S^2(z)-\tilde{S}^2)^{1/2}} \tag{43}$$

and

$$T(S,\tilde{S}) = A + \tilde{S}R = 2\int_{\tilde{z}^-}^{\tilde{z}^+} \frac{S^2(z)dz}{(S^2(z)-\tilde{S}^2)^{1/2}} \tag{44}$$

so that at the stationary point,

$$x = nR, \quad \tau_n = nT. \tag{45}$$

These definitions have a simple ray-geometric interpretation (MWW, p. 44): R is the double (upper plus lower) loop range and T is the double loop travel time. These are precisely the results for ray arrivals, and we can identify the integer n with the number of ray double loops for a ray with axial inclination θ_0. As an aside, note that (44) is the Hamiltonian for ray acoustics [see Miller (1986)] and will be used extensively in Sec. VI in the discussion on perturbation analysis.

Returning to the four ray arrivals described by (32), we have written $p(t)$ for the four constituents p_a, p_b, p_c, p_d of a wave group (MWW, p. 49). A group is characterized by the number n of double (upper and lower) loops. A group with n double loops always has $n^- = n$ *lower* loops, and has $n^+ = n, n, n-1, n+1$ *upper* loops. The first two constituents have the same number of upper loops as lower loops, but correspond to an upward and a downward launch angle. For an axial source and receiver they have the same travel time, τ_n. The last two constituents have one less and one more upper loop than lower loop and correspond to downward and upward launch angles, respectively.

C. Synthetic arrival pattern

The synthetic arrival pattern previously shown in Fig. 3 (lower panel) was computed for the experimental conditions described in Sec. II. The *a priori* sound–speed field was derived using the Del Grosso equation from temperatures and salinities in the WOA-94 July climatology. Mode shapes and wave numbers were calculated by a spectral approximation method with Chebyshev polynomials as basis functions [Dzieciuch (1993)]. Mode amplitudes and phases were computed over the entire aperture of the array at 1 Hz intervals across the transmission bandwidth. Since the results (amplitudes and group velocities) vary slowly with frequency, they can be accurately interpolated onto a finer grid (necessary to span the entire arrival time, since $\delta f = 1/T$), increasing computational efficiency. The arrival pattern was then synthesized with the Fourier components across the bandwidth. The pattern was computed assuming adiabatic propagation.

The synthetic arrival pattern shown in Fig. 3 (lower panel) for the 28 Hz signal as received in Hawaii contains time fronts that correspond to rays in the early portion of the pattern. Mode arrivals are clearly recognized in the later portion by the null pattern which delineates the nodes of individual modes in depth.

In the subsequent analysis the following procedure was used. For a selected axial inclination θ_0, form the beam given by Eq. (12) with time delays for each hydrophone given by (7) for the turning-point filter or by (8) for the linear beamformer. For the linear beamformer all hydrophones were delayed and then summed. For the turning-point filter, only the hydrophones within turning-point depths corresponding to the axial inclination were included. As the inclination approaches zero, the turning-point depths converge towards the sound–speed axis and fewer hydrophones are therefore included in the sum.

Figure 9 compares the performance of the linear beamformer (upper panel) with that of the turning-point filter (middle panel); the bottom panel will be considered in Sec. V. The gap in the turning-point filter at zero axial inclination is real and is associated with the fact that modes are standing waves composed of upward-going and downward-going waves. Overlaid on both panels is the expected travel–time versus axial-inclination dispersion curve from the bottom panel of Fig. 6. Ideally, peaks in both the linear beamformer and the turning-point filter outputs should lie along the dispersion curve.

The peaks in the linear beamformer output do not lie exactly along the dispersion curve. This bias is greater for the late arrivals at low inclination as predicted by the angle, depth diagram of Fig. 4. Although the peaks of the linear beamformer output are biased with respect to the dispersion curve, the root mean square (rms) travel–time error for peaks identified as ray arrivals is only 6.5 ms with respect to the ray travel times computed from the synthetic wave front. For travel times exceeding 2361 s, energy leaks towards small θ_0, resulting in sidelobes. These are the result of the mismatched linear lags and are not present for the turning-point filter (middle panel).

The peaks of the turning-point filter output lie along the travel-time, inclination curve as they should. Inspecting the panel, one can identify ray arrivals as the pairs of peaks arriving before 2363.5 s. The rms differences between the ray peaks in the turning-point filter output and the ray travel times computed from the synthetic wave front are 4.7 ms for this particular experimental geometry.

Looking toward the latest arriving energy, peaks associated with modes one and two are identifiable. The travel–time difference between the turning-point filter output associated with mode one and the mode one travel time computed from the synthetic wave front is 4 ms.

In the transition zone, between rays and modes, there is energy in the turning-point filter output that cannot be identified with a particular ray or mode. The interference pattern still produces peaks though and these can be compared to the predicted dispersion curve. This simple interpretation allows

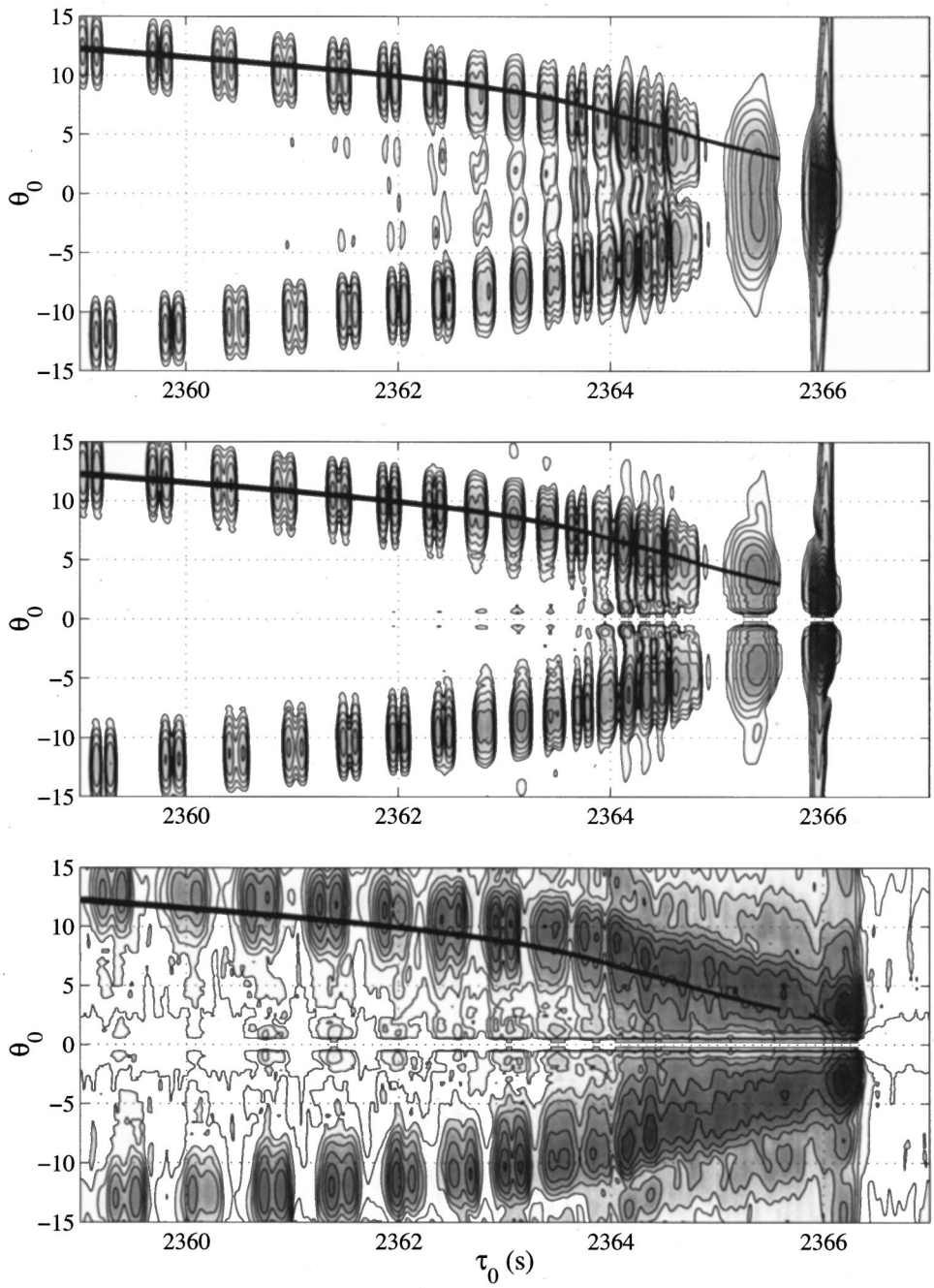


FIG. 9. Intensity in axial travel-time, inclination space for the 28 Hz Pioneer Seamount to Hawaii path contoured at 3 dB intervals with high intensity darkened. The heavy dark line is the expected dispersion from Fig. 6, using the Levitus climatology (Fig. 2). Upper panel: Beam pattern formed with linear time lags (7) operating on synthetic arrival pattern data (Fig. 3, upper panel). Middle panel: Turning-point filter beam pattern formed with nonlinear time lags (6) operating on synthetic arrival pattern data. Lower panel: Turning-point filter beam pattern formed, operating on the observed arrival pattern data (Fig. 3, lower panel). The early pattern is associated with separate ray arrivals. The late pattern is associated with low mode numbers.

analysis of nonidentifiable shallow axial-inclination data, providing additional information on the sound–speed profile near the axis. Hence, the turning-point filter contains information on rays, modes, and the transition zone.

V. EXPERIMENTAL RESULTS

In Sec. IV, we analyzed a synthetic (but realistic) time, depth arrival pattern. Now we compare the measured and predicted arrival patterns (Fig. 3). The measured data are the incoherent average of 40 transmissions at 28 Hz. The close resemblance in the overall structure, the number of wave fronts and the spacing between them, allows a detailed comparison of features.

As previously noted, the most important difference between the measured and computed patterns is the time shift. This of course is the basis for acoustic thermometry. The measurements lag the expected pattern by about 0.25 s, implying that the temperature of the ocean water along the path is colder than the climatology would indicate. But some of this difference might be due to an error in the sound–speed equation at high pressure.

The second most important difference is that the measured pattern is noticeably diffuse compared to the predicted. Ocean internal waves are responsible for this effect [see Colosi *et al.* (1994)].

One can easily identify the ray fronts in both the measured and computed patterns on the left side (earlier arrival times) of Fig. 3. At some point later in the arrival pattern the ability to identify ray fronts is lost. The latest arriving energy can be identified with low modes (or shallow inclinations) in the computed pattern. In the measured pattern this is not

possible except perhaps for the gravest mode.

To extract observables from the later portion of the arrival pattern, turning-point filtering has been applied. The actual data (Fig. 9, lower panel) compares favorably to the synthetic data (middle panel). The measurements are delayed relative to the prediction, and internal wave scattering has diffused the energy. The scattered energy broadens the pulses and generates sidelobes, increasing the measured variance in travel time and inclination. The scattering increases for the late-arriving, low-angle arrivals.

Several observables can be extracted from these data. Early arrivals can be identified with particular rays, and thus travel time and inclination can be measured for a particular ray. The large final peak is the lowest mode, so its travel time and inclination can be measured as well. Between the early rays and late modes there is a ridge of energy that cannot be distinctly identified with any particular ray or mode. Peaks located along this ridge can be used as observables without identification as a particular ray or mode. The distance of the data peaks from the overlaid expected dispersion curve of Fig. 6 can be measured in both travel time and inclination.

VI. PERTURBATION ANALYSIS

In the Sec. IV, we have considered how to calculate the location of peaks in travel-time, inclination space. In an experimental situation the peak locations are the observables that are measured. Differences between the observations and the predictions in both travel time and inclination are the information used in the tomographic inverse problem to infer ocean parameters, heat content, for example. In the following analysis, we relate the perturbations in travel time and inclination to perturbations in ocean sound-speed.

Suppose the ocean along our experimental path characterized by climatological data (Fig. 2) has been perturbed by a warming, decaying exponentially down from the surface (Fig. 10). After applying the turning-point filter to both the unperturbed and perturbed cases, the travel-time/inclination data are displayed in Fig. 11. The curve itself is not readily observable but peaks that lie along it are. The peaks can be classified into three sets: ray peaks, mode peaks, and (nonidentifiable) transition peaks that lie in between.

For each of the three types of peaks, there are two measurable quantities, the travel time and the inclination. Inverse methods require expressions for the relationship between these measurements and the ocean properties to be inferred. We derive the first order terms of the Taylor expansion for range-independent perturbations to a range-independent ocean. Adiabatic extensions to the theory could be derived like those in other work [Miller (1986) or Wunsch (1987)]. An important point is that perturbations in travel time and inclination are a function only of θ_0 (Fig. 11) and it is appropriate to use WKBJ theory.

We consider travel-time changes for ray peaks, transition peaks, and mode peaks in turn. We then consider inclination changes for each case (Fig. 12). Although travel time has historically been the most useful tool for tomographic experiments, measurements of inclination could in principle also provide useful tomographic

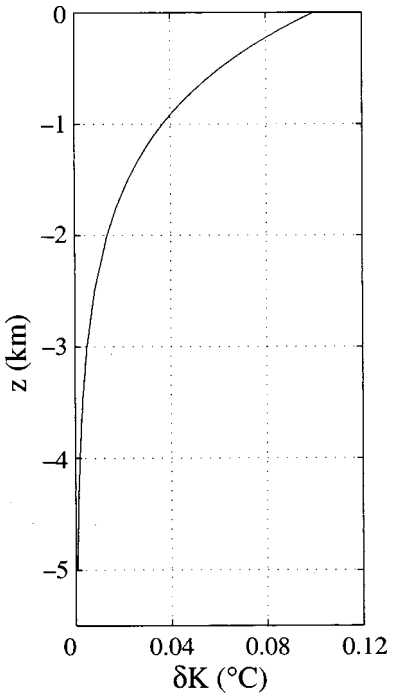


FIG. 10. Temperature perturbation of 0.1 °C and exponential decay length of 1 km.

information. In practice, the usefulness of inclination measurements is problematic, since the precision of such measurements depends on both the array length and on the signal-to-noise ratio. We will proceed as if the inclination measurements are meaningful, but the cautious experimenter will give the noise field detailed consideration to guarantee that the travel-time and inclination measurements are independent.

A. Travel time

First consider the travel-time equation from (41) and (43),

$$\tau = x s_g = x\left(\tilde{S} + \frac{A}{R}\right). \tag{46}$$

This equation is the WKBJ approximation to the travel-time, inclination curve shown in the lower panel of Fig. 6.

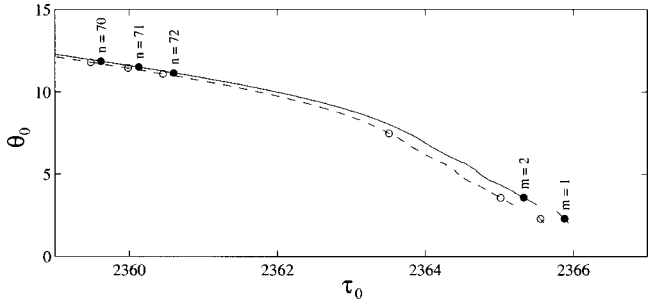


FIG. 11. Computed dispersion in τ_0, θ_0 space. The solid line shows the expected dispersion using the Levitus climatology (Fig. 2). The dashed line shows the dispersion with an added temperature perturbation (Fig. 10). The perturbed dispersion curves again condense along a single curve consistent with the WKBJ approximation. Rays, $n = 70$, 71, and 72 and modes $m = 1$, and 2 are indicated. A nonidentifiable transition peak is shown also.

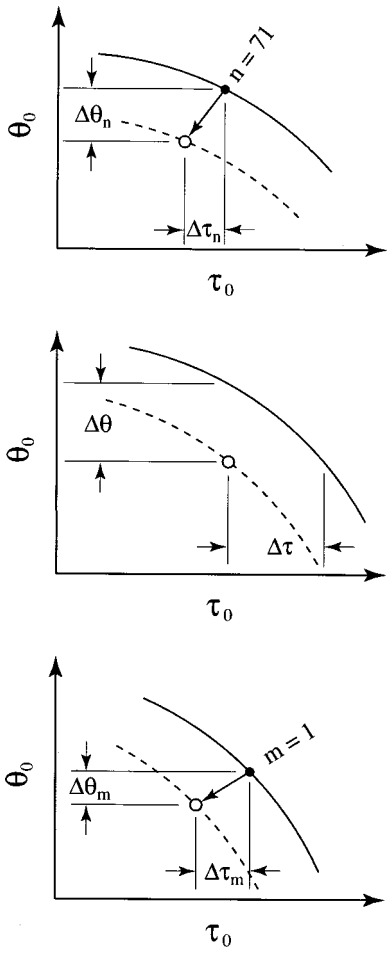


FIG. 12. Cartoon showing an expanded view of Fig. 11, where the solid lines show expected dispersion curves and the dashed lines show perturbed dispersion curves. The closed circles are expected peaks and the open circles are perturbed peaks: (upper panel) ray peak, (middle panel) nonidentified transition peak, and (lower panel) mode peak.

1. Ray peaks

The travel-time perturbation of a ray is measured relative to its predicted travel time as in Fig. 12 (upper panel). From Eq. (45), at a fixed receiver range, for a particular ray with n double loops,

$$\Delta x = n\Delta R + R\Delta n = 0. \tag{47}$$

Since the ray has an unchanging n,

$$\Delta n = 0 \Rightarrow \Delta R = 0, \tag{48}$$

thus

$$\Delta\tau_n = x\Delta s_g = x\left(\Delta\tilde{S} + \frac{\Delta A}{R} - \frac{A}{R^2}\Delta R\right). \tag{49}$$

Using $\Delta R = 0$ and the action definition,

$$A(S,\tilde{S}) = 2\int_{\tilde{z}^-}^{\tilde{z}^+} (S^2(z) - \tilde{S}^2)^{1/2} dz, \tag{50}$$

(49) can be expanded as

$$\Delta\tau_n = x\left[\Delta\tilde{S} + \frac{1}{R}\left(\frac{\partial A}{\partial\tilde{S}}\Delta\tilde{S} + \frac{\partial A}{\partial S}\delta S\right)\right], \tag{51}$$

and because $\partial A/\partial\tilde{S} = -R$, simplified to

$$\Delta\tau_n = n\frac{\partial A}{\partial S}\delta S. \tag{52}$$

The δ notation denotes a functional perturbation so

$$\frac{\partial A}{\partial S}\delta S = 2\int_{\tilde{z}^-}^{\tilde{z}^+} \frac{S\,\delta S(z)dz}{(S^2(z) - \tilde{S}^2)^{1/2}}. \tag{53}$$

It is convenient to parameterize sound-slowness perturbations by writing the slowness as the finite-dimensional sum of a reference and a perturbation.

$$S(z) = \tilde{S}(z) + \sum_i l_i F_i(z). \tag{54}$$

Thus sound-slowness perturbations,

$$\delta S = \sum_i F_i(z)\Delta l_i, \tag{55}$$

are characterized by the set of Δl_i. Equations (52), (53), and (55) are then combined in the matrix equation,

$$\Delta\tau_n = \sum_i n W^{\tau}_{\text{ray}}(\tilde{S},F_i)\Delta l_i, \tag{56}$$

with weighting function

$$W^{\tau}_{\text{ray}}(\tilde{S},F_i) = 2\int_{\tilde{z}^-}^{\tilde{z}^+} \frac{SF_i(z)dz}{(S^2(z) - \tilde{S}^2)^{1/2}}. \tag{57}$$

This parameterization is the basis of an inverse solution.

The integrand in (57) is singular at the endpoints so care is needed in its evaluation; nonetheless for physically reasonable choices of $F_i(z)$ the integral converges. This is the standard equation found in the work of MWW (p. 51) and has a simple interpretation that the ray travel time observable is the integral of the sound-slowness change along the unperturbed ray path.

Given the data $\Delta\tau_n$, solving the inverse problem (56) is covered in great detail in MWW and will not be repeated here. The solution requires specification of the noise covariance and the *a priori* variation of the sound–speed field, which would require a more oceanographic focus than intended.

2. Transition peaks

In this case, the travel time observable is defined to be the difference in travel time between the transition peak and the reference dispersion curve at constant θ_0 (or equivalently $\tilde{S}$) [see Fig. 12 (middle panel)]. So at a fixed range,

$$\Delta x = 0, \quad \Delta\tilde{S} = 0. \tag{58}$$

Starting with the travel-time equation (46), remembering the definition of $R(S,\tilde{S})$ (43), and proceeding as before, (49) can be simplified to

$$\Delta\tau = \frac{x}{R}\left(\frac{\partial A}{\partial S}\delta S - \frac{A}{R}\frac{\partial R}{\partial S}\delta S\right). \tag{59}$$

The first term is the same as that for rays (52), but because a ray has not been identified, R is not fixed, and thus further simplification is not possible. The second term contains the functional perturbation of R,

$$\frac{\partial R}{\partial S}\,\delta S = 2\int_{\tilde{z}^-}^{\tilde{z}^+} \frac{\tilde{S}S\,\delta S(z)\,dz}{(S^2(z)-\tilde{S}^2)^{3/2}}. \tag{60}$$

As in (57), the integrand in (60) is singular at the endpoints but for reasonable choices of δS it converges.

3. Mode peaks

Finally consider a mode travel time observable as in Fig. 12 (lower panel). In this instance, measure the perturbation in the mode peak travel time from the expected. We will assume that the source bandwidth is large enough so that there is a mode travel-time peak and that it can be resolved from other modes. In practice this may require a mode-inverse filter (see Sutton *et al.* (1994)]. For certain ocean sound–speed profiles and ranges this may not be possible [see Brown *et al.* (1996) for further discussion of these issues].

In this first-order analysis, the travel-time perturbation of the mode peak due to environmental perturbations is characterized by the group-slowness perturbation at the carrier frequency f_c of a broadband source. Differentiating (36) gives

$$\Delta A = \frac{2\pi}{\omega_c}\Delta m - \frac{2\pi}{\omega_c^2}\left(m-\frac{1}{2}\right)\Delta\omega_c. \tag{61}$$

So for a particular mode, at carrier frequency $\omega_c = 2\pi f_c$,

$$\Delta m = 0, \quad \Delta\omega_c = 0 \Rightarrow \Delta A = 0, \tag{62}$$

and thus the differential of the travel-time equation (49) can be written as

$$\Delta\tau = x\left(\Delta\tilde{S} - \frac{A}{R^2}\Delta R\right). \tag{63}$$

Using the fact that the total derivative of (50) can be written as

$$\Delta A = 0 = \frac{\partial A}{\partial S}\,\delta S + \frac{\partial A}{\partial \tilde{S}}\Delta\tilde{S} \Rightarrow \Delta\tilde{S} = \frac{1}{R}\frac{\partial A}{\partial S}\,\delta S, \tag{64}$$

the mode travel-time perturbation is therefore

$$\Delta\tau_m = \frac{x}{R}\left[\left(1-\frac{A}{R^2}\frac{\partial R}{\partial \tilde{S}}\right)\frac{\partial A}{\partial S}\,\delta S - \frac{A}{R}\frac{\partial R}{\partial S}\,\delta S\right]. \tag{65}$$

This result is the same as (59) with the addition of a third term. With a modern computer and reasonable choices for δS it can be easily evaluated. As an aside, note that if the source is continuous wave (cw), then mode phase can be used as an observable via $\Delta\phi_m = 2\pi f_c \Delta\tau_m$.

B. Inclination angle (turning-point slowness)

Now consider inclination measurements. The travel-time equation (46) and its constituents A and R are functions of $\tilde{S}$, which is related to θ_0 through the axial sound-slowness by $\tilde{S} = S_0\cos(\theta_0)$. All figures in this paper drawn with a θ_0 ordinate could have (and perhaps should have) been drawn with an $\tilde{S}$ ordinate. Thus the observable is $\Delta\tilde{S}$.

Again consider in turn the measurable changes in $\tilde{S}$ for ray, transition, and mode peaks. We do not attempt to address the issue of the signal-to-noise ratio of such measurements, i.e., the size of $\Delta\tilde{S}$ for a typical ocean variation compared to the array angular resolution. As far as the authors know, the following expressions relating inclination changes to ocean sound-slowness changes have not appeared in the literature previously.

1. Ray peaks

The change in the ray-peak inclination observable is measured as in Fig. 12 (upper panel). The turning-point slowness observable would then be defined as $\Delta\tilde{S} \equiv -S_0\sin(\theta_0)\Delta\theta_0$. For ray peaks with a constant n, (48) gives $\Delta R = 0$, so differentiating (44) gives

$$\Delta T = \Delta A + R\Delta\tilde{S}. \tag{66}$$

Since T is a function of $\tilde{S}$ and S, both sides of (66) can be expanded as

$$\Delta T = \frac{\partial T}{\partial S}\,\delta S + \frac{\partial T}{\partial \tilde{S}}\Delta\tilde{S} = \frac{\partial A}{\partial S}\,\delta S + \frac{\partial A}{\partial \tilde{S}}\Delta\tilde{S} + R\Delta\tilde{S}. \tag{67}$$

The last two terms cancel so (67) can be rearranged as

$$\Delta\tilde{S}_n = \left(\frac{\partial T}{\partial \tilde{S}}\right)^{-1}\left(\frac{\partial A}{\partial S}\,\delta S - \frac{\partial T}{\partial S}\,\delta S\right). \tag{68}$$

By using

$$\frac{\partial T}{\partial \tilde{S}} = \tilde{S}\frac{\partial R}{\partial \tilde{S}}, \tag{69}$$

and

$$\frac{\partial T}{\partial S}\,\delta S = \frac{\partial A}{\partial S}\,\delta S + \tilde{S}\frac{\partial R}{\partial S}\,\delta S, \tag{70}$$

one gets

$$\Delta\tilde{S}_n = -\left(\frac{\partial R}{\partial \tilde{S}}\right)^{-1}\frac{\partial R}{\partial S}\,\delta S. \tag{71}$$

This expression relates the measured change in ray turning-point slowness to changes in the ocean environment using the functional perturbation of R as in (60).

2. Transition peaks

The difference in inclination (or, more correctly, $\tilde{S}$) between the nonidentified transition peak and the expected dispersion curve at a constant travel time (the travel time of the peak) can be measured as shown in Fig. 12 (middle panel) and also used as an observable.

Once again, an expression relating $\Delta\tilde{S}$ to δS is needed, so using $\Delta\tau = 0$ and (49) gives

$$\Delta s_g = \Delta\tilde{S} + \frac{\Delta A}{R} - \frac{A}{R^2}\Delta R = 0. \tag{72}$$

By expanding,

$$\Delta\tilde{S} + \frac{1}{R}\frac{\partial A}{\partial \tilde{S}}\Delta\tilde{S} + \frac{1}{R}\frac{\partial A}{\partial S}\delta S - \frac{A}{R^2}\frac{\partial R}{\partial \tilde{S}}\Delta\tilde{S} - \frac{A}{R^2}\frac{\partial R}{\partial S}\delta S = 0. \tag{73}$$

The first two terms cancel so (73) can be rearranged as

$$\Delta\tilde{S} = -\left(\frac{\partial R}{\partial \tilde{S}}\right)^{-1}\left(\frac{\partial R}{\partial S}\delta S - \frac{R}{A}\frac{\partial A}{\partial S}\delta S\right). \tag{74}$$

This result contains the first term of (71) with the addition of a second term containing the functional perturbation of A (53).

3. Mode peaks

For a mode peak as shown in Fig. 12 (lower panel), $\Delta A = 0$, so

$$\Delta A = \frac{\partial A}{\partial S}\delta S + \frac{\partial A}{\partial \tilde{S}}\Delta\tilde{S} = 0. \tag{75}$$

Combining (75) with the definition of R gives

$$\Delta\tilde{S}_m = \frac{1}{R}\frac{\partial A}{\partial S}\delta S, \tag{76}$$

and again it is expressed in terms of a functional perturbation of A (53). This is a most interesting result. It asserts that the perturbation in mode turning-point slowness is nearly equivalent to the perturbation in travel time for an equivalent ray (52): the integral of the sound-slowness perturbation along an unperturbed ray path with an $\tilde{S}$ equal to the modal $\tilde{S}$ divided by range equals the perturbation in mode turning-point slowness!

Implicit in this broadband analysis is the framework appropriate for a single frequency experiment. In such an instance, mode travel time would not be resolvable, but mode turning-point slowness (i.e., inclination) would still be measurable. Turning-point slowness resolution (almost the same as inclination resolution) would likely become an issue, and high resolution methods could be necessary. Nonetheless, perhaps the experimental focus should not be on measuring mode travel time, or modal phase perturbations at a single frequency [see Shang (1989), for example], but on mode turning-point slowness.

At this point, it is perhaps worthwhile to remember that all the above analysis has been for the first-order term of the range-independent sound-slowness perturbation. No attempt has been made to characterize higher-order terms in the range-independent expansion or for range-dependent cases.

VII. DISCUSSION

Using a turning-point filter it is possible to extract robust observables from the entire arrival pattern. The observables are peaks in τ_0, θ_0 space identified with particular rays or modes, as well as nonidentifiable transition peaks. The turning-point filter is not biased and all peaks lie on the travel-time, inclination dispersion curve.

The turning-point filter approaches a linear beamformer when the time front curvature is small. The linear beamformer pattern is poorly focused for the late arrivals when the time front curvature is large, thus it is hard to separate observed scattering from the expected diffraction. For the turning-point filter the pattern is much better focused so scattering can be attributed to real ocean processes (such as internal waves).

The measurable properties of the observables can be expressed in terms of variable ocean properties, so using inverse theory, travel time and inclination can be used to investigate those ocean properties. The inclusion of nonidentifiable peaks in the set of observables increases the amount of information that can be extracted from tomographic experiments, particularly at small inclinations where there is a lack of identifiable rays. At present, the only assumption is that the propagation can be modeled as adiabatic and thus acoustic energy is not exchanged between modes traveling with different wave numbers. This appears to be a reasonable assumption for the low-frequency (28 Hz) data presented here, but may not hold for higher frequencies.

The turning-point filter implicitly includes modal dispersion and unambiguously provides an observable for broadband modal arrivals. There is no need to separate modes since those with the same turning-point slowness carry the same information about the ocean and an inverse solution would recombine them anyway. The turning-point filter includes single frequency mode experiments as a special case. As far as the authors know the turning-point filter results, particularly those regarding the inclination observable, have not previously appeared in the literature.

ACKNOWLEDGMENTS

The authors wish to thank the members of the Alternate Source Test (AST) Group (Theodore G. Birdsall, Bruce M. Howe, James A. Mercer, Kurt Metzger, and Robert C. Spindel) for making available the data used in this article. B. Betts and E. Blackmore assisted with preparation of the manuscript. Two HLF-6A low-frequency acoustic sources were modified for the AST as part of the Joint Environmental Test Initiative sponsored by SPAWAR PMW-182. The AST was supported by the Office of Naval Research (Grant No. N00014-95-1-0800 to APL-UW and Grant No. N00014-95-1-0589 to SIO) and by the Strategic Environmental Research and Development Program through Defense Advanced Research Project Agency Grant No. MDA972-93-1-0003. This research was also supported in part by Office of Naval Research Grant No. N00014-97-1-0258. One of the authors (M.D.) received additional support from NASA, through MIT Subcontract No. 5700000137. One of the authors (W.M.) holds a Secretary of the Navy Research chair in Oceanography. The authors thank the two reviewers for many helpful suggestions.

ATOC Consortium, Baggeroer, A. B., Birdsall, T. G., Clark, C., Colosi, J. A., Cornuelle, B. D., Costa, D., Dushaw, B. D., Dzieciuch, M., Forbes, A. M. G., Hill, C., Howe, B. M., Marshall, J., Menemenlis, D., Mercer, J. A., Metzger, K., Munk, W., Spindel, R. C., Stammer, D., Worcester, P. F., and Wunsch, C. (**1998**). ''Ocean climate change: Comparison of acoustic tomography, satellite altimetry, and modeling,'' Science **281**, 1327–1332.

Brekhovskikh, L. (**1980**). *Waves in Layered Media*, 2nd ed. (Academic, New York).

Brekhovskikh, L., and Lysanov, L. (**1991**). *Fundamentals of Ocean Acoustics*, 2nd ed. (Springer, New York).

Brown, M. G., Viechnicki, J., and Tappert, F. D. (**1996**). ''On the measurement of modal group time delays in the deep ocean,'' J. Acoust. Soc. Am. **100**, 2093–2102.

Colosi, J. A., and the ATOC Group. (**1999**). ''A review of recent results on ocean acoustic wave propagation in random media: Basin scales,'' IEEE J. Ocean Eng. **24**, 138–155.

Colosi, J. A., Flatté, S. M., and Bracher, C. (**1994**). ''Internal-wave effects on 1000-km oceanic acoustic pulse propagation,'' J. Acoust. Soc. Am. **96**, 452–468.

Cox, H., Zeskind, R. M., and Meyers, M. (**1990**). ''A subarray approach to matched-field processing,'' J. Acoust. Soc. Am. **87**, 168–178.

Dzieciuch, M. (**1993**). ''Numerical solution of the acoustic wave equation using Chebyshev polynomials with application of global acoustics,'' *Proceedings of IEEE Oceans '93* (IEEE, Piscataway, NJ), Vol. No. 1, pp. 267–271.

Guthrie, K. M., and Tindle, C. (**1976**). ''Ray effects in the normal modes approach to underwater acoustics,'' J. Sound Vib. **47**, 403–413.

Levitus, S. (**1994**). *World Ocean Atlas, Vol. 4 Temperature*, NOAA Atlas NESDIS No. 4 (NOAA, Washington DC).

Levitus, S., Burgett, R., and Boyer, T. (**1994**). *World Ocean Atlas, Vol. 3 Salinity*, NOAA Atlas NESDIS No. 3 (NOAA, Washington D.C.).

McDonald, B. E., Collins, M. D., and Kuperman, W. A. (**1994**). ''Comparison of data and model predictions for Heard Island acoustic transmissions,'' J. Acoust. Soc. Am. **96**, 2357–2370.

Metzger, K. (**1983**). ''Signal processing equipment and techniques for use in measuring ocean acoustic multipath structures,'' Ph.D. thesis, University of Michigan, Ann Arbor, MI.

Miller, J. C. (**1986**). ''Hamiltonian perturbation theory for acoustic rays in a range-dependent sound channel,'' J. Acoust. Soc. Am. **79**, 338–346.

Munk, W. H., Worcester, P. F., and Wunsch, C. (**1995**). *Ocean Acoustic Tomography* (Cambridge University Press, Cambridge).

Roden, G. I. (**1975**). ''On North Pacific temperature, salinity, sound velocity and density fronts and their relation to the wind and energy flux fields,'' J. Phys. Oceanogr. **5**, 557–571.

Shang, E. C. (**1989**). ''Ocean acoustic tomography based on adiabatic mode theory,'' J. Acoust. Soc. Am. **85**, 1531–1537.

Shang, E. C., Wang, Y. Y., and Georges, T. M. (**1994**). ''Dispersion and repopulation of Heard–Ascension modes,'' J. Acoust. Soc. Am. **96**, 2371–2379.

Smith, W. H. F., and Sandwell, D. T. (**1997**). ''Global sea floor topography from satellite altimetry and ship depth soundings,'' Science **277**, 1956–1962.

Sutton, P., Morawitz, W. M. L., Cornuelle, B. D., Masters, G., and Worcester, P. F. (**1994**). ''Incorporation of acoustic normal mode data into tomographic inversions in the Greenland Sea,'' J. Geophys. Res., **99**(C6), 12487–12502.

Worcester, P. F., Cornuelle, B. D., Dzieciuch, M. A., Munk, W. H., Howe, B. M., Mercer, J. A., Spindel, R. C., Colosi, J., Metzger, K., Birdsall, T. G., and Baggeroer, A. B. (**1999**). ''A test of basin-scale acoustic thermometry using a large-aperture vertical array at 3250-km range in the eastern North Pacific Ocean,'' J. Acoust. Soc. Am. **105**, 3185–3201.

Wunsch, C. (**1987**). ''Acoustic tomography by Hamiltonian methods including the adiabatic perturbation,'' Rev. Geophys. **25**, 41–53.

Broadband parametric imaging of breaking ocean waves

Rex K. Andrew[a)]
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105-6698

David M. Farmer[b)]
Institute of Ocean Sciences, P.O. Box 6000, Sidney, British Columbia V8L 4B2, Canada

R. Lynn Kirlin[c)]
Department of Electrical and Computer Engineering, University of Victoria, P.O. Box 3055, Victoria, British Columbia V8W 3P6, Canada

(Received 22 September 1999; revised 6 March 2001; accepted 26 March 2001)

An acoustic array was deployed in the near-surface layer of a fetch-limited coastal inlet to image breaking waves using only the sound radiated in the band (400 Hz to 2000 Hz) from the breaking region. The breakers were assumed to possess predominantly spiller characteristics. For this frequency band, the wavelength of sound in bubble-free water is much larger than the surface wave height and the depth of the breaker bubble plume, so both were considered insignificant. The 15-element array was configured as a sparse horizontal cross with an 8 m aperture, bottom moored, and positioned nominally 3 m beneath the surface. Propagation from the source to the array elements assumed dipole sources, an acoustically flat surface, and an acoustically thin bubble plume. The radiating region was parameterized by a broadband two-dimensional Gaussian profile: information from up to six independent frequencies was combined to yield a maximum-likelihood image. Analysis shows that the images align closely with the wind and can be observed moving downwind with a speed roughly equal to 70% of the phase speed of the dominant wind waves. A model of acoustic source strength which is linear in log frequency is found fit the data reasonably well, and model parameters are provided for a single wind speed. Unlike other imaging experiments, this technique provides measurements of the size and shape of the bubble-creation region at or near the peak of the radiated autospectrum.
[DOI: 10.1121/1.1377870]

PACS numbers: 43.30.Pc, 43.30.Nb, 43.60.Gk [DLB]

I. INTRODUCTION

It is now well established that breaking wind waves are significant sources of ambient sound in the ocean. This acoustic radiation was originally considered *noise* by the sonar community; now, we are using it as a *signal* as we try to understand and quantify the fundamental processes involved in wave breaking.

As examples, the broadband acoustical intensity radiated from individual breaking waves in a laboratory wave tank has been shown to be well correlated with the mechanical energy dissipated by the breaking wave,[1] thereby suggesting a method for estimating the surface wave field dissipation. Likewise, the spatial statistics of waves tracked acoustically across the ocean surface can provide key parameters needed in models of surface wave field growth.[2] The radiated autospectrum may in addition be related to the bubble size distribution within the breaking region, thus yielding an indirect measure of the volume of gas entrained by the breaker.[3,4]

A fundamental issue is the size, shape, and behavior of the acoustically active source region. Whereas in previous modeling, breaking events could be adequately represented by *point* sources, and this poses a true *imaging* problem. Crowther and Hansla[5] made the first measurements utilizing a seven-beam sonar system configured like an insect eye. The system was bottom moored in 85 m of water and the beams provided seven overlapping footprints at the surface, with each footprint diameter approximately 9.7 m at 24 kHz. Their results showed the acoustically active region to be considerably larger than the apparent visual size of the white cap. In a second experiment, Epifanio and Buckingham[6] used a sonar system containing 126 beams at a slant range from the surface of 45 m, operating in the range of 8–80 kHz. Each beam had a footprint diameter of approximately 0.75 m at 80 kHz. Their results showed that a sea surface with no apparent breaking activity still had an acoustically active spatial structure that evolved on time scales at least as short as 40 ms.

Both of these experiments imaged the breakers at frequencies well above the peak of their radiated spectral signature. There is, however, mounting evidence that the fundamental physical processes of surface wave field dissipation, turbulent dissipation, and air–sea gas transfer are closely linked to the energetics associated with the shape of the radiated autospectrum around its maximum. As examples, freshly entrained ''big'' bubbles, which resonate individually slightly above the spectral peak, essentially domi-

[a)]Author to whom correspondence should be addressed. Electronic mail: randrew@apl.washington.edu
[b)]Electronic mail: farmerd@dfo-mpo.gc.ca
[c)]Electronic mail: kirlin@ece.uvic.ca

nate the total contribution of atmospheric gas entrained during the break.[8] Spectral energy below the peak may indicate large scale plunging,[9] which may bound the outer scales of turbulent mixing during breaking. There may also be a connection among the bubble size distribution, bubble breakup, and turbulent dissipation.[10] Therefore, breaking wave images formed at frequencies near the peak of the radiated spectral signature, as opposed to frequencies well above this peak, are observations of the acoustical representation of these fundamental physical processes at the particular frequencies that ought to be most directly linked to these processes.

Imaging near the spectral peak, however, presents a new set of challenges. The expected dimensions of the radiating regions of coastal breakers are approximately an acoustic wave length at 500 Hz in bubble-free seawater; this implies that the source is acoustically compact, and that conventional bottom-moored imaging systems cannot resolve the spatial structure of the source region.

The conventional approach was therefore abandoned in favor of deploying the imaging array in the near field, at a nominal depth less than the aperture size. This near-field array has a much smaller field of view (about $10\,\mathrm{m}\times 10\,\mathrm{m}$) than those of previous imaging experiments, but has theoretical resolution of about 0.2 m at 2000 Hz, which is generally smaller than the size scales of fully developed coastal or oceanic white caps.

The imaging problem, as presented here, is fundamentally an inverse problem, and some form of regularization is required to yield stable solutions. In this article, the imaging problem is stabilized by assuming the source is from a class of space–time separable sources,[11] and further parameterizing the spatial source profile with a two-dimensional (2D) Gaussian function.

The imaging problem is formulated by combining the fundamental source features, summarized in Sec. II, with propagation physics and a parametric likelihood-based statistical statement, described in Sec. III. Experimental data were gathered during a storm described in Sec. IV and processed according to details given in Sec. V. Results are presented in Sec. VI, including parametric images of breaking waves and estimated source strengths. The validity of the technique is demonstrated in Sec. VI, where the images are used to compute wind and wave field parameters consistent with those in effect during the storm. The key features of this approach are summarized in Sec. VII.

II. SOURCE MECHANISM

In a passive imaging application, the problem is to utilize measurements of an acoustic field $p_t(\mathbf{r},t)$ governed by an inhomogeneous wave equation,

$$\nabla^2 p_t(\mathbf{r},t) - \frac{1}{c(\mathbf{r})^2}\frac{\partial^2}{\partial t^2} p_t(\mathbf{r},t) = f(\mathbf{r},t), \tag{1}$$

to image the inhomogeneous term $f(\mathbf{r},t)$. In this particular application, there is considerable literature devoted to understanding the source mechanism represented by the inhomogeneous term, and this research can be used to establish a model for the inhomogeneous term and an algorithm to recover the parameters of the model. The key features of the source mechanism are presented below, and a plausible model for the inhomogeneous term is developed in Sec. III.

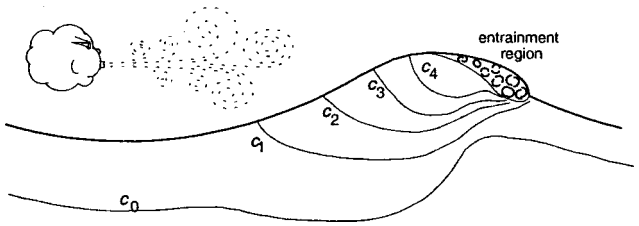


FIG. 1. Bulk sound speed analog for a breaking wave. A bulk or effective sound speed field $c(\mathbf{r})$, suggested here by representative isovelocity contours in an $x-z$ ''slice.''

A. Ringing bubbles

Bubbles are known to play an essential role in underwater sound generation.[12] Bubbles oscillate upon formation like impulsively forced mass-spring systems, giving rise to exponentially decaying sinusoidal acoustic pressure signals.[13,14,17] This can also occur in bubble breakup or coalescence.[15,16] The oscillation frequency is inversely related to the bubble radius, a relationship known as the Minnaert formula,[12] and experimental observations[14,18] suggest decay constants of the order of tens of milliseconds for the sizes of bubbles ordinarily encountered in the ocean under natural conditions.

Bubble size measurements of the freshly created, ringing bubbles is a problematic task: most early efforts took measurements outside the bubble-creation region, thus measuring the population of ''mature bubbles'' (see below) rather than that of active, ringing bubbles. Recent optical measurements inside the breaker identified bubbles with radii up to several millimeters, and air-filled ''filaments'' up to an order larger.[8,19] Passive acoustic methods,[4,18,20] which typically isolate the signatures of individual bubbles (and therefore are most useful for small-scale breakers), have identified bubbles with radii as large as approximately 7 mm, which, by the Minnaert formula, will resonate at about 400 Hz. This value coincides roughly with the peak of the radiated autospectrum. Deane[8,19] suggested that for radius $a>1$ mm, the distribution goes as $a^{-4.5}$, i.e., the bubble size distribution decreases for increasing bubble radius, so that the largest bubbles, those associated with the greatest acoustic intensity, are actually the least numerous.

B. Mature bubbles

Once the initial impulse response dies away, a bubble becomes ''mature'' and is no longer a primary radiator. Bubbles initially mature locally near the point of creation, producing the ''bubble plume'' associated with the breaker. After the wave travels on, the mature bubbles may linger to join the continuum bubble layer that builds up under the surface during windy conditions.[34] The presence of mature bubbles in the plume and the continuum layer introduces an effective local sound speed and absorption that is considerably different than the sound speed and absorption in bubble-free seawater.[35–39] This is shown schematically in Fig. 1.

Physically, the individual bubbles in a bubbly fluid scatter an ''averaged field,'' and so may be acting like secondary sources. This may produce a complicated combined refractive and masking problem between the source and a nearby imaging system. These mature bubbles, however, are no longer primary sources and should not be considered part of the source mechanism to be imaged.

C. Breaking waves

Bubbles are broadly distinguished into two categories.[21]

(1) **Plunging**. Finite amplitude water waves have been shown[22–24] to form a fluid jet near the crest. When the jet falls or ''plunges'' forward, it intercepts the smooth upstream fluid and traps a volume of air. This volume of air fractures into bubbles. The collapsing jet continues to drive the bubbles below the surface to create a high void fraction subsurface plume (i.e., a localized volume with a large concentration of bubbles.)
(2) **Spilling**. For less energetic waves, ''vortex ripples'' can occur on the forward face of a wave[25–30] which grow and eventually collapse into a turbulent two-phase flow (sometimes called a ''roller'' or a ''bore'') that rides down the forward face of the wave. Bubbles formed in the turbulent bore are entrained into the wake at the lower boundary of the bore.[31]

There is no exact operational procedure for distinguishing between spillers and plungers: at-sea breakers will typically possess characteristics of both. Visual observations suggest that vigorously plunging breakers away from shoals, the shore, and riptides are quite rare.[32,33] Dockside observations of breaking waves during the primary storm event of this experiment rarely if ever noted the sort of plunging breaker commonly associated with a breaking surf. Therefore, the breakers observed in this experiment are assumed to have predominantly ''spiller'' characteristics.

Deane[8,19] has made a thorough investigation of the acoustically active bubble region for plunging breakers, particularly shoaling breakers in surf. These plungers generate a high void fraction plume, as high as 30% to even 50%. (Similar values apply for at-sea breakers.[38]) Acoustic absorption within this plume is so high that only newly oscillating bubbles on the plume periphery are expected to emit acoustic radiation that remains audible in the far field—radiation from active bubbles in the plume interior is nearly completely masked by local absorption.

On the other hand, nominally ''mature'' bubbles advected into regions of high turbulence may break apart, or touch and coalesce, and these events will also be the site of acoustic emissions.[15,16]

The extent to which these features transfer to spilling breakers is not obvious. Bubbles formed at the forward face of the bore will radiate freely forward and downward into bubble-free water. Bubbles formed at the surface by vortex ripples ride atop a layer of advected freshly mature bubbles which will mask downward radiation, but this layer is also acoustically thin compared to a wavelength (in bubble-free water) at the frequencies under consideration. Bubbles advected away from the face of the bore or transported through the bore to the less-bubbly fluid below may re-radiate into the less-bubbly fluid below.

Unfortunately, however, not enough information is available regarding the refracting and masking properties of the fluid in a spilling breaker, so we have assumed for simplicity that bubbles formed at the front of and along the bore are both audible in the water below. Consequences of this assumption are discussed in Sec. VII.

D. Low- to midfrequency regime

Bubbles with radii larger than about 7 mm have thus far eluded detection; this leaves open the question of the source mechanism in the so-called low- to midfrequency (LMF) regime[40] of about 10 Hz to the peak of the autospectrum. Investigators have proposed turbulence amplification,[41,42] nonlinear acoustical interactions[43] and collective oscillations.[42,44]

This uncertainty has consequences for the imaging problem. A key feature of this work will be to combine solutions at multiple frequencies. This should constitute a valid technique as long as the inhomogeneous term obeys certain restrictions. According to an argument advanced below, these restrictions are plausible above the spectral peak; in the LMF regime, however, the restrictions may or may not hold (depending on the nature of the forcing function). Therefore, pending a more comprehensive understanding of the term $f(\mathbf{r},t)$ in the LMF regime, the approach adopted here will exclude LMF data and derive a broadband image valid only above the LMF regime.

E. Source summary

For frequencies above the LMF regime, a source model can be constructed based on the following assumptions.

(1) Assumption 1. The source excitation is due to individual bubbles radiating at their Minnaert frequency as they reconfigure into spheroidal shapes.
(2) Assumption 2. Bubble forcing is oscillatory in character and will time average to zero.
(3) Assumption 3. Bubbles contribute independently to the sound field, so that, as sources, they are statistically independent in space and frequency.
(4) Assumption 4. The actively radiating bubbles form an extended source that is distributed horizontally but, for spilling breakers, is acoustically thin in the vertical direction, creating essentially a 2D source.

A physical picture of this model is shown in Fig. 1.

III. PARAMETRIC IMAGING

The standard imaging approach is to deploy a sensor array, designed with adequate spatial resolution at emitted source frequencies, opposite the source region and beam form across the source. This approach would not have yielded satisfactory solutions in our situation because (1) the sensor constellation was sparse at frequencies where near-field resolution was reasonable, and (2) the near-field resolu-

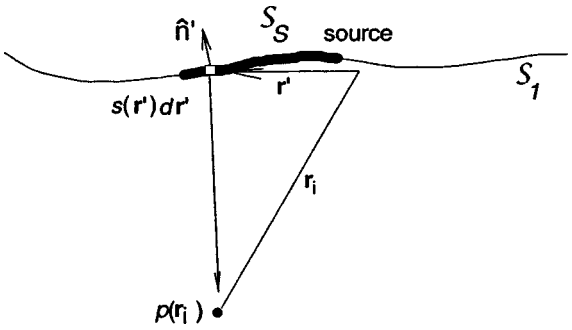


FIG. 2. Radiation from a source element.

tion was inadequate at frequencies where the sensor constellation was not sparse. An alternate approach is to model the source with a parametric function possessing at best a small number of free parameters. This approach is outlined in Sec. III.

A. Narrowband maximum-likelihood solution

Consider the acoustic field pressure at field point $\mathbf{r}$ due to a frequency-domain excitation $s(\omega;\mathbf{r}')d\mathbf{r}'$ from a source element at $\mathbf{r}'$ as shown in Fig. 2. The source element is assumed to be located on a boundary $\mathcal{S}_1$ where homogeneous Dirichlet conditions hold everywhere except over the region $\mathcal{S}_s$.

Each elemental source excitation is a random function: under the assumptions of Sec. II E,

$$\langle s(\omega;\mathbf{r})\rangle = 0, \tag{2}$$

$$\langle s^*(\omega;\mathbf{r})s(\omega';\mathbf{r}')\rangle = \Gamma_{ss}(\mathbf{r},\mathbf{r}';\omega)\,\delta(\omega-\omega'), \tag{3}$$

$$\Gamma_{ss}(\mathbf{r},\mathbf{r}';\omega) = \Psi(\mathbf{r};\omega)\,\delta(\mathbf{r}-\mathbf{r}'). \tag{4}$$

The notation $\langle\cdot\rangle$ denotes ensemble averaging, or, invoking ergodicity, time averaging: Eq. (2) therefore asserts that the excitation has zero mean, and Eq. (3) that the elemental excitations, even from the same source element, are statistically uncorrelated at different frequencies. $\Gamma_{ss}(\mathbf{r},\mathbf{r}';\omega)$ is the source mutual spectral density (MSD), and Eq. (4) indicates that the source mechanism is spatially uncorrelated within the source. The frequency-dependent intensity profile $\Psi(\mathbf{r};\omega)$ defines the (narrowband) acoustic shape of the breaking wave.

The total acoustic field at the field point at frequency ω is the sum of contributions from each elemental source. For an extended source, this ''sum'' takes the form of an integral over the 2D source region. It can be shown[11] that the cross-spectral density $\Gamma_{m,n}(\omega)$ between sensors m and n is

$$\Gamma_{m,n}(\omega) = \int_{\mathcal{S}_s} dS'\,\Psi(\mathbf{r}';\omega)K(\mathbf{r}';\mathbf{r}_m,\mathbf{r}_n;\omega), \tag{5}$$

where[45]

$$K(\mathbf{r}';\mathbf{r}_m,\mathbf{r}_n;\omega) = \left(\frac{\partial}{\partial n'}g^*(\mathbf{r}_m,\mathbf{r}';\omega)\right) \times\left(\frac{\partial}{\partial n'}g(\mathbf{r}_n,\mathbf{r}';\omega)\right) \tag{6}$$

is the ''kernel'' that propagates source statistics $\Psi(\mathbf{r}';\omega)$ to field cross statistics. The function $g(\mathbf{r}_m,\mathbf{r}';\omega)$ is the (deterministic) Green function for the problem. The derivative is with respect to the local normal (see Fig. 2). $\mathcal{S}_s$ is the source region. Note that the MSD between sensors m and n is also the m,nth element of a ''signal'' cross-spectral density matrix (CSDM) $\mathbf{Q}_{SS}(\omega;\Psi)$. This is then a fundamental result: each element of the signal CSDM is in fact a particular functional of the intensity profile.

The signal received at a field sensor is the sum of the random contributions from the many differential areal elements in the source region. This suggests that the distribution of the received signal is, according to the central limit theorem, well approximated by a Gaussian distribution (at least for short time intervals.) In this situation, the sample narrowband CSDM $\mathbf{S}_\omega$ will have the complex Wishart distribution,[47]

$$f_{CW}(\mathbf{S}_\omega|\mathbf{Q}_{SS},N_{\text{samp}};\omega) = C|\mathbf{Q}_{SS}|^{-N_{\text{samp}}} \times\exp(-N_{\text{samp}}\,\text{tr}\,\mathbf{Q}_{SS}^{-1}\mathbf{S}_\omega),$$

where C contains terms unrelated to $\mathbf{Q}_{SS}(\omega;\Psi)$ (but possibly a function of $\mathbf{S}_\omega$). The narrowband log-likelihood function $L_\omega(\Psi)$ is then

$$L_\omega(\Psi)|\mathbf{S}_\omega,N_{\text{samp}};\omega) = C - N_{\text{samp}}\ln|\mathbf{Q}_{SS}(\omega;\Psi)| - N_{\text{samp}}\,\text{tr}\,\mathbf{Q}_{SS}^{-1}(\omega;\Psi)\mathbf{S}_\omega, \tag{7}$$

where C is a different constant, although still not a function of $\mathbf{Q}_{SS}$.

This provides a ''goodness-of-fit'' function for Ψ based directly on a distributional argument: the narrowband log likelihood will reach a maximum when a model CSDM $\tilde{\mathbf{Q}}_{SS}(\omega;\hat{\Psi})$ based on an estimate $\hat{\Psi}$ of the true intensity profile Ψ attains the best fit to the available data CSDM $\mathbf{S}_\omega$. The profile $\hat{\Psi}_{\text{ML}}$ that maximizes this goodness-of-fit measure is of course the maximum-likelihood (ML) profile.

This prescription is fundamentally an inverse problem: the available data, for a P-element array, consists of the P^2 cross- and self-terms in $\mathbf{S}_\omega$, but the intensity profile is a distribution with essentially an infinite number of degrees of freedom. This kind of problem cannot be solved uniquely without applying some a priori information to bound the solution space. One technique is to assume the source intensity profile belongs to a class of functions parameterized by a few ''shape parameters.'' We choose here a 2D Gaussian function,

$$W(u,v;\boldsymbol{\theta}) = \exp\left\{-\left(\frac{u^2}{w^2}+\frac{v^2}{l^2}\right)\right\}, \tag{8}$$

where $(w,\,l)$ are, following Wu,[48] the cross-wind dimension and down-wind dimension scales, respectively, and where (u,v) are transformed source plane coordinates,

$$u = (x-x_0)\cos\phi + (y-y_0)\sin\phi,$$

$$v = -(x-x_0)\sin\phi + (y-y_0)\cos\phi.$$

Here (x_0,y_0) are the x-axis and y-axis centroids (first moments) of the profile. The angle ϕ is the angle between the u axis and the array x axis: given the array bearing, ϕ relates

the orientation of the breaker long axis to true geographical coordinates. The parameter vector $\boldsymbol{\theta}=[x_0,y_0,w,l,\phi]^T$. The narrowband imaging problem is then reduced to computing the ML estimator $\hat{\boldsymbol{\theta}}_{\mathrm{ML}}$: the resulting parametric narrowband ML image is the Gaussian profile of Eq. (8) with $\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\mathrm{ML}}$.

B. Broadband solution

As mentioned earlier, the conventional (narrowband) imaging problem suffers from (1) array sparseness (hence spatial aliasing) at high frequencies and (2) loss of resolution at low frequencies. In the parametric problem, high frequency performance results in decent estimates of profile shape $\{w,l,\phi\}$ but not of location, while low frequency performance results in decent estimates of location $\{x_0,y_0\}$ but not of profile shape. A compromise solution is to seek a broadband image consistent with the low and high frequency data. A summary of this technique is given below; details are provided elsewhere.[11,49,50]

Broadband imaging applies here only when the following additional assumption is in force.

Assumption 5. The shape of the bubble-''creation'' size distribution (as opposed to the size distribution of mature or ''adult'' subsurface bubbles) is the same throughout the entrainment region (although the rate of bubble creation may vary.)

There is no direct evidence that assumption 5 is valid: in fact, there is very little evidence whatsoever regarding the dependence of the bubble size distribution $n(R;\mathbf{r},t)$ on space and time within the bore. Kerman[51] has advanced a theory regarding the evolution of $n(R;\mathbf{r},t)$, speculating that the distribution of bubbles, following principles of fracture mechanics, depends mostly on the size of each bubble, i.e., a property with strong local dependence.

Assumption 5 permits the source MSD to be written as a separable function:

$$\Gamma_{ss}(\mathbf{r},\mathbf{r}';\omega)=S_{ss}(\omega)W(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}'), \tag{9}$$

the product of a frequency-dependent autsopectral density source strength $S_{ss}(\omega)$ and a broadband acoustic shape function $W(\mathbf{r})$. This has the intuitive appeal that the fluid is acoustically quiet (at all frequencies) outside the region of support of $W(\mathbf{r})$, which corresponds to the notion of localized bubble production. Others[8] have also invoked this assumption in the theoretical modeling of breaking waves.

Inserting Eq. (9) into Eq. (5) yields

$$\Gamma_{i,j}(\omega)=S_{ss}(\omega)\int_{\mathcal{S}_s} dS'\,W(\mathbf{r}')K(\mathbf{r}';\mathbf{r}_i,\mathbf{r}_j;\omega). \tag{10}$$

The source spectral strength is parameterized as

$$S_{ss}(\omega)=S_0 e^{20+\nu(\omega)}, \tag{11}$$

where $S_0=1\ \mu\mathrm{Pa}^2/\mathrm{Hz}$ at 1 m, as required for dimensional accuracy. If the function $\nu(\omega)$ were known, Eq. (10) would be a Fredholm integral equation of the first kind for the shape function. However, this term is not known, so the problem is a bit more complicated than that posed by Eq. (5).

Assume now that measurements are made at N_f different frequencies. The values $\nu_1=\nu(\omega_1),\nu_2,\ldots,\nu_{N_f}$ represent nuisance parameters that must be determined but are not of direct interest (in the imaging problem, but see Sec. VI D). Under assumption 3, the information at each of these different frequencies is independent, so the total ''broadband density'' is formally

$$f_{CW}(\mathbf{S}_{\omega_1},\mathbf{S}_{\omega_2},\ldots,\mathbf{S}_{\omega_{N_f}})=\prod_{i=1}^{N_f} f_{CW}(\mathbf{S}_{\omega_i}), \tag{12}$$

and the ''broadband log likelihood'' is now

$$L_{\mathrm{BB}}(\boldsymbol{\theta}|\boldsymbol{\nu})=\sum_{i=1}^{N_f} L_{\omega_i}(\mathbf{Q}_{SS}(\boldsymbol{\theta}|\boldsymbol{\nu})|\mathbf{S}_{\omega_i},N_{\mathrm{samp}};\omega_i), \tag{13}$$

where it has been assumed that N_{samp} samples are collected at each frequency and $\boldsymbol{\nu}=[\nu_1,\nu_2,\ldots,\nu_{N_f}]^T$ is the vector of source strength nuisance parameters. Then, the ''maximum-likelihood image'' is the image $W_{\mathrm{ML}}(u,v|\hat{\boldsymbol{\theta}}_{\mathrm{ML}})$ based on the parameter vector $\hat{\boldsymbol{\theta}}_{\mathrm{ML}}$ which in turn maximizes the broadband log-likelihood L_{BB}:

$$\hat{\boldsymbol{\theta}}_{\mathrm{ML}}=\arg\max_{\boldsymbol{\theta},\boldsymbol{\nu}} L_{\mathrm{BB}}.$$

This is in general a constrained nonlinear optimization problem, where the constraints require that the shape dimensions (w, l) be positive.

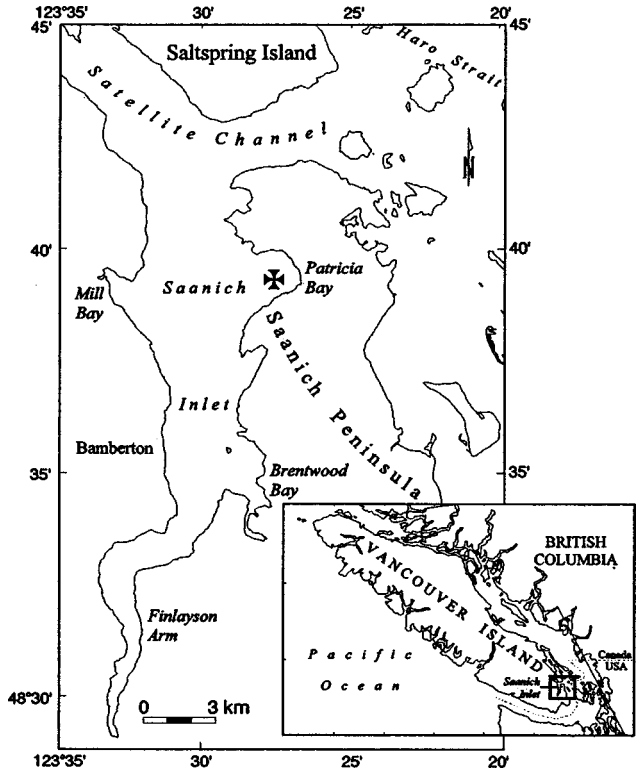


FIG. 3. Experiment coastal location, Patricia Bay (the ''cross''), Saanich Inlet, Vancouver Island, Canada.

IV. THE EXPERIMENT

The data were acquired by the instrument deployed in Patricia Bay, Saanich Inlet, British Columbia (see Fig. 3) for 57 days in the spring of 1995. The deployment site has a fetch of about 5 km to the west and a water depth of about

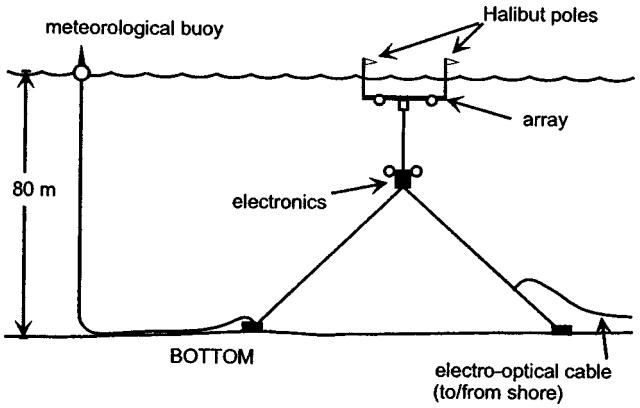


FIG. 4. Two-point bottom-mounted mooring scheme.

80 m. A Coastal Climate Systems meteological buoy was moored nearby to provide near-surface environmental parameters.

A diagram of the actual mooring scheme is shown in Fig. 4. The acoustic array was positively buoyant and bottom moored so that the sensor plane had a nominal depth of about 3 m. The local sound speed was measured by a conductivity-temperature-depth (CTD) mounted on the array frame. The system was cabled to shore. In addition, 6 m ''halibut'' poles were attached to the end of each arm of the array in order to aid in recovery and to warn off local boaters.

Surface drift currents had a considerable effect upon the array: on-board tilt sensors recorded angles up to $\pm 8°$. Strumming due to aerodynamic flow around the halibut poles and hydrodynamic flow around the array structure itself introduced acoustic contamination. In addition, wave orbital motion caused the array structure to flex, causing further radiated and structure-borne mechanical noise.

The sensor constellation, Fig. 5, was originally designed for a direction-finding experiment. The arm-end sensors were half wavelength spaced at the system's highest frequency of operation, but the constellation, due to cost and complexity, was sparse along each arm and in each quadrant.

A storm event on April 13, 1995 produced data with high enough quality for imaging. The vector-averaged wind speed during this event is shown in Fig. 6. The breaking events selected for imaging were acquired at approximately 15:00–16:30 Pacific Daylight Time (PDT), when the wind speed had been relatively constant for about 5 h. The air–sea temperature (Fig. 7) shows that this was a warm front, and the wind direction shown in Fig. 8 was quite consistent at about 260°; this corresponds to a fetch of about 5 km. The wave field was essentially fully developed for these acquisitions, and a hindcast[52] predicts the period and the significant wave height of the dominant spectral component at 2.34 s and 0.38 m, respectively.

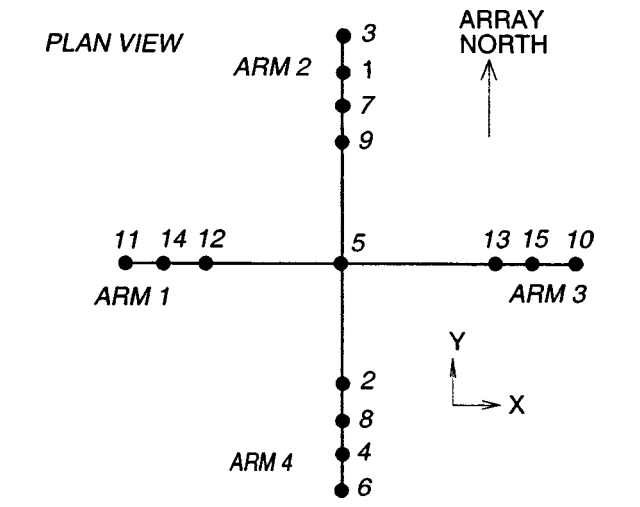


FIG. 5. Sensor constellation. The sensors are located in the horizontal array plane. The interelement distance in the arm-end clusters is a half wavelength at the highest frequency of operation.

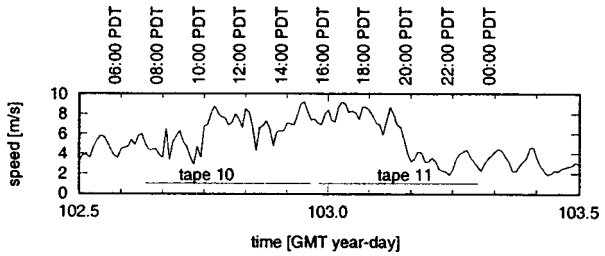


FIG. 6. Vector-averaged wind speed (measured at 3 m height).

Figure 9 shows a multichannel broadband acoustic signature of a wave breaking over the array. Judging solely by power levels, the wave appears to have broken somewhere between the center sensor and arms 3 (channels 2, 8, 4, and 6) and 4 (channels 13, 15, and 10.) The dipole character of the acoustic radiation attenuates the contribution to arms 1 and 2 to the extent that the broadband signal is scarcely noticeable above the background noise.

V. SIGNAL PROCESSING

A. CSDM estimation

The calculation of quality CSDM estimators requires some manual intervention. For example, the power signature from the center hydrophone for wave ''W3'' is shown in Fig. 10. The data have been filtered with a fifth-order Butterworth high pass filter with corner frequency at 160 Hz to remove a strum component and then converted to units of μPa, squared, and averaged over consecutive 11 ms blocks. The record shows the onset of the acoustic signal at about 5.5 s (arbitrary start), peaking at about 6.3 s, and dying away by about 8 s. Several spikes due to mooring noise can also be seen. The background ambient level can be estimated by the signal level prior to the event onset; this is about 97 dB. The

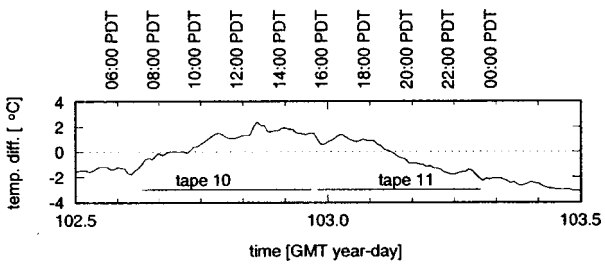


FIG. 7. Air–sea temperature difference.

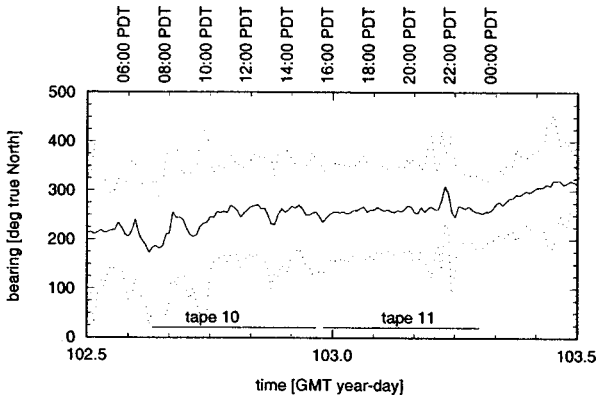


FIG. 8. Wind direction. The solid line is the bearing averaged over 10 min intervals, and the dotted lines represent ±1 standard deviation.

breaker signal has bursts to about 109 dB, so this event had, at sensor 5, a maximum signal to noise ratio (SNR) of about 10 dB.

Figure 10 is also annotated with lines indicating the selection of ambient and ambient+signal (labeled signal for brevity). The selection process was done by eye. The radiated signature is clearly a nonstationary signal; however, an examination of the data showed that the signal could be considered quasistationary over time scales of about a half second. Care was taken to avoid segments in which the power appeared to take an obvious jump, or contained ''bursts,'' probably mooring noise.

Sample CSDM computation used a fast Fourier transform (FFT)-based multiple taper technique;[46,53,54] this technique produces (1) tapers that are optimally concentrated in frequency and (2) auto-and cross spectral estimates that are efficient. The latter feature was particularly important for the short-time record quasistationary data available here. The

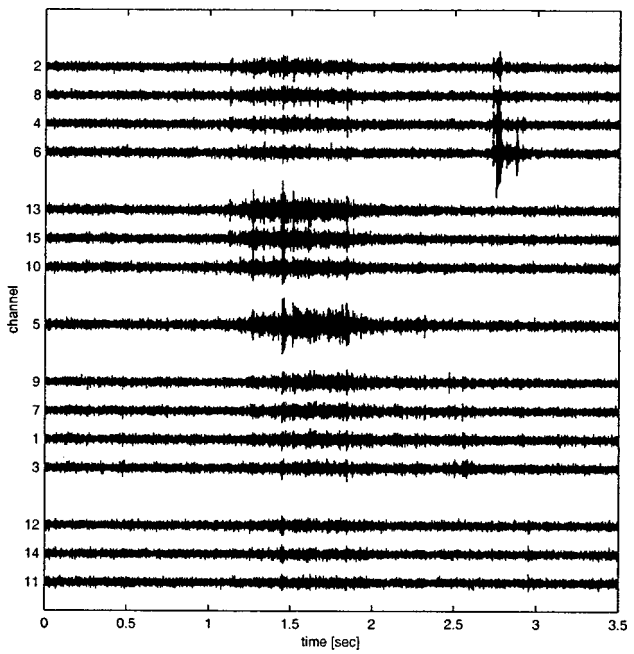


FIG. 9. Broadband acoustic signals per channel for a typical wave. The channels are grouped according to arms. The signals have been prefiltered to remove low frequency vibrational contamination; a structure-borne transient is evident on channels 2, 8, 4, and 6 at the top.

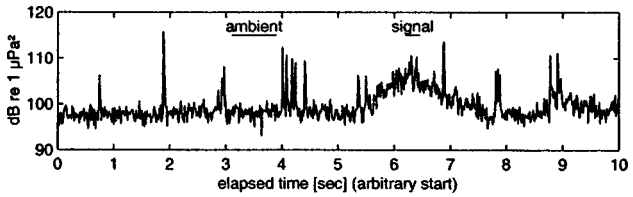


FIG. 10. Wave W3 short-time power: 15:24:01 PDT.

taper family depends on the product NW of the sample size N and the resolution bandwidth W. The parameters chosen for this work are shown in Table I. The time–bandwidth product $NW=3$ produces a resolution bandwidth of about 258 Hz, which is about 4 discrete Fourier transform (DFT) bins. Six bands, with the center frequencies given in Table I, provided roughly nonoverlapping coverage of the frequency band [474 Hz, 2024 Hz] (mutual independence is required for Eq. (12) to hold). Thus, the ''signal'' data segment shown in Fig. 10 would yield six sample 15×15 CSDMs (one at each frequency).

(The data were often contaminated by a strong strum component, particularly on sensors at the arm ends. This manifested itself most strongly in the lowest frequency bands. For acoustic events contaminated by this noise, imaging solutions using all six bands did not converge properly, but successful convergence could usually be achieved when some of the lowest bands were ignored. Imaging solutions presented here always used three or more bands.)

B. Model CSDM computation

Several key steps are required to compute the model CSDMs $\mathbf{Q}_{SS}(\omega;\Psi)$: a propagation model must be chosen so as to identify the appropriate Green function $g(\mathbf{r}_i,\mathbf{r}';\omega)$, and the integral of Eq. (5) must be evaluated. These are discussed below.[50]

Propagation model: Neither the surface geometry nor the sound speed profile $c(\mathbf{r})$ was measured during the experiment. However, at the frequencies of interest here, it is an adequate approximation to assume the surface $\mathcal{S}_1$ to be the $z=0$ plane, and to assign the ocean halfspace $\mathcal{V}$ a constant sound speed c_0. (Supporting arguments are presented in the Appendix.) These assumptions allow the use of a greatly simplified Dirichlet–Green function[11] that uses simple spherical spreading from a vertical dipole.

Numerical integration: To evaluate Eq. (10) for various trial solutions $W(\mathbf{r})$, the range of integration was limited to an overhead ''field of view,'' with surface sources outside this region considered part of the background noise. The

TABLE I. CSDM estimation parameters.

Parameter	Value
Sampling frequency (Hz)	5512.5
N	64
NW	3
Bandwidth (W) (Hz)	258
Center frequencies (Hz)	603, 861, 1120, 1378, 1637, 1895

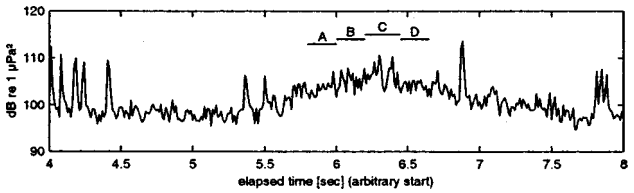


FIG. 11. Short-time segmentation of wave W3. End times, in relative seconds: A (5.8, 6.0); B (6.0, 6.2); C (6.2, 6.45); D (6.45, 6.65). The time axis is referenced to 15:24:01 PDT.

Dirichlet–Green function allows arbitrary $\{\mathbf{r}_i\}$ (i.e., the array platform orientation was specified by empirical pitch, roll, and yaw angles.)

The integral, Eq. (10), was approximated by the sum of 2D subintegrals over the field of view: each 2D subintegral was estimated by a standard cubature technique.[55] Each element of $\mathbf{Q}_{SS}(\omega;\Psi)$ required one evaluation of Eq. (10).

VI. OBSERVATIONS

Thirty distinct and well-defined acoustic events were identified in the data and processed. Twenty-one yielded convergent parametric image solutions. In Sec. VI A we discuss these images. These images, all associated with a fully developed fetch-limited sea and a wind speed of about 7 ms^{-1}, are used in Secs. VI B and VI C to infer two simple oceanographic parameters. The source strength, only a nuisance parameter for the imaging problem, is analyzed separately in Sec. VI D.

A. Parametric images

Waves W3 and W5 were chosen for sequential imaging analysis to validate the performance of the broadband imaging technique. These waves were chosen because they allowed full broadband solutions (i.e., all measurement bands contributed to the solutions and should therefore yield more stable estimates) and had significant SNR over their duration. In each case, the acoustic record was partitioned more or less ''by eye'' into segments approximately 0.2–0.4 s long, and a solution computed for each segment.

Wave W3 contained four usable segments, A–D, that yielded stable solutions. The segmentation plan is shown in Fig. 11. A parametric solution over each segment produces the solutions in Table II. Figure 12 shows a visualization of the shape solutions and includes the wind direction, corrected to the array field-of-view frame of reference. The radiant region is clearly seen to be moving almost directly downwind.

TABLE II. Maximum-likelihood parameter evolution for wave W3. The second and third columns provide the source centroids at the nominal fluid surface. Columns 4 and 5 provide the major and minor axis scales. Column 6 provides the orientation of the wave's major axis relative to the array's frame of reference.

Wave	x_0 (m)	y_0 (m)	w (m)	l (m)	ϕ (deg)
3A	−2.62	1.39	1.74	1.30	64.2
3B	−1.95	1.46	1.61	1.11	83.8
3C	−1.43	1.33	1.80	1.23	87.1
3D	−0.77	1.30	1.92	1.03	80.2

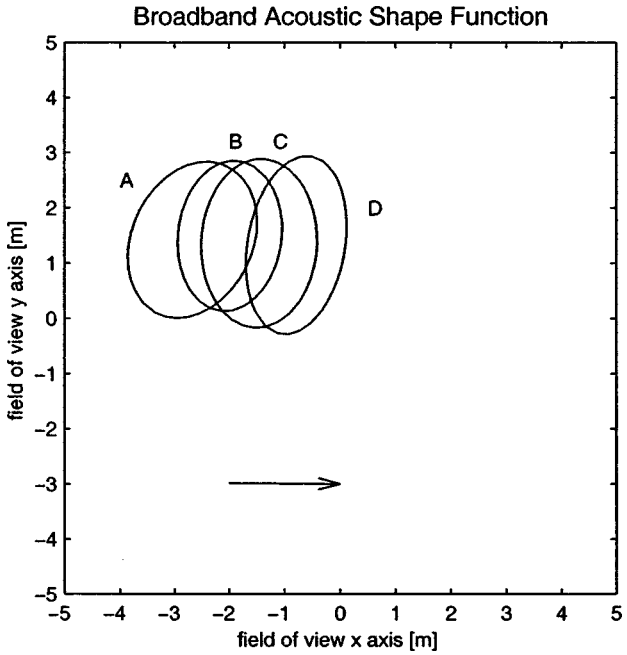


FIG. 12. Sequence of solutions for wave W3, the array's frame of reference. Solution evolves from A to D: only the half-power contour (intensity = 0.5) solution is shown. The approximate wind direction is also indicated by the arrow.

Wave W5 contained five segments, A–E, that yielded stable solutions. The segmentation plan is shown in Fig. 13. The ML shape parameter estimators are shown in Table III. Figure 14 shows a solution visualization.

In both cases, the radiant region is clearly seen to be moving almost directly downwind. Furthermore, it appears that the radiant region is wider along the cross-wind direction, which conforms to visual observations that active breakers have a wider cross-wind dimension. These observations suggest that the technique forms accurate parametric ''images'' of breaking waves.

The ensemble of all convergent image solutions obtained during this storm are shown in Fig. 15; solution parameters are shown in Table IV. Most solutions are more or less aligned (see Sec. VI C). Most of the solutions are also quite elliptical. Several solutions seem to have outlier parameters, in that they are either unreasonably large or quite small. These features are discussed further in Sec. VII.

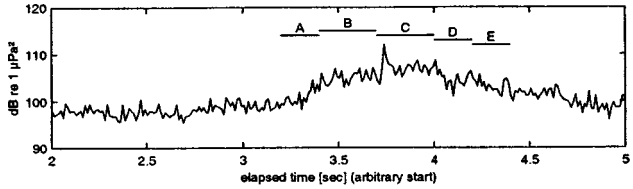


FIG. 13. Short-time segmentation of wave W5. End times, in relative seconds: A (3.2, 3.4); B (3.4, 3.7); C (3.7, 4.0); D (4.0, 4.2); E (4.2, 4.4). The time axis is referenced to 15:25:16 PDT.

TABLE III. Maximum-likelihood parameter evolution for wave W5. The second and third columns provide the source centroids at the nominal fluid surface. Columns 4 and 5 provide the major and minor axis scales. Column 6 provides the orientation of the wave's major axis relative to the array's frame of reference.

Wave	x_0 (m)	y_0 (m)	w (m)	l (m)	ϕ (deg)
5A	−1.00	1.50	1.84	1.00	90.4
5B	−1.17	1.54	1.77	1.20	92.2
5C	0.01	1.55	1.64	1.22	59.5
5D	0.67	1.44	1.57	1.32	72.1
5E	1.42	1.30	1.70	1.22	130.0

B. Wind speed correspondence

The mean speed of each breaker can be computed by measuring the Euclidean distance between the centroids of the first and last shape functions and dividing by the total segmentation duration. Using this formula, the mean speed $\overline{c_{br}}$ of the wave W3 breaker is 2.85 ms^{-1}, and the mean speed of the wave W5 breaker is 2.43 ms^{-1}. (The mean speed of wave W5 would be greater than 3 ms^{-1} if measured over segments B–E.) In contrast, the phase speed of the dominant wind waves,[56]

$$c_p = \sqrt{g/k},$$

is 3.65 ms^{-1}, where k is the wave number calculated from the hindcast. This is consistent with previous acoustics measurements[2,38] of mean breaking speed which showed that $\overline{c_{br}} \approx 45\% - 75\%$ of c_p.

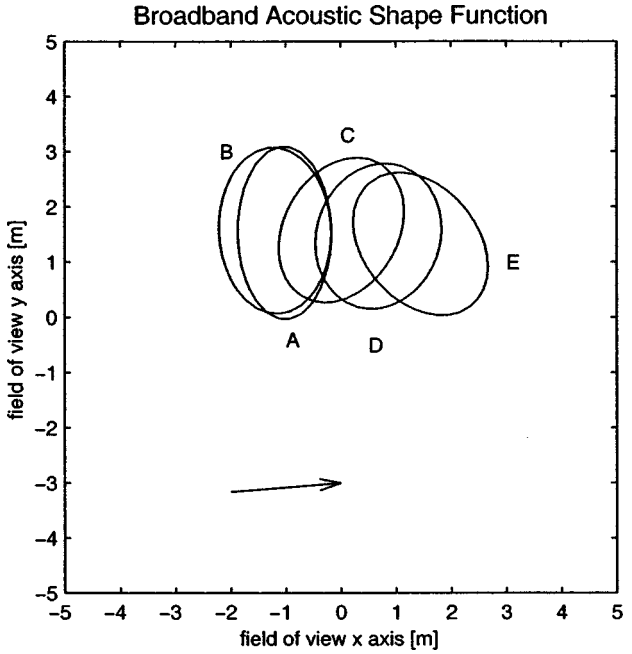


FIG. 14. Sequence of solutions for wave W5, the array's frame of reference. Solution evolves from A to E: only the half-power contour (intensity = 0.5) solution is shown. The approximate wind direction is also indicated by the arrow.

TABLE IV. Maximum-likelihood parameters for the broadband wave shapes. x_0 and y_0 are the source centroids at the nominal fluid surface. w and l are the major (along the crest) and minor (across the crest) axis scales. The orientation here is the angle of the wave's major axis relative to true North (TN). Only those parameters for convergent solutions are shown.

Wave	x_0 (m)	y_0 (m)	w (m)	l (m)	Orientation (deg TN)
2	0.53	−0.85	1.99	0.55	184.4
3	−1.43	1.33	1.80	1.23	172.5
4	−3.67	−2.61	3.81	1.02	184.5
5	−1.17	−1.54	1.77	1.20	182.7
6	−3.07	−0.44	2.90	2.28	151.8
7	−0.31	0.94	1.56	0.78	179.3
8	1.24	−2.40	2.70	0.91	203.9
9	−0.70	0.20	1.19	0.46	304.8
11	0.00	−1.20	1.74	0.49	171.8
13	0.08	−0.47	2.05	0.67	135.7
14	1.37	−2.85	3.42	1.44	201.6
17	2.15	−1.62	3.28	0.06	175.9
18	0.05	0.27	2.04	0.87	206.8
20	1.27	−2.59	3.64	0.57	178.1
21	−2.51	0.45	1.22	0.73	206.5
22	0.56	0.86	2.16	1.27	173.4
23	0.56	−0.93	1.35	0.58	216.8
24	1.32	2.61	9.47	1.44	207.7
27	−0.5	−0.43	1.49	0.62	135.5
28	1.06	−2.14	3.60	3.00	217.2
30	0.74	−2.30	5.60	1.28	198.5

C. Wind direction correspondence

Solution trajectory provides an estimate of wind direction. An alternative technique is to assume that the breaking region is generally elongated along the crests of waves: the mean orientation of the solutions (i.e., the angle of the major axis) should provide a second indicator of wind direction.

The wind direction and speed are seen from Figs. 6 and 8 to be roughly stationary during this period. The orientation solutions for the 21 convergent wave solutions therefore comprise samples from approximately the same distribution. These samples are shown in Fig. 16 as triangles, plotted versus time: the estimation errors on each point are too small to be seen on this scale. Also plotted are the vector-averaged

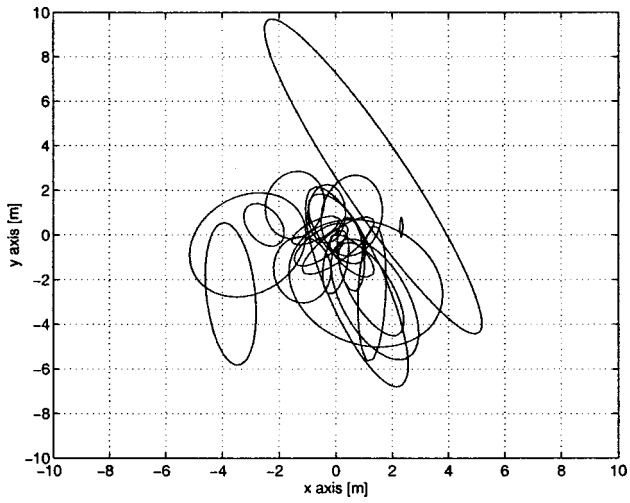


FIG. 15. All convergent imaging solutions from the storm event. Only half-power contours (intensity = 0.5) are shown.

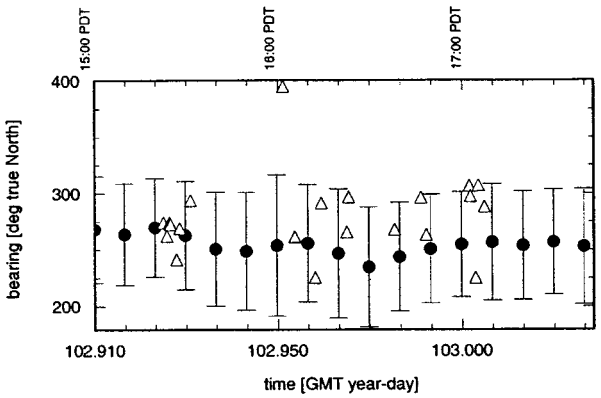


FIG. 16. Sequence of inferred wind directions. The triangles are derived from the orientation estimators $\hat{\phi}$ of the wave shape function, and the circles are from the nearby buoy. The vertical error bars on the circles represent 1 standard error.

wind direction estimates and standard errors generated every 10 min by the nearby meteorological buoy.

Inspection of Fig. 16 shows that there is clearly one significant outlier at about 16:02. This is the solution for wave W9 which appeared somewhat unstable at the lower frequencies. If this point is deleted from the data set, the mean orientation and bootstrap[57] standard error using the remaining solutions are $272.3° \pm 4.6°$. This is well within the standard error of the surface wind direction measurements.

D. Source strength

Although the source strength is a nuisance parameter for the imaging problem, this parameter is a fundamental input for the radiated intensity of breaking waves. Using the same argument as that presented in Sec. VI C, the sequence of source strength solutions may be considered samples associated with a single wind speed: from Fig. 6, this is approximately 7–8 $\mathrm{m\,s^{-1}}$. All available solutions are plotted together in Fig. 17: solution points at different frequencies for the same wave are connected by simple line segments. The most prominent feature of Fig. 17 is the appearance that the frequency dependence of the source strength parameters, over the available frequency band, is predominantly linear in

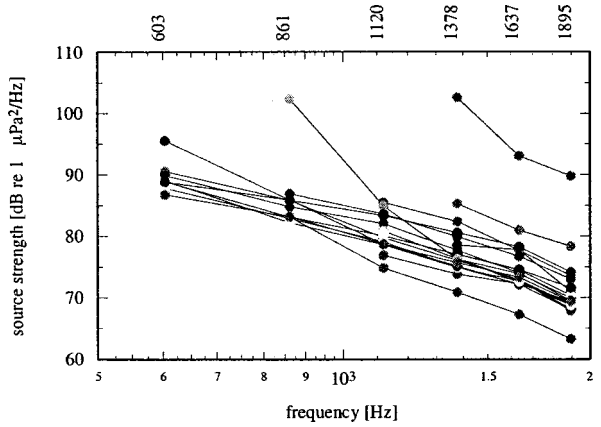


FIG. 17. Source strength parameters, 21 waves with 3 or more solutions. Solutions at different frequencies for the same wave are connected by line segments.

log frequency. (The deviations from this linearity at the lower frequencies are thought to be related to poor convergence problems in the algorithm due to noisy data.) This suggests that the source strength parameter be modeled as

$$\nu(f) = b_0 + b_1 \ln f, \tag{14}$$

where b_0 is a random process that will vary from wave to wave since each varies in size and breaking intensity. (One might suppose the distribution of b_0 to be strongly wind dependent; unfortunately, we only have data at one wind speed.) The parameter b_1, however, might be considered a more universal function, so it is reasonable to combine all available solutions for the estimation of b_1.

A linear regression was performed on each available solution against a model which was first order in log frequency. This produced 21 ''raw'' values of $\hat{b}_0$ and $\hat{b}_1$. The set of raw estimates $\{\hat{b}_1\}$ contained evidence of a few outliers, so the mean of this set was computed using

$$\overline{b_1} = \mathrm{median}(\{\hat{b}_1\}). \tag{15}$$

The standard error of the estimator $\overline{b_1}$ was again approximated with the bootstrap standard error (se): the value of $\overline{b_1}$ was estimated at $\overline{b_1} \pm \widehat{se} = -4.55 \pm 0.47$.

As mentioned above, the distribution of b_0 should be a function of wind speed. If the source strength curves are assumed to have a slope of $\overline{b_1}$, then raw estimates of b_0 can be made for each available curve. This sample set shows evidence of outliers, but there are not enough raw values to sketch out the shape of this distribution. Thus, the b_0 distribution is represented by a single parameter, here chosen to be the median: median $(\overline{b_0})|(b_1 = \overline{b_1}) = 30.4$.

How does this solution for the source strength of an individual breaker relate to background ambient levels? According to Urick,[58] the surface component of the ambient sound spectrum from about 500 Hz to about 50 kHz has a shape known as the Knudsen spectrum. The slope of the Knudsen spectrum is nearly constant at about -6 dB per octave, or about -20 dB per decade. This surface component is generally attributed to natural sea–surface processes, and, in particular, breaking waves. There must therefore be a connection between the spectral shape of the sound radiated from individual breaking waves and the Knudsen spectrum.

Several investigators have in fact had good success in modeling the shape of the Knudsen ambient sound curves by considering the ambient curve to have a shape similar to the radiated spectrum of individual breaking waves and then assuming that the latter is due to the incoherent contributions from a population of damped ringing bubbles within the breaker. Medwin and Daniel[4] and Medwin and Beaky[20] placed a hydrophone under small breaking waves in laboratory tanks and measured acoustic autospectra with an approximate Knudsen slope from about 500 Hz to about 50 kHz. Loewen and Melville[3] used Monte Carlo simulations to predict the radiated acoustic field from a region of active bubbles and also achieved a curve with similar slope and amplitude from about 400 Hz to above 50 kHz.

The source levels estimated in this section are related *by construction* to the radiated spectra of individual breaking

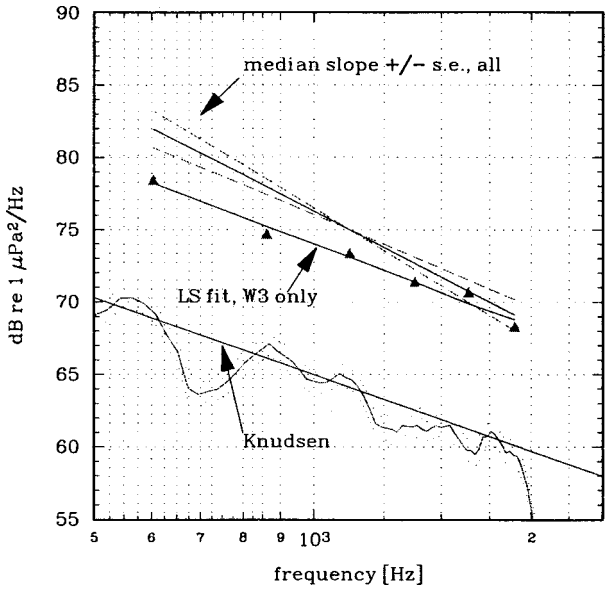


FIG. 18. Relation of source levels to actual and Knudsen ambient sound levels compared via a fictitious sensor at 3 m depth. The triangles were computed using the actual solutions from W3. The two curves at the bottom are the Knudsen level and a short-time estimate with 90% confidence intervals of the actual ambient autospectrum. Two regression lines are shown: the one labeled LS fit, W3 only used the slope $\hat{b}_1$ from wave W3. The other, labeled median slope, used $\overline{b_1}$, and also shows ± 1 se curves.

waves: the two are connected by the dipole propagation model. If the radiated spectra have shapes similar to that of the ambient autospectrum, then the mean source level and the ambient autospectrum should be similarly connected by the dipole propagation model. This can be checked directly.

An estimate of the one-sided ambient autospectrum is shown in Fig. 18. The antialias filter cutoff is evident near 2000 Hz. Below about 500 Hz, the ambient level plateaus at about 70 dB. This plateau is a well-known feature of ambient sound spectra[58] and marks the region where manmade (shipping) noise overwhelms the Knudsen shape.

A Knudsen curve fixed at a level of 65 dB at 1000 Hz is also shown in Fig. 18: it is not extended below about 500 Hz as it is not clear whether it remains a valid model below this point. The ambient autospectrum follows this Knudsen curve in general, although there is evidence of some statistically significant fine structure.

Three model radiated spectra are shown in Fig. 18; all were computed for a fictitious sensor located 3 m directly below the centroid of a patch of acoustically active fluid with the parametric broadband shape function of wave W3. The source region was energized with (1) wave W3 source strengths (at the six analysis frequencies), (2) a regression (least squares) fit of Eq. (14) onto wave W3 source strength parameters alone, and (3) the mean source strength model involving $\overline{b_0}$ and $\overline{b_1}$. Although the fictitious received levels are louder than the background ambient levels (due to the proximity of the wave to the sensor), there is good agreement in slope among them. The solution based on W3 parameters fits very well, and the median solution fits within the estimated standard error.

This supports the idea that the mean source strength parameterization, when modulated by a dipole propagation model, yields a mean radiated spectrum which has the same spectral shape as the surface component of the background ambient noise over the frequency regime of 400 Hz–2 kHz.

Another feature of interest is the low-frequency cutoff in source level: in principle, there ought not be individual bubbles with resonance frequencies below some appropriate low frequency. The analysis here does not extend down into the regime but, over the available two octaves, the source level shows no sign whatsoever of a low-frequency rolloff or cutoff.[7]

VII. CONCLUDING REMARKS

The acoustically active region of breaking fetch-limited waves can be imaged at midfrequencies just above the peak of the radiated spectrum. In this frequency regime, the radiated sound is believed to be the incoherent contribution of individual newly formed bubbles ringing at their resonant frequencies in the turbulent two-phase bore.

Far-field resolution limitations and spatial aliasing were overcome by positioning the array near the surface (i.e., a near-field array), assuming the active region has an identical intensity over frequency, assigning a parametric model to the wave, and forming a broadband solution. The radiated signal is inherently a stochastic process, so the imaging problem is formulated in a probabilistic framework, thereby yielding, under common distributional arguments, a maximum-likelihood image.

These parametric images appear to be accurate representations of the white caps in that their mean velocity, trajectory, and orientation equate closely to the corresponding oceanographic parameters of the wave and wind field. This correspondence could be reversed, allowing estimation of the wave and wind field from parametric images of breaking ocean waves.

Parametric imaging has the advantage of limiting the dimension of the solution space (here, to the five elements of $\boldsymbol{\theta}$) but also brings disadvantages: unmodeled physical features remain unobserved. In particular, the Gaussian profile cannot accurately represent a multiple source region, a region that splits up, or multiple regions that coalesce. This insensitivity to more complex source shapes may in part be responsible for the unusually large or small parametric images depicted in Fig. 15.

The Gaussian profile may also not accurately model the distribution of acoustically active bubbles. If, for example, an advected layer of mature bubbles masks the downward radiation of all newly created bubbles in the wake of bore, then the apparent image of the source region might be better modeled by a cresent or arc, which would represent only that part of the bubble-creation region along the forward face of the bore. However, a crescent or arc model would be equally poor in modeling an extended surface patch of active bubbles. Based on the conjecture that, at least at the lower frequencies, the advected plume is acoustically thin and that the bubble-creation region itself is not concentrated at the leading edge of the bore but in fact spatially extended rearward by advection, the Gaussian profile seemed a practical compromise.

The validity of assumption 4 (an acoustically thin high void-fraction plume) certainly cannot be resolved with a parametric imager, and probably cannot be resolved by acoustics alone. If the plume is acoustically opaque, with only new bubbles at the bore face radiating downward, a true imager would only observe sources at the bore's forward edge, as discussed above. Alternatively, if bubbles under the plume are being stimulated into re-emission, they will be out of focus by an imager focused at the fluid surface. (Vertical translation of the actual fluid surface about the mean level will also cause this.) If the plume is acoustically translucent, scattering from the plume may yield a ''halo'' around the actual source region, giving a false representation of the source region size. These effects cannot be investigated with a parametric imager.

A better imaging solution would be to use an approach free of parametric constraints. As outlined in Sec. I, this is a challenging endeavor. However, better measurements and modeling of the propagation and masking properties of the spiller plume will also be needed before imaging solutions of either kind can be confidently linked to the physics of the bubble-creation region at the surface, and hence to studies of turbulence and gas transfer at the sea surface.

ACKNOWLEDGMENTS

The authors thank Oceanetic Measurement, Ltd. for their technical acumen, and the staff of the Acoustical Oceanographic Research Group at the Institute of Ocean Sciences for their support. One of the authors (R.K.A.) was supported by ONR Grant No. N00014-92-J-1688.

APPENDIX

Several model approximations have been made that require justification.

Flat surface model. The applicability of a flat surface approximation can be judged by utilizing the Rayleigh roughness parameter,[59] $\zeta = 2k\eta_{\mathrm{rms}}\cos\theta$ where k is the wave number $= \omega/c$, η_{rms} is the root mean square (rms) surface height, and θ is an ''angle of incidence.'' This has no meaning here, so assume the worst case value of $\theta = 0$.

The rms surface height η_{rms} for the storm event was calculated by hindcast to be 0.09 m; the corresponding Rayleigh roughness parameters at the low and high ends of the system operating range are shown in Table AI. The surface is considered rough if $\zeta \gg 1$ and smooth if $\zeta \ll 1$. Thus it is seen that the flat surface is a good approximation over the frequency band considered here.

Uniform sound speed. Measurements have suggested that the bubble plume and the continuum bubble layer modify the bulk sound speed by an amount $\Delta c \ll c_0$, and that this anomaly is virtually undetectable more than several significant waveheights below the surface.[39] This would imply that the sound speed in the fluid surrounding the array is, to a good approximation, c_0 and that this approximation remains valid above the array up to at least about 1 m below the surface.

Directly under the breaking wave, however, the high void fraction entrainment region supports a fluid ''inclusion'' with a significantly different sound speed over a vertical scale of about a wave height. The refractive properties of this inclusion have been thoroughly investigated for certain ideal geometries,[60] and it must be acknowledged that the source information will be modified by propagation out from such an inclusion. Unfortunately, the space–time subsurface bubble field above the array was not measured, and therefore propagation anomalies cannot be corrected. More accurate imaging must include this knowledge and use a more appropriate Green's function.

TABLE AI. Rayleigh roughness parameters for the fully developed fetch-limited wave field of the experiment.

Frequency (Hz)	Rayleigh parameter
500	0.4
2000	1.5

[1] M. R. Loewen and W. K. Melville, ''Microwave backscatter and acoustic radiation from breaking waves,'' J. Fluid Mech. **224**, 601–623 (1991).
[2] L. Ding and D. Farmer, ''Observations of breaking surface wave statistics,'' J. Phys. Oceanogr. **24**, 1368–1387 (1994).
[3] M. R. Loewen and W. K. Melville, ''A model of the sound generated by breaking waves,'' J. Acoust. Soc. Am. **90**, 2075–2080 (1991).
[4] H. Medwin and A. C. Daniel, ''Acoustical measurements of bubble production by spilling breakers,'' J. Acoust. Soc. Am. **88**, 408–412 (1990).
[5] P. A. Crowther and A. Hansla, ''The lifetimes, velocities and probable origin of sonic and ultrasonic noise sources in the ocean,'' in *Natural Physical Sources of Underwater Sound*, edited by B. R. Kerman (Kluwer Academic, Dordrecht, The Netherlands, 1993), pp. 379–392.
[6] C. L. Epifanio and M. J. Buckingham, ''Imaging the acoustical sources beneath individual breaking waves,'' J. Acoust. Soc. Am. **100**, 2841–2842 (1996).
[7] R. K. Andrew, ''Radiated autospectra from 160 Hz to 2000 Hz of individual breaking ocean waves,'' Canadian Acoustics **27** (3), 28–29 (1999).
[8] G. B. Deane, ''Sound generation and air entrainment by breaking waves in the surf zone,'' J. Acoust. Soc. Am. **102**, 2671–2689 (1997).
[9] R. D. Hollett, ''Observations of underwater sound at frequencies below 1500 Hz from breaking waves at sea,'' J. Acoust. Soc. Am. **95**, 165–170 (1994).
[10] C. Garrett, M. Li, and D. Farmer, ''The connection between bubble size spectra and energy dissipation rates in the upper ocean,'' J. Phys. Oceanogr. **30**, 2163–2171 (2000).
[11] R. K. Andrew and R. L. Kirlin, ''A broadband maximum likelihood imager for a class of extended space-time spearable sources,'' IEEE Trans. Signal Process. **48**, 1287–1294 (2000).
[12] T. G. Leighton, *The Acoustic Bubble* (Academic, London, 1994).
[13] M. M. Minnaert, ''On musical air bubbles and the sounds of running water,'' Philos. Mag. **16**, 235–248 (1933).
[14] M. Strasberg, ''Gas bubbles as sources of sound in liquids,'' J. Acoust. Soc. Am. **28**, 20–26 (1956).
[15] T. Berger, Hydrodynamic Properties of Air-Entraining Flows: A Study Using New Acoustic Techniques, Ph.D. thesis, University of California, San Diego, California, 1999.
[16] A. R. Kolaini and A. G. Goumilevski, ''Acoustic characterization of an adault bubble injected into a fully developed turbulent flow field,'' J. Acoust. Soc. Am. **101**, 218–226 (1997).
[17] G. E. Updegraff and V. C. Anderson, ''Bubble noise and wavelet spills recorded 1 m below the ocean surface,'' J. Acoust. Soc. Am. **89**, 2264–2279 (1991).
[18] G. E. Updegraff, Technical Report No. MPL-U-78/89, Marine Physical Laboratory and Scripps Institution of Oceanography, SIO reference No. 89-21.
[19] G. B. Deane and M. D. Stokes, ''Air entrainment and bubble size distribution in the surf zone,'' J. Phys. Oceanogr. **29**, 1393–1403 (1999).
[20] H. Medwin and M. M. Beaky, ''Bubble sources of the Knudsen sea noise spectra,'' J. Acoust. Soc. Am. **86**, 1124–1130 (1989).

[21] D. H. Peregrine, ''Mechanism of water-wave breaking,'' in *Breaking Waves*, IUTAM Symposium, Sydney, Australia, 1991, edited by M. L. Banner and R. H. J. Grimshaw (Springer, Berlin, 1992), pp. 39–53.
[22] T. B. Benjamin and J. E. Feir, ''The disintegration of wave trains in deep water. Part 1, Theory,'' J. Fluid Mech. **27**, 417–430 (1967).
[23] M. S. Longuet-Higgins, ''The instabilities of gravity waves of finite amplitude in deep water. I. Superharmonics,'' Proc. R. Soc. London, Ser. A **360**, 471–488 (1978).
[24] M. S. Longuet-Higgins and E. D. Cokelet, ''The deformation of steep surface waves on water. II. Growth of normal-mode instabilities,'' Proc. R. Soc. London, Ser. A **364**, 1–28 (1978).
[25] K. H. Ebuchi, H. Kawamura, and Y. Toba, ''Fine structure of laboratory wind-wave surfaces studied using an optical method,'' Boundary-Layer Meteorol. **39**, 133–152 (1987).
[26] M. S. Longuet-Higgins and R. P. Cleaver, ''Crest instabilities of gravity waves. I. The inner solution,'' J. Fluid Mech. **258**, 115–129 (1994).
[27] M. S. Longuet-Higgins, R. P. Cleaver, and M. J. H. Fox, ''Crest instabilities of gravity waves. II. Matching and asymptotic analysis,'' J. Fluid Mech. **259**, 333–344 (1994).
[28] M. S. Longuet-Higgins, ''Shear instability in spilling breakers,'' Proc. R. Soc. London, Ser. A **446**, 397–410 (1994).
[29] J. H. Duncan, V. Philomin, M. Behres, and J. Kimmel, ''The formation of spilling breaking water waves,'' Phys. Fluids **6**, 2558–2560 (1994).
[30] J. H. Duncan, H. Qiao, V. Philomin, and A. Wenz, ''Gently spilling breaker crest profile evolution,'' J. Fluid Mech. **379**, 191–222 (1999).
[31] J. A. Battjes and T. Sakai, ''Velocity field in a steady breaker,'' J. Fluid Mech. **111**, 421–438 (1981).
[32] M. L. Banner and D. H. Peregrine, ''Wave breaking in deep water,'' in *Annual Review of Fluid Mechanics*, edited by J. L. Lumley, M. van Dyke, and H. L. Reed (Annual Reviews, Palo Alto, CA, 1993), Vol. 25, pp. 373–398.
[33] M. S. Longuet-Higgins, ''New insights into breaking waves and bubble entrainment,'' in *Sea Surface Sound '94*, edited by M. J. Buchingham and J. R. Potter, Proceedings of the IIIrd International Meeting on Natural Physical Processes Related to Sea Surface Sound, University of California, Lake Arrowhead, 1994 (World Scientific, Singapore, 1995).
[34] D. M. Farmer and D. D. Lemon, ''The influence of bubbles on ambient noise in the ocean at high wind speeds,'' J. Phys. Oceanogr. **14**, 1762–1778 (1984).
[35] A. B. Wood, *A Textbook of Sound* (MacMillan, New York, 1941), pp. 360–363.
[36] K. W. Commander and A. Prosperetti, ''Linear pressure waves in bubbly fluids: Comparison between theory and experiments,'' J. Acoust. Soc. Am. **85**, 732–746 (1989).
[37] E. Lamarre and W. K. Melville, ''Instrumentation for the measurement of void-fraction in breaking waves: Laboratory and field results,'' IEEE J. Ocean Eng. **17**, 204–215 (1992).
[38] E. Lamarre and W. K. Melville, ''Void-fraction measurements and sound speed fields in bubble-plumes generated by breaking waves,'' J. Acoust. Soc. Am. **95**, 1317–1328 (1994).
[39] E. Lamarre and W. K. Melville, ''Sound-speed measurements near the ocean surface,'' J. Acoust. Soc. Am. **96**, 3605–3616 (1994).
[40] W. M. Carey, ''Low- to mid-frequency oceanic noise,'' in *Sea Surface Sound '97*, Proceedings of the IVth International Meeting on Natural Physical Processes Related to Sea Surface Sound, Institute for Sound and Vibration Research, Southampton, UK, 1997.
[41] B. R. Kerman, ''Underwater Sound Generation by Breaking Wind Waves,'' J. Acoust. Soc. Am. **75**, 149–165 (1984).
[42] A. Prosperetti, ''Bubble-related ambient noise in the ocean,'' J. Acoust. Soc. Am. **84**, 1042–1054 (1988).
[43] I. N. Kozhevnnikova and L. Bjørnø, ''Near sea surface bubble cloud oscillations as potential sources of ambient noise,'' in Ref. 5, pp. 339–348.
[44] W. M. Carey, J. W. Fitzgerald, E. C. Monahan, and Q. Wang, ''Measurement of the sound produced by a tipping trough with fresh and salt water,'' J. Acoust. Soc. Am. **93**, 3178–3192 (1993).
[45] G. Barton, *Elements of Green's Functions and Propagation* (Oxford University Press, Oxford, 1989).
[46] J. P. Bendat and A. G. Piersol, *Random Data*, 2nd ed. (Wiley–Interscience, New York, 1986).
[47] N. R. Goodman, ''Statistical analysis based on a certain multivariate Gaussian distribution (an introduction),'' Ann. Math. Stat. **34**, 152–177 (1963).
[48] J. Wu, ''Individual characterstics of whitecaps and volumetric description of bubbles,'' IEEE J. Ocean Eng. **17**, 150–158 (1992).
[49] R. K. Andrew, D. M. Farmer, and R. L. Kirlin, ''Broadband acoustical imaging of breaking ocean waves,'' J. Acoust. Soc. Am. **101**, 3032 (1997).
[50] R. K. Andrew, Broadband Acoustical Superresolution Imaging of Breaking Ocean Waves, Ph.D. thesis, University of Victoria, Victoria, BC, Canada, 1997.
[51] B. R. Kerman, ''On the distribution of bubbles near the ocean surface,'' in *Sea Surface Sound*, edited by B. R. Kerman (Kluwer Academic, Dordrecht, The Netherlands, 1988), pp. 185–196.
[52] *Shore Protection Manual*, Coastal Engineering Research Center, Department of the Army, Waterways Experiment Station, Corps of Engineers, P.O. Box 631, Vicksburg, MS, 1984, Vols. 1 and 2.
[53] D. J. Thomson, ''Spectrum estimation and harmonic analysis,'' Proc. IEEE **70**, 1055–1096 (1982).
[54] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications* (Cambridge University Press, Cambridge, 1993).
[55] J. Berntsen and T. Espelid, ''DCUTRI: An algorithm for adaptive cubature over a collection of triangles,'' ACM Trans. Math. Softw. **18**(3), 329–342 (1992), algorithm No. 706.
[56] P. H. LeBlond and L. A. Mysak, *Waves in the Ocean* (Elsevier, Amsterdam, 1978).
[57] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap* (Chapman and Hall, New York, 1993).
[58] R. J. Urick, *Principles of Underwater Sound*, 3rd ed. (McGraw–Hill, New York, 1983).
[59] L. Brekhovskikh and Y. Lysanov, *Fundamentals of Ocean Acoustics* (Springer, Berlin, 1982).
[60] H. N. Oğuz, ''Bubble clouds as sources of underwater noise,'' J. Acoust. Soc. Am. **95**, 1895–1912 (1994).

Further analysis of intensity fluctuations from a 3252-km acoustic propagation experiment in the eastern North Pacific Ocean

John A. Colosi
Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543

Fred Tappert
Division of Applied Marine Physics, University of Miami, RSMAS, Miami, Florida 33149

Matthew Dzieciuch
Scripps Institution of Oceanography, La Jolla, California 92093-0225

(Received 14 June 2000; revised 23 January 2001; accepted 6 March 2001)

In the Acoustic Thermometry of Ocean Climate (ATOC) program's Acoustic Engineering Test (AET), broadband 75-Hz center frequency transmissions were recorded on a 700-m-long vertical array, 3252 km distant from a midwater source suspended from R/P FLIP. The transmissions occurred over a 6-day period. Previously reported results from the AET using 12.7-min averaged data by Colosi *et al.* [J. Acoust. Soc. Am. **105**(6), 3202–3218 (1999)], hereafter referred to as Colosi99) revealed surprisingly weak acoustic scattering for early arriving identifiable wavefronts. Colosi99 found pulse time spreads on the order of 0–5 ms and the probability density function (PDF) of peak intensity was close to log normal. In this paper these results are confirmed using 1.8-min averaged data. It is also shown that scintillation index (SI) is a strong function of position along the pulse with the smallest values occurring at the peak and larger values occurring at the tails. Intensity PDFs of identifiable wavefronts are reanalyzed in terms of both peak intensity and integrated pulse energy (IE) where the integration is over ± 50 ms from the wavefront peaks. While SI for the IE are somewhat smaller than for the peak intensity, the PDFs are both very closely log normal. Regarding multipathing along the wavefronts, it is found that on average there are 1.7 peaks per wavefront segment per hydrophone and the intensity PDF of all multipath peaks is log normal. The combined observation of weak scattering and multipathing is a novel result. A reanalysis of the scintillations in the AET transmission finale where no wavefronts are evident is presented. Colosi99 analyzed the finale in terms of peak scintillations and found a near log-normal intensity PDF. Reprocessing the full field without limiting data to intensity peaks and accounting for mean intensity nonstationarity yields an intensity PDF which is much closer to the exponential distribution associated with full saturation; these results show that the finale region can be expected to behave much like Gaussian random noise.
[DOI: 10.1121/1.1369100]

PACS numbers: 43.30.Re, 43.30.Cq, 43.30.Qd [DLB]

I. INTRODUCTION

In November of 1994 an Acoustic Engineering Test (AET) was carried out by the Acoustic Thermometry of Ocean Climate (ATOC) program to examine the precision to which basin scale temperatures can be measured acoustically, and to study acoustic variability attributed to internal waves. A comparison between these observations and acoustic fluctuation theory based on internal wave dominance was reported by Colosi *et al.*[1] (hereafter referred to as Colosi99) and the analysis revealed surprisingly weak acoustic fluctuations for early arriving identifiable wavefronts. Observed pulse time spreads were on the order 0 to 5 ms rms and the probability density function (PDF) for wavefront peak intensity was close to log normal. These results were in plain disagreement with theoretical calculations presented in Colosi99, which predicted pulse time spreads of 0.2 to 1 s and an intensity PDF near exponential (Rayleigh in amplitude). In addition, Colosi99 showed that the intensity PDF of peaks in the wavefront finale region also had a near log-normal shape.

In this paper the analysis of intensity variability from the AET is extended to address several important issues. Examples of the intensity patterns from one AET transmission for the finale and wavefront regions are shown in Fig. 1.

First, long-range acoustic transmission data require pulse compression and time integration to obtain adequate SNRs from peak-power limited sources. Therefore, the effects of signal processing on observed signal statistics are important, especially when a comparison is made to theoretical expressions which do not account for specific experimental processing. With this in mind, it is conceivable that the observation of weak scattering in Colosi99 is a result of the 12.7-min averaging done in the signal processing which can diminish highly scattered incoherent energy with time scales less than 12.7 min. In this paper the effects of signal integration time on the signal statistics are examined, and

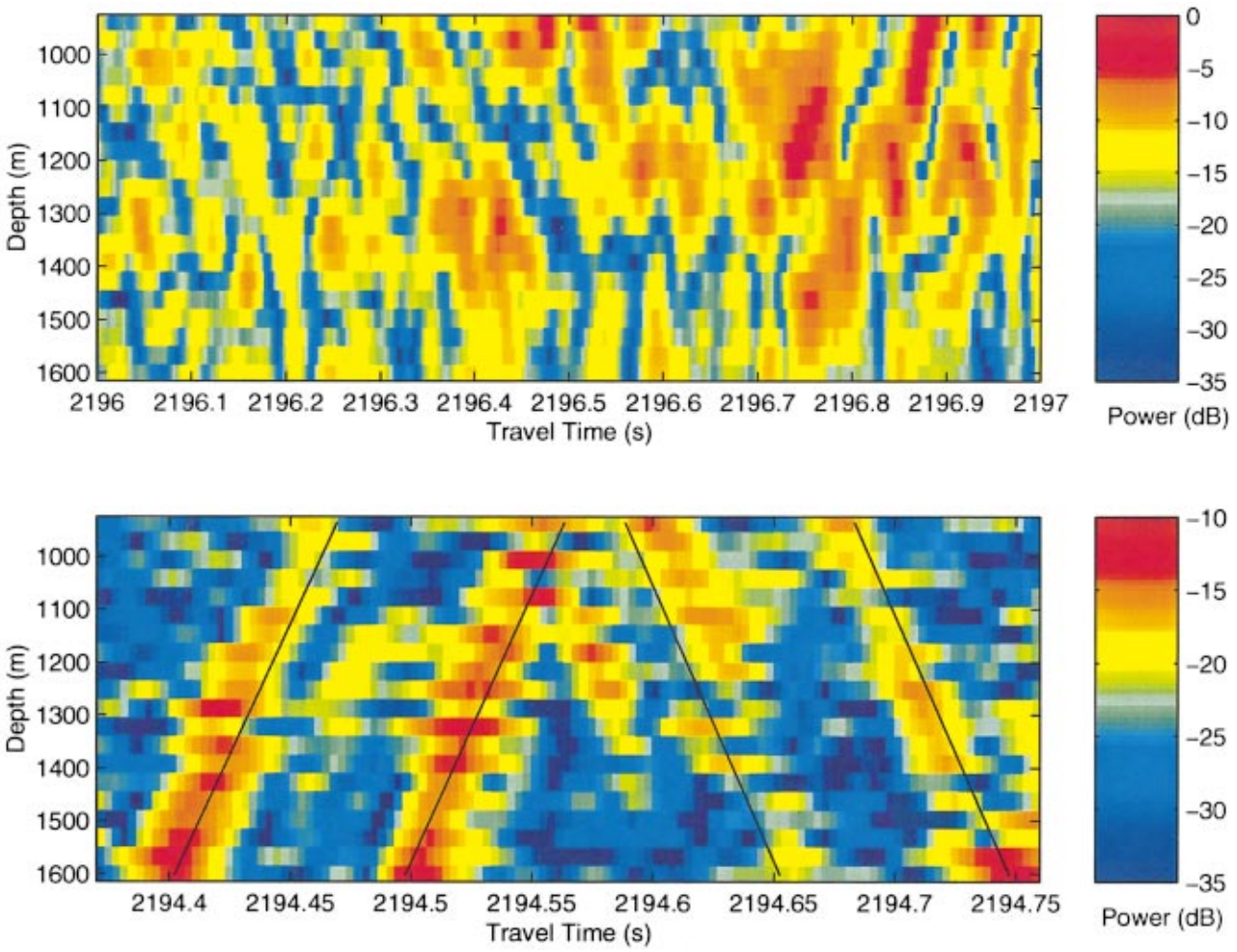


FIG. 1. Examples of intensity variability from AET data taken on yearday 322, hour 4. A total of 28 m-sequences was processed to generate this data. The maximum SNR for this pulse is 32 dB; therefore, the bottom of the color scale roughly represents the noise level. The upper panel shows intensity variability in the arrival finale and the lower panel shows intensity variability along four wavefront segments. Solid black lines in the lower panel are geometrical acoustics wavefront predictions.

small effects on all statistics are found with averaging times of 12.7, 7.25, 3.63, and 1.81 min.

Second, a reanalysis of the intensity PDF of identifiable wavefronts is presented which uses the integrated energy (IE) over the pulse within ± 50 ms around the peak. This approach is taken since the resulting fluctuation might be more sensitive to small variations in the pulse shape. However, it is found that the IE and peak intensity PDFs are both very closely log normal, and they are relatively insensitive to signal integration times. Also, it is found that scintillation index (SI) for the IE are less than those for the peak intensities.

Third, the effect of wavefront multipathing is addressed. Figure 1 shows that in the wavefront region there can be multiple peaks associated with a single wavefront arm or segment, and these multiple peaks are within tens of milliseconds of the segment. It is found that on average there are 1.7 peaks per wavefront segment per hydrophone, and multipathing is responsible for the observed ''tails'' on the ensemble average pulses presented in Colosi99 and here. In addition, multipathing results in an increase in the scintillation index (SI) in the neighborhood of a peak. It is found here that minimum SI values occur at the pulse peak, and this value more than doubles 30 ms away from the peak. Further, the intensity PDFs presented in Colosi99 used only one peak per wavefront per hydrophone and it is shown here that if all the multipath peaks are considered the PDF maintains its log-normal form.

Fourth, a reanalysis of the scintillations in the transmission finale where no identifiable wavefronts are evident is presented. Figure 1 shows the intensity pattern in the finale of one transmission: Strong focuses and deep fades are apparent. Colosi99 discussed finale PDFs based on peak intensities. Here, this restriction is relaxed and an analysis of the full field in the finale is presented. This analysis reveals that the intensity PDF in the finale is much closer to the exponential distribution associated with full saturation, and therefore the acoustic field in the finale can be expected to behave like Gaussian random noise. While an exponential intensity PDF does not fit the data well, a modulated exponential (ME) model based on small variations in the mean intensity for the exponential distribution does fit the data well. These results are insensitive to signal integration times. The observation of a near-exponential PDF for the finale is consistent with our understanding of the strong mode coupling[2] or ray chaos[3,4] in the finale.

The organization of this paper is as follows. Section II gives a brief explanation of AET signal processing. The discussion is then separated in terms of wavefront statistics which are presented in Sec. III and wave-field scintillations in the finale region of the pulse which are presented in Sec. IV. In Sec. V a summary and conclusions are given.

II. AET DATA PROCESSING

The AET data processing is explained in detail in the companion paper to Colosi99 by Worcester *et al.*[5] and other discussions of acoustic tomography signal processing can be found in the monograph by Munk[6] and the references therein. The data presented in Colosi99 were obtained by coherently processing (with Doppler corrections) 28 replicas of the signal where each replica was of 27.280-s duration. In actuality 40 replicas were transmitted, but for replica averages greater than 28 the gain was no longer characterized by ideal linear growth. In this paper the AET data is reprocessed exactly as was done in Colosi99, except 28, 16, 8, and 4 period integration times are considered, and many of the statistics discussed in Colosi99 are recomputed for these integration times.

III. WAVEFRONT STATISTICS

A. Average pulse shape

Average pulse shape for 28, 16, 8, and 4 period data averages are computed slightly differently here than in Colosi99. For the present computation no intensity threshold was used (a 12-dB SNR threshold was used in Colosi99). The intensity threshold was eliminated so as to include lower SNR peaks which might have more time spread. Also, because of lower SNRs for the lower number of sequence averages, the average pulse shapes are only computed for the higher wavefront IDs ± 138 through ± 135 (in Colosi99 the average was over IDs ± 138 through ± 126). Figure 2 shows the average pulse shapes for the four different averaging intervals. These curves are to be compared with the upper panel of Fig. 12 in Colosi99. Clearly the different signal integration times have a small effect on the pulse shapes. There is a small trend for the pulses to be narrower for the shorter time integrations. This effect may be due to uncorrected Doppler spread caused by mooring/source motion, or it may be due to changes in the signal structure.[4] Also, in comparing to Fig. 12 of Colosi99 it is seen that the intensity threshold does not make a big difference in the average pulse shape.

Because the average pulse shapes do not change dramatically from 28 to 4 period averages, then if there is an incoherent part of the signal that is being suppressed by the signal processing its time scale must be significantly less than 2 min. This result sets an interesting upper bound on the incoherent field time scale. The signal cannot be analyzed using less than 4 sequences because of low signal to noise level.

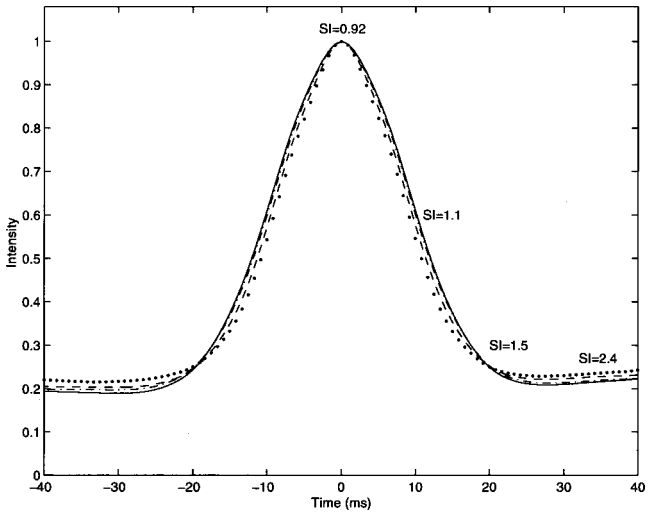


FIG. 2. AET ensemble averaged pulse for four different m-sequence averages. The solid curve is the 28 period average and the dash-dot, dash, and dot curves are the 16, 8, and 4 period averages, respectively. The average pulse is computed over all 47 transmissions, 20 depths, and over wavefront IDs ± 138 through ± 135. No intensity threshold was used. Values of the scintillation index as a function of time delay along the ensemble average pulse are denoted for the 28 period average data for wavefront IDs ± 137 and 138 (as in Table I). SI values correspond to time delays of 0, ± 10, ± 20, and ± 33 ms.

B. Peak intensity and integrated energy PDFs

Next, wavefront scintillations are examined and two observables are considered. First, wavefront peak intensity is treated and then time-integrated energy (IE) is treated. The IE is computed for each wavefront arrival by time integrating the wavefront intensity from ± 50 ms around the peak intensity. From Fig. 2 it can be seen that this time integration extends well into the tails of the average pulse.

The scintillation index (SI) for both peak intensity and IE is defined as

$$\mathrm{SI} = \frac{\langle I^2 \rangle}{\langle I \rangle^2} - 1, \tag{1}$$

and Table I shows specific values as a function of signal integration time. For the SI computations in Table I an average over wavefront IDs ± 138 and ± 137 is used and again, no intensity threshold is imposed. However, in keeping with the conventions of Colosi99, peaks near wavefront crossing regions are excluded, and only one peak per wavefront per hydrophone is considered. Note that the SI value of 0.92 for the 28 sequence average is larger than the value of 0.74 (for ID $+138$) presented in Colosi99. This difference can be at-

TABLE I. Scintillation index as a function of number of m-sequences averaged. For peaks and integrated energy the scintillation index is computed over the wavefront IDs ± 138 and ± 137. Square brackets show the number of data used in the scintillation index calculation.

	Number of m-sequences			
Observable	28	16	8	4
Peak energy	0.92 [3231]	0.96 [3169]	0.91 [3204]	0.79 [3239]
Integrated energy	0.65 [3231]	0.67 [3169]	0.55 [3204]	0.52 [3239]
Finale	1.39 [19740]	1.38 [19320]	1.39 [19320]	1.38 [19320]

TABLE II. Kolmogorov–Smirnov test probability (percent) for rejecting the null hypothesis for the log-normal distribution (peaks and integrated energy) and the modulated exponential (ME) distribution (finale) as a function of the number of m-sequences averaged. For peaks and integrated energy the statistics are computed over the wavefront IDs ± 138 and ± 137. The probability of rejecting the null hypothesis for the exponential distribution is essentially zero.

Observable	Number of m-sequences			
	28	16	8	4
Peak energy (log normal)	99.9	63.5	30.3	33.8
Integrated energy (log normal)	90.6	76.8	54.9	80.0
Finale (ME)	22.6	7.3	4.4	0.4

tributed to the 12-dB threshold used in Colosi99. Table I shows that there is very little dependence of SI on the number of m-sequences averaged.

To examine the intensity PDFs of the peaks and IE, a Kolmogorov–Smirnov (KS) test[7] is performed to compute the probability of rejecting the null hypothesis for two candidate distributions which are the log-normal and exponential distributions. Table II shows the probabilities (given in percent) of rejecting the null hypothesis for the data/log-normal distribution comparison. Similar probabilities for the data/exponential distribution comparison yield essentially zero probability; that is, the exponential distribution can be rejected as a candidate distribution for the observations. Table II shows that for the intensity peaks and IE, signal time integration effects are very small and the observations can be fit fairly well by a log-normal distribution. Because of the sensitivity of the KS test, probability values of order 10 percent are considered good and variations within this order are not terribly significant.

C. Multipathing effects

From Fig. 1 there is clear evidence of multipathing within a specific wavefront segment, and in Fig. 2 it is seen that a result of the multipathing process is to produce tails on the average pulse. Figure 2 also shows SI values (computed from 28 sequence averages) as a function of time delay relative to the peak travel time. Moving away from the peak time there is an increase in the SI values with the maximum values occurring on the pulse tails. From Table I in Colosi99 it can be seen that the energy in the pulse tails is still over 10 dB above the noise level, so noise effects can be ignored. A KS test comparing the intensity PDF on the tails (33-ms time delay) to a log-normal or an exponential distribution rejects both distributions. Evidently the average pulse tails are made up of highly variable acoustic energy, which is distinct from the more stable peak energy, and therefore the tail energy may be considered to behave as if it were in a different wave propagation regime (namely the partially saturated regime) than the unsaturated energy which makes up the peak.

Because of the multipathing effect, the treatment in Colosi99 of using only one peak per wavefront per hydrophone is examined. Using 28 sequence-averaged data all peaks are selected within a ± 50 ms window of our template.[1] The intensity PDFs for these two cases are shown in Fig. 3. For only one peak per wavefront per hydrophone there are 3231 points, but including all peaks there are 5556 points, which gives an average of 1.7 peaks per wavefront per hydrophone. Figure 3 shows that the PDF is still closely log normal even when multiple peaks are considered.

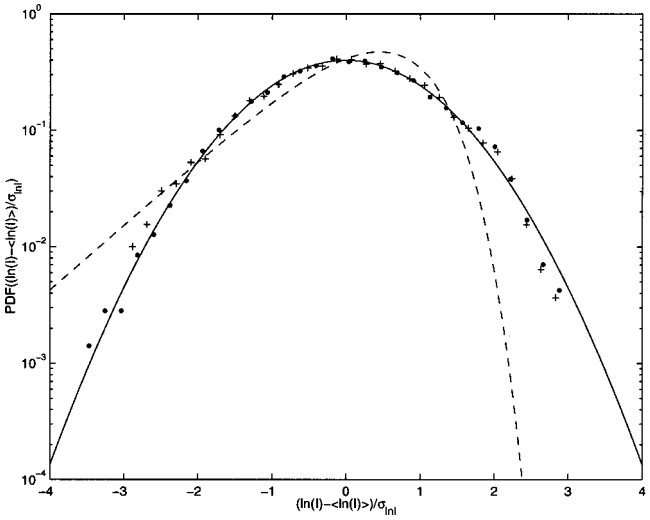


FIG. 3. Intensity PDF for wavefronts ± 137 and ± 138 using only one peak per wavefront per hydrophone (+) and using all peaks (•). Theoretical PDFs (normal-solid and exponential-dash) are plotted for comparison. For the one peak per wavefront per hydrophone there are 3231 data and for all peaks there are 5556 data, for an average of 1.7 peaks per wavefront per hydrophone. These PDFs were generated using data from 28 m-sequence averages.

IV. FINALE SCINTILLATIONS

A. Modulated exponential distribution

In this section the modulated exponential (ME) distribution function[8] for the intensity is introduced by first analyzing AET noise data. AET noise data were tabulated starting at a travel time of 2197.75 s and points every 66.6 ms (20 samples) were collected out to a travel time of 2204.416 s for each of the 20 hydrophones and each of the 47 transmissions. The starting point of the noise data collection (2197.75 s) occurred roughly 0.75 s after the transmission finale. The mean noise intensity was computed over all 20 hydrophones for each transmission and the results are shown in Fig. 4. Clear nonstationarity in the mean noise intensity is evident.

Figure 5 shows the probability density functions (PDF)s of the noise field. In the upper panel the noise PDF for the entire duration of the experiment is displayed, SI for this case is 1.35, and the exponential distribution function for intensity does not model the observations well. In the lower panel the PDF of the noise is displayed over an approximately stationary period using the first two transmissions of yearday 322. In this stationary case SI$=1.00$ and the PDF is modeled very well by a pure exponential distribution function for intensity. The case of intensity nonstationary can be treated, however, of which an extensive literature exists in optical propagation (see Churnside and Hill[8] and references therein).

The exponential intensity distribution (Rayleigh in amplitude) is given by

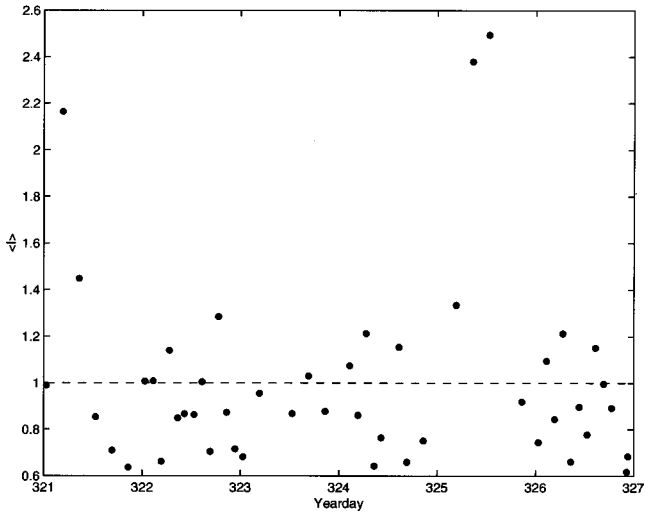


FIG. 4. AET mean noise intensity as a function of yearday. The value of b, the ratio of the mean intensity variance divided by the square of the mean intensity, is 0.164.

$$P(I)=\frac{1}{\langle I\rangle}\exp\left(-\frac{I}{\langle I\rangle}\right). \tag{2}$$

Taking $\langle I\rangle = I_0 + a$, where it is assumed $a/I_0 \ll 1$, then Eq. (2) can be expanded to second order in a/I_0, yielding

$$P(I;a)\simeq P_0(x)\left[1-\frac{a}{I_0}(1-x)+\frac{a^2}{I_0^2}(1-2x+x^2/2)\right], \tag{3}$$

where $x=I/I_0$, and $P_0(x)=(1/I_0)\exp(-x)$. Next, the parameter a is considered a random variable that is normally distributed with mean zero and variance σ^2. The expected value of the PDF can be calculated analytically, yielding

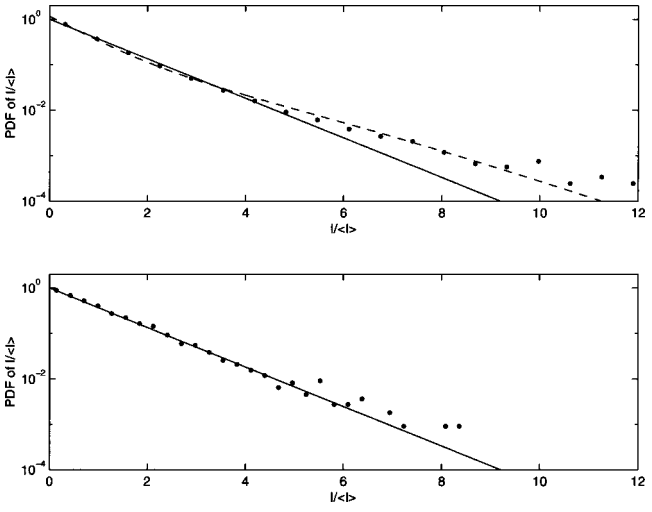


FIG. 5. AET noise intensity PDF for the entire duration of the experiment (upper panel •) and an AET noise PDF for a 2-h stationary period using the first two transmissions of yearday 322 (lower panel •). In the upper panel the dash curve is the ME distribution ($b=0.164$) and the solid curve is the exponential distribution. For the upper panel SI=1.35 and the modulated exponential (ME) model predicts SI=1.328. For the lower panel SI=1.00.

$$\langle P(I)\rangle = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} P(I;a)\exp\left(-\frac{a^2}{2\sigma^2}\right)da$$

$$= P_0(x)\left[1+\frac{b}{2}(x^2-4x+2)\right], \tag{4}$$

where $b=\sigma^2/I_0^2$. The scintillation index for this model is SI$=1+2b$.

From Fig. 4 the value of b is computed to be 0.164 (including the points in Fig. 4 which might be considered outliers for a Gaussian distribution), which yields an estimate of 1.33 for SI that compares favorably with the observed value of 1.35. The upper panel of Fig. 5 compares the noise PDF for the entire, nonstationary, duration of the experiment with the ME model distribution, and the ME distribution is a significant improvement over the pure exponential model.

B. AET finale

In the present analysis of the wavefront finale, the approach is changed significantly from Colosi99 and scintillations of the full field are considered not just intensity peaks. The finale region is defined similarly to Colosi99 as the travel time region between 2195.75 to 2197.0 s for all 20 hydrophone depths. The observed acoustic intensity at each depth is subsampled in time by starting at 2195.75 s and advancing 20 samples (66.7 ms) at a time until the travel time of 2197 is reached. Denote this data as $I_k(z,t)$, where there are 20 depths (z), 21 time points (t), and 47 transmissions (k). This procedure gives independent samples in both depth and time. Next, the mean intensity is computed for each depth and each time point

$$\langle I\rangle_{z,t}=\frac{1}{47}\sum_{k=1}^{47} I_k(z,t), \tag{5}$$

and finally the intensity quantity for which the PDF is computed is defined as

$$I_k(z,t)/\langle I\rangle_{z,t}. \tag{6}$$

This method accounts for the variation of the mean intensity as a function of position in the finale region.

Figure 6 shows the observed PDF for the finale region with the pure exponential and the ME PDFs. In the computation of the data PDFs no intensity threshold is used; in Colosi99 a 14-dB SNR threshold was used on the intensity peaks. The ME PDF for $b=1.9$ fits the observations well and a KS test gives a 22.6 percent probability for rejecting the null hypothesis. Note also that the ME model predicts the distinctive crossing point between the observations and the exponential model at $I/\langle I\rangle = 2\pm 2^{1/2}$. The finale PDF in Fig. 5 is almost identical to the noise PDF before mean intensity nonstationarity is removed (see the upper panel of Fig. 5). The finale SI value is 1.38 and the noise value is 1.34. Computing the finale PDF over 1/2 and 1/4 of the data record length does not show a PDF converging to the pure exponential model. Therefore, if intensity mean nonstationarity is the explanation for the finale PDF, then the nonstationarity

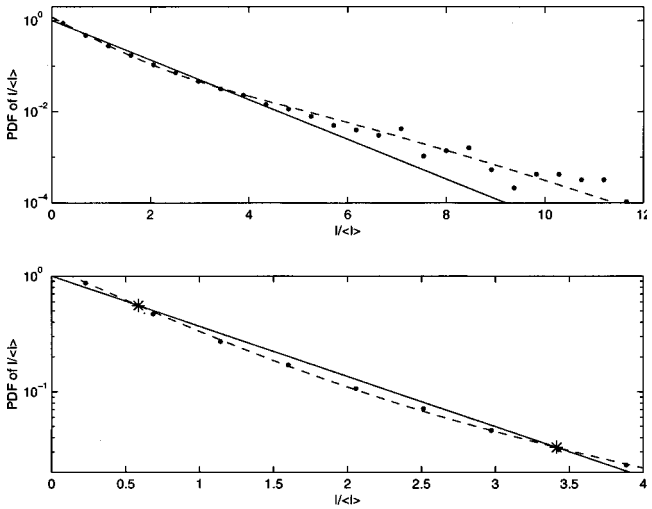


FIG. 6. AET intensity PDF for the finale. Observations (•) are for 28 m-sequence averages and there is no intensity threshold used. The solid curve is the PDF for the exponential distribution and the dash curve is for the ME distribution with $b=0.19$. The lower panel is an expanded view for the small $I/\langle I\rangle$ region showing the distinctive crossing (*) of the ME and exponential distributions at $I/\langle I\rangle = 2 \pm 2^{1/2}$. The Kolmogorov–Smirnov test probability for rejecting the null hypothesis for the ME distribution is 22.6% and the probability for rejecting the null hypothesis for the exponential distribution is essentially zero.

must take place on a time scale of less than 1–2 days. PDFs for time periods less than this cannot be computed because the evaluation of Eq. (6) becomes poor.

Finally, Table I shows that the value of SI in the finale region is insensitive to the signal averaging time, but Table II shows a significant reduction in the quality of fit for the ME distribution as signal averaging time is reduced. This reduction in the quality of fit may be due to noise sensitivity in that it is seen in general that all of the observables in Table II have a reduced quality of fit to their candidate distributions.

V. SUMMARY

Four main issues have been addressed in this paper relating to intensity variability of 3250-km range, 75-Hz broadband transmissions from the AET.

Regarding the effects of coherent signal processing on intensity variability, it is found that the average pulse shape is unchanged when signal integration times from 12.7 to 1.8 min are considered. This means that if indeed signal processing is averaging out an incoherent component of the acoustic field, the time scale of that incoherent component is much less than a few minutes. Furthermore, it is found that intensity PDFs and SI values are insensitive to the signal-processing durations chosen.

On the subject of wavefront PDFs it is found that observables of peak intensity and time-integrated energy, IE, are both closely log normal. The observation of a log-normal PDF may have interesting connections to recent ray-dynamics results which show that the intensity PDF of chaotic rays at finite range is also log normal.[9] Interestingly, recent analysis of wavefront PDFs where neither peaks nor IE are considered but intensities along a wavefront template (determined by a turning point filter[10]) are the quantities for which the PDF is calculated show a PDF closer to exponential (M. Dzieciuch, personal communication).

Wavefront multipathing or fracturing is an important element in this analysis. It is found that on average there are 1.7 peaks per wavefront per hydrophone, yet the intensity PDF is log normal regardless if all peaks are considered or only one peak per hydrophone per wavefront is considered. In addition, multipathing is the cause of the distinct tails that are observed on the average pulse. It is interesting that the multipathing pattern is such that multipaths rarely overlap in time; that is, the average pulse is still thin in the middle but at larger time delays multipaths fill in the energy. Also, it is found that multipathing causes an increase in the intensity variability as a function of time delay from the pulse peak.

The aforementioned results supplement the analysis of Colosi99 and confirm that the AET wavefronts have very little time spread and the wavefront intensity PDFs are very close to log normal. These are conditions typical of weak scattering or the unsaturated wave propagation regime,[11] although wavefront SI values are between 0.52 and 0.96; these are values which are much larger than traditional weak scattering limits of 0.3 or so. However, the unsaturated wave propagation regime is not characterized by multipathing and therefore our observations present a picture of acoustic variability which is interestingly inconsistent with established wisdom.[11]

Finally, reprocessing the AET finale data in terms of the full field shows an intensity PDF that is very well described by a modified exponential, ME, distribution which takes into account mean intensity nonstationarity. The quality of the fit of the ME distribution suggests that the statistics of the finale are indeed described by the pure exponential distribution (Rayleigh in amplitude) if the finale could be observed over a stationary time period. It is estimated that the stationarity period is significantly less than 1–2 days, and the cause could be one or more of several processes. First, the AET had source and receiver motion which could affect mean intensity. Other candidate processes are the internal tide or small changes in the mesoscale field which were observed by Worcester *et al.*[5] It is reiterated that the observation of Rayleigh statistics for the finale are consistent with our understanding of the strong mode coupling[2] and ray chaos[3,4] in the finale region; however, it must be noted that even in strong scattering cases the approach to pure exponential is quite slow if analogies to optical propagation are any guide.[12]

The transition from weak scattering in the wavefront region to the strong scattering in the finale is a fascinating topic for future study.

ACKNOWLEDGMENTS

This work was supported by the Strategic Environmental Research and Development Program through Defense Advanced Research Projects Agency (DARPA) Grant No. MDA 972-93-1-0003. This work was partially supported by the Office of Naval Research through the Ocean Acoustics Program Grants N00014-97-1-0068 and N00014-96-1-0254. John Colosi wishes to acknowledge support from an ONR Young Investigator Award. The authors would also like to acknowledge useful comments from Mike Brown, Stanley

Flatté, Frank Henyey, Jeff Simmen, Steve Tomsovic, Anitoly Virovlyansky, Mike Wolfson, and George Zaslavsky. This is Woods Hole Oceanographic Institution Contribution Number 10162.

[1] J. A. Colosi, E. K. Scheer, S. M. Flatté, B. D. Cornuelle, M. A. Dzieciuch, W. H. Munk, P. F. Worcester, B. M. Howe, J. A. Mercer, R. C. Spindel, K. Metzger, T. G. Birdsall, and A. B. Baggeroer. ''Comparisons of measured and predicted acoustic fluctuations for a 3250-km range propagation experiment in the eastern North Pacific Ocean,'' J. Acoust. Soc. Am. **105**, 3202–3218 (1999).

[2] J. A. Colosi and S. M. Flatté, ''Mode coupling by internal waves for multi-megameter acoustic propagation in the ocean,'' J. Acoust. Soc. Am. **100**, 3607–3620 (1996).

[3] J. Simmen, S. M. Flatté, and G.-Y. Wang, ''Wavefront folding, chaos, and diffraction for sound propagation through ocean internal waves,'' J. Acoust. Soc. Am. **102**, 239–255 (1997).

[4] F. D. Tappert and X. Tang, ''Ray chaos and eigenrays,'' J. Acoust. Soc. Am. **99**, 185–195 (1996).

[5] P. F. Worcester, B. D. Cornuelle, M. A. Dzieciuch, W. H. Munk, B. M. Howe, J. A. Mercer, R. C. Spindel, J. A. Colosi, K. Metzger, T. G. Birdsall, and A. G. Baggeroer, ''A test of basin-scale acoustic thermometry using a large-aperture vertical array at 3250-km range in the eastern North Pacific Ocean,'' J. Acoust. Soc. Am. **105**, 3185–3201 (1999).

[6] W. Munk, P. Worcester, and C. Wunsch, *Ocean Acoustic Tomography* (Cambridge University Press, Cambridge, 1995).

[7] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipies in* FORTRAN: *The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1992).

[8] J. Churnside and R. Hill, ''Probability density of irradiance scintillations for strong path-integrated refractive turbulence,'' J. Opt. Soc. Am. A **4**, 727–733 (1987).

[9] M. Wolfson and S. Tomsovic, ''On the stability of long-range sound propagation through a structured ocean,'' J. Acoust. Soc. Am. (Submitted).

[10] M. Dzieciuch, W. Munk, and P. Worcester, ''Basin scale acoustics: Turning point filters,'' J. Acoust. Soc. Am. (Submitted).

[11] S. Flatté, R. Dashen, W. Munk, K. Watson, and F. Zachariasen, *Sound Transmission through a Fluctuating Ocean* (Cambridge University Press, Cambridge, 1979).

[12] S. M. Flatté, C. Bracher, and G.-Y. Wang, ''Probability-density functions of irradience for waves in atmospheric turbulence calculated by numerical simulation,'' J. Opt. Soc. Am. A **11**, 2080–2092 (1994).

Numerical simulation of optimal deconvolution in a shallow-water environment

Ben S. Cazzolato,[a] Philip Nelson, and Phillip Joseph
Fluid Dynamics and Acoustics Group, Institute of Sound and Vibration Research, University of Southampton, Southampton SO17 1BJ, United Kingdom

Richard J. Brind
DERA, Winfrith, Dorset, United Kingdom

(Received 5 June 2000; revised 12 March 2001; accepted 26 March 2001)

A fast technique for deconvolving signals in a dispersive multipath shallow-water environment using inverse filters is compared with the more commonly used deconvolution technique of time reversal (also known as phase conjugation). The objective of such techniques is to improve the accuracy of sound transmission from a source array to some receiving space. Time reversal provides benefits in this regard but here the additional performance that can be gained from inverse filters is examined. Several strategies for obtaining a set of inverse filters are discussed, each aimed at improving the accuracy of the reconstruction of the desired time signals through inverse techniques. It will be shown that an ''optimal'' inversion (in the sense of achieving a flat system response in the frequency domain) does not necessarily achieve a realizable time domain filter. A fast field model (using OASES) of the Giglio Basin shallow-water test facility is used as the basis for evaluating the various focusing strategies for single receiver locations. It will be seen that inverse filter arrays provide enhanced temporal and spatial focusing when compared to time reversal arrays. In addition, inverse filtering allows multiple receivers to be used, thereby increasing bandwidth or improving redundancy.
[DOI: 10.1121/1.1379081]

PACS numbers: 43.30.Vh, 43.30.Dr, 43.20.Fn, 43.30.Hw [DLB]

I. INTRODUCTION

The transmission of acoustic signals in shallow-water environments is often poor due to the dispersive multipath nature of the waveguide. This results in *intersymbol interference*,[1–3] which subsequently limits the capacity of underwater communication systems. Parvulescu and Clay[4] reported an experiment in which a *time-reversed* (or *phase-conjugated*) retransmission between a source and receiver compensated for the multipath effects. Several authors[5–15] have successfully extended the technique to multielement sensor arrays. Such *phase-conjugate arrays* also achieve spatial focusing in addition to overcoming the problems associated with refraction/reflection within the waveguide, unknown source/receiver response, and unknown array deformation. This type of technique has obvious applications in active sonar and in the improvement of underwater communication systems. Such is the interest currently being shown in this method that full-scale sea trials have already been undertaken with considerable success.[12,14]

In the field of ultrasonics, Tanter *et al.*[16,17] successfully extended the time-reversal technique by not only compensating for phase aberrations but also magnitude attenuation arising through absorption of an ultrasonic beam passing through a skull. They found that the zone of focus was improved over that of conventional time reversal.

It will be shown here that *inverse filters*, which also compensate for both phase and amplitude aberrations introduced by both the environment and transducers, greatly improve the signal quality within the focus zone compared to conventional time reversal. There is considerable literature on using optimal inverse filters for compensating for room responses (also known as *equalization*) in multichannel sound reproduction.[18–26] The same technique has been used here to compensate for the multiple path propagation in a shallow-water waveguide. It should be noted that most waveguides are dispersive, whereas a room response is not. However, it will be seen here that inverse filtering also performs well in dispersive media.

A typical communications scenario involves a single source sending out a ''ping'' to a source–receiver array from which the transfer functions from the source to the array elements are determined. Assuming reciprocity, the path from the array to the source is also known. This paper will examine the application of inverse filtering for the purpose of improving underwater communications by deconvolving the received signal from the multipath response of the waveguide.

In Sec. II the theory of the two deconvolution techniques, namely phase conjugation and optimal inverse filtering, are reviewed. A numerical model of the shallow-water test facility in the Giglio basin off the west coast of Italy[12,14,27] was created using OASES[28,29] and is discussed in Sec. III. The results of applying the deconvolution techniques are then compared against each other with the waveguide model in Sec. IV. The optimization of the inverse filters is then addressed and the temporal and spatial performance is quantified.

[a]Now at the Dept. of Mechanical Engineering, The University of Adelaide, SA 5005, Australia.

 0001-4966/2001/110(1)/170/16/$18.00

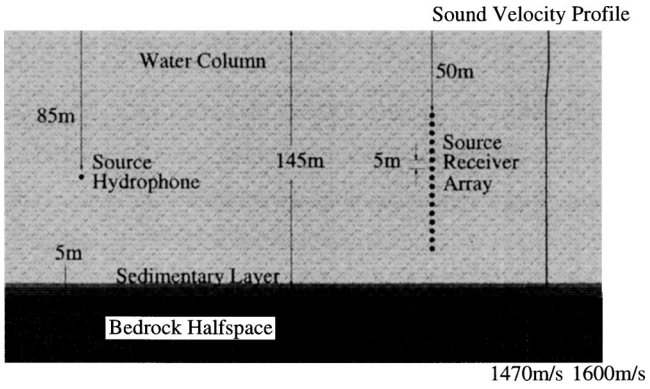


FIG. 1. Schematic of the numerical model used for the inverse filter experiment.

II. BACKGROUND TO THE THEORY OF SPATIAL AND TEMPORAL FOCUSING

Figure 1 shows a typical transmission problem where some array is used to send a signal to a point receiver. With both time reversal and inverse filtering, initially a system identification signal is sent from the probe source to the receive elements of the source–receiver array (SRA) to ''identify'' the medium (plant matrix) through which the sound is transmitted. Once the system identification has been performed, the transmit elements of the SRA are then used to transmit a deconvolved signal to the probe source. Reciprocity between the probe source and the SRA is a necessary condition for either of these two methods to work.

The equivalent block diagram of the transmission problem is shown in Fig. 2. A prefilter matrix $\mathbf{H}(\omega)$ is required, which when applied to the source input signals, $\mathbf{v}(\omega)$, leads to an improvement (based on some performance metric) in the reproduced signals $\mathbf{w}(\omega)$. This schematic assumes that the plant transfer function matrix has already been estimated. In practice this would be done by sending a system identification signal $\mathbf{u}(\omega)$ from the probe source, through the plant $\mathbf{C}(\omega)$ via an identity $\mathbf{H}$ matrix to the receive elements of the array $\mathbf{w}(\omega)$.

Two techniques for determining $\mathbf{H}(\omega)$ will now be reviewed; time reversal and optimal inverse filtering.

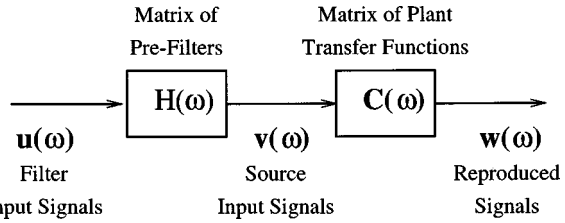


FIG. 2. System diagram showing prefilters and plant matrix operators. The variables are defined as follows: $\mathbf{u}(\omega)$ is the column vector of n_f filter input signals (representing the signal to be transmitted), $\mathbf{v}(\omega)$ is the column vector of n_s source input signals (representing the transmit end of the source–receiver array), and $\mathbf{w}(\omega)$ is the column vector of n_r reproduced signals at the receivers. In practice, the number of prefilters equals the number of elements in the source, i.e., $n_f = n_s$. The matrices $\mathbf{H}(\omega)$ and $\mathbf{C}(\omega)$ are the $(n_s \times n_s)$ prefilter matrix and the $(n_r \times n_s)$ plant transfer function matrix, respectively.

A. Time reversal

Time reversal (which is also the basis of *matched filter theory*) was first investigated for improving the transmission of continuous wave signals in a waveguide with a single source and a single receiver.[30–32] This is illustrated in Fig. 3 for a multipath single element SRA.

Time reversal involves initially sending out a signal from a probe source to the receive elements of an SRA to characterize the transmission path. The pressure, $w(t)$, due to an impulse from the source, $u(t)=\nu(t)=\delta(t)$, at a receiver is given by the convolution of the impulse (or signal) and the impulse response function between the source and the receiver, $c(t)$. That is, the signal at the receiver is given by

$$w(t)=\int_{-\infty}^{\infty}\delta(t-\tau)c(\tau)d\tau=c(t). \tag{1}$$

Tolstoy and Clay[33] showed that if a *matched filter h(t)* with an impulse response function given by the delayed and time-reversed plant response, $h(t)=c(t_0-t)$, is applied to a delta function, $\delta(t)$, prior to transmission from the SRA through the waveguide, then the signal at the receiver is given by

$$w(t)=\int_{-\infty}^{\infty}c(t_0-t-\tau)c(\tau)d\tau=\varphi_{11}(t-t_0). \tag{2}$$

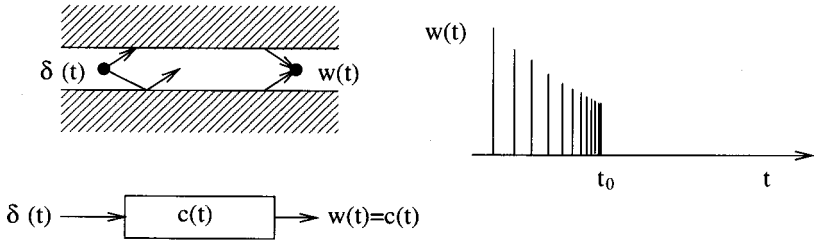
Equation (2) shows that the result of passing a signal through the matched filter $h(t)=c(t_0-t)$ prior to transmission is equivalent to the autocorrelation function, $\varphi_{11}(t)$, of the waveguide impulse response function delayed by some time t_0 [see Fig. 3(b)]. For time reversal to work the following conditions are required: there can be no losses in the medium, and also there must be time invariance of the transmission path under the time-reversal operation.[34] If either of these two constraints is not met then the performance of the array will be degraded.[6–8]

This simple single-input single-output example shows why time reversal works so effectively. The signal which is reproduced when employing time-reversal filters is simply the autocorrelation function of the plant matrix impulse response function. Thus, the reproduced signal consists of an impulse at $t=t_0$ surrounded by smaller impulses which typically decay away from $t=t_0$.

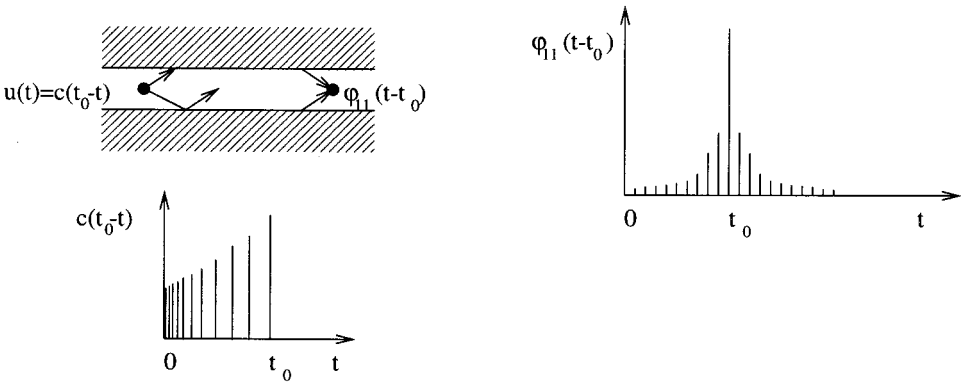
Time reversal (in the temporal domain) is equivalent to phase conjugation in the frequency domain[5,12] since the Fourier transform of $c(t_0-t)$ is simply given by $C^*(\omega)e^{j\omega t_0}$, where the term $e^{j\omega t_0}$ is associated with the delay in order to retain causality. With reference to Fig. 2 and Eq. (2), where $\mathbf{w}(\omega)=\mathbf{C}(\omega)\mathbf{H}(\omega)\mathbf{u}(\omega)$, the time-reversal process can be expressed by[12]

$$\mathbf{w}(\omega)=\mathbf{C}(\omega)\mathbf{C}^*(\omega)e^{j\omega t_0}\mathbf{u}(\omega)=|\mathbf{C}(\omega)|^2e^{j\omega t_0}\mathbf{u}(\omega). \tag{3}$$

Equation (3) identifies some of the benefits associated with using time-reversal arrays as opposed to single elements. For a single element, the response of the received signal when using time reversal, $|\mathbf{C}(\omega)|^2\mathbf{u}(\omega)$, has the square (or double in a logarithmic scale) of the magnitude aberration of a broadside transmission $\mathbf{C}(\omega)\mathbf{u}(\omega)$. Therefore,



(a) Impulse source $\delta(t)$ with waveguide as a filter



(b) Waveguide as a matched filter to signal $c(t_0 - t)$

FIG. 3. Application of time reversal to propagation in a waveguide. From Tolstoy and Clay (Ref. 33).

at nulls in the transfer function caused by destructive interference, there will be significant ''dropout'' in the signal. However, with an array, time reversal should provide an improved magnitude response compared to the broadside transmission since the transfer matrix $|\mathbf{C}(\omega)|^2$ is positive definite, whereas in the case of broadside beamforming $\mathbf{C}(\omega)$ is not; hence, significant destructive interference may still occur.

The variable magnitude with respect to frequency given by Eq. (3) may reduce the quality of FM signals, particularly if either of the side bands lie near a trough in the magnitude of the frequency response function. This concern was also raised by Parvulescu,[35] with regards to transmission of a continuous wave pulse which was accidently set to a ''cancellation frequency'' (or plant matrix zero). He noted that under such circumstances the time-reversed signal would not represent the desired signal at all.

Iterative time reversal has been used to improve the performance of single-frequency time reversal transmission by repeatedly using the above operator.[14] For example, repeating the transmission/collection/retransmission process n times gives

$$\mathbf{w}(\omega) = |\mathbf{C}(\omega)|^{2n}\mathbf{u}(\omega). \tag{4}$$

Although this process improves the accuracy in terms of spatial and temporal focusing by allocating greatest effort to the most effective modes for transmission, it severely degrades the quality of the reproduced signal.

The application of the time-reversal procedure will be demonstrated in Sec. IV.

B. An introduction to optimal inverse filtering

Prefiltering is a common technique used with linear systems to improve the response of a system, for example, compensating for a nonuniform transducer response. When the prefilter completely compensates for the plant transfer function (including transducer responses), such that the receiver signal is equal to the source input signals, then the process is referred to as *inverse filtering* or *equalization* and is commonly used for auralization in rooms and the free field.[36–39]

For example, Fig. 2 shows a typical arrangement where there is some set of filters (one per source) which operates on the input signal prior to transmission through the plant. Clearly, if the prefilter matrix $\mathbf{H}$ is set equal to the inverse of the plant matrix (assuming $\mathbf{C}$ is square), such that

$$\mathbf{H}(\omega) = \mathbf{C}^{-1}(\omega), \tag{5}$$

then $\mathbf{w}(\omega) = \mathbf{C}(\omega)\mathbf{C}^{-1}(\omega)\mathbf{u}(\omega) = \mathbf{u}(\omega)$ an ideal waveform is recovered. If the system is nonsquare it is necessary to use the least-squares approximation

$$\mathbf{H}(\omega) = [\mathbf{C}^H(\omega)\mathbf{C}(\omega)]^{-1}\mathbf{C}^H(\omega). \tag{6}$$

However, in general it is not possible to realize a causal and stable inverse filter using the above approach.[40] This is because of two main difficulties.

(1) Minimum phase zeros close to the unit circle when inverted give rise to inverse filters of long duration in forward time. This results in filters that ''wrap around'' when the time-domain filter is Fourier synthesized.

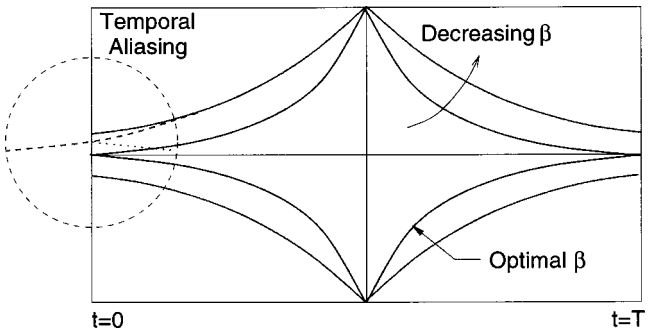


FIG. 4. Schematic illustrating the effect that β has on the inverse filter. The time window shows the duration of the inverse filter. Too-small β leads to temporal aliasing outside the time window, leading to the signal folding back in on itself in the region circled.

(2) Nonminimum phase zeros close to the unit circle when inverted give rise to inverse filters of long duration in reverse time. This can create *noncausal* filters which will also wrap around.

There are a variety of techniques used to overcome the difficulties of realizing causal and stable filters defined by Eq. (5) and are well summarized by Kahana *et al.*[41]

1. Optimal filtering using Tikhonov regularization

The frequency-domain formulation of Eqs. (5) and (6) do not account for the *temporal aliasing* or *wrap around* that can occur when synthesizing a time-domain representation. The temporal aliasing happens when the filter time constant is too low and the filters ring beyond the end of the time window and fold back. This is illustrated in Fig. 4. This leads to discontinuities at the beginning and end of the filter much like aliasing in the frequency domain when sampling time series without antialiasing filters. The problem of temporal aliasing becomes an important issue when the zeros of the transfer function close to the unit circle are inverted to produce unstable poles in the inverse filter.[22]

Temporal aliasing is sometimes overcome by increasing the number of time samples[42] of the inverse filter impulse response function (IRF). This can be achieved through the use of longer time samples or zero padding of the measured IRFs in order to accommodate the filter decay constant. This technique, however, can result in finite impulse response (FIR) filters with a very large number of taps.

An alternative is to use *regularization*,[43] a process often used for solving ill-conditioned and singular linear systems, which limits the duration of the filter IRFs. *Tikhonov regularization* has been used here although any suitable regularization technique, such as singular value discarding,[17] could have been used instead. The inverse (least-squares) matrix with Tikhonov regularization is given by[25,37]

$$\mathbf{H}(\omega)=[\mathbf{C}^H(\omega)\mathbf{C}(\omega)+\beta(\omega)\mathbf{I}]^{-1}\mathbf{C}^H(\omega), \tag{7}$$

where $\beta(\omega)$ is the (frequency-dependent) *regularization coefficient*. The diagonal *regularization matrix* $\beta\mathbf{I}$ simply limits the singular values of $\mathbf{C}^H\mathbf{C}+\beta\mathbf{I}$ to a minimum value of β. Clearly, if $\beta=0$, then Eq. (7) reduces to Eq. (6).

Using inverse filters to focus ultrasonic waves, Tanter *et al.*[17] used regularization to improve the invertibility of the plant matrix $\mathbf{C}$ and reduce the sensitivity of the inverse filters to noise. However, since they were using essentially a monochromatic source, the formulation was completely performed in the frequency domain and consequently they failed to realize or at least acknowledge the difficulties of Fourier synthesizing the time-domain form of the inverse filters from the inverse matrix $\mathbf{H}$. Kirkeby *et al.*,[25,26] working over a broad bandwidth also used regularization, not only to improve invertibility of the *inverse filter operator* ($\mathbf{C}^H\mathbf{C}$), but also to limit the duration of the inverse filters and thus reduce the temporal aliasing (see Fig. 4). If $\beta(\omega)$ is set too large then the optimal filter matrix $\mathbf{H}(\omega)$ no longer resembles $\mathbf{C}^{-1}(\omega)$, and if $\beta(\omega)$ is set too small then wraparound is significant. The objective is to find the optimal value of $\beta(\omega)$ to improve the invertibility of the matrix and reduce aliasing without severely degrading the performance of the inversion procedure. It will be seen in Sec. IV A that there exists an optimal value for β that minimizes the least-squared error between the reproduced and desired signals.

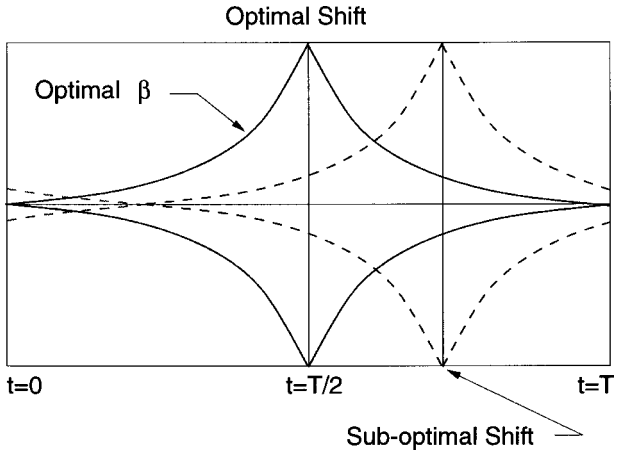


FIG. 5. Schematic illustrating the effect that cyclic shifting has on the inverse filter. The time window shows the duration of the inverse filter. The optimal delay is approximately given by half the time window.

2. Causality and cyclic shifting

Causality is generally achieved by introducing a phase delay as is the case with the time-reversal filters. This can be seen in Eq. (3), which gives $\mathbf{w}(\omega)=\mathbf{C}(\omega)\times\mathbf{C}^{-1}(\omega)e^{j\omega t_0}\mathbf{u}(\omega)=e^{j\omega t_0}\mathbf{u}(\omega)$. The required delay depends on the proximity of the nonminimum phase zeros of the plant matrix to the unit circle.

The issue of causality is closely associated with temporal aliasing of the inverse filters in the time domain. Aliasing often happens when the peak of the inverse filter IRF is too near the limits of the time window, as illustrated in Fig. 5. Aliasing may be partially overcome by *cyclic shifting*[25] the IRF so that it is centrally located in the time window.

There are several ways in which the appropriate cyclic shift can be calculated. One way is to locate the time at which the maximum level in the IRF is received. An additional delay can then be added by cyclic shifting the IRF so that the peak of the IRF resides in the middle of the time window. Alternatively, in the frequency domain, one can cal-

TABLE I. Properties of the shallow water channel model.

Layer	Depth (m)	Wave speed (m/s)		Attenuation (dB/λ)		Density (kg/m^3)
		Comp.	Shear	Comp.	Shear	
Water	0–20	1512	0	0	0	1000
	20–40	1512–1507.5	0	0	0	1000
	40–145	1507.5–1508.5	0	0	0	1000
Sediment	145–150	1470–1490	0	0.054	0	1500
Bedrock	150–∞	1600	0	0.135	0	1800

culate the mean group delay over the bandwidth of interest $(\omega_{m_2}-\omega_{m_1})$ for all sensors. This is given by

$$\overline{\tau}_{gd}=\frac{1}{m_2-m_1+1}\sum_{\omega_m=\omega_{m_1}}^{\omega_{m_2}}\tau_{gd}(\omega_m), \tag{8}$$

where $\tau_{gd}(\omega)=-\partial\phi(\omega)/\partial\omega$ is the group delay. For optimal filtering this should be approximately half the duration of the time sample. The transfer functions may then be phase shifted by multiplying the difference between then measured group delay and the desired delay (half the time window).

III. SHALLOW-WATER MODEL

The shallow-water environment in the Giglio basin off the west coast of Italy provided the basis for the shallow-water model, since this is the likely future site for the experimental validation of the technique. The configuration for the numerical experiment is shown schematically in Fig. 1 and is similar to the experimental configurations used by Kuperman *et al.*[12] and Song *et al.*[14] for the phase-conjugation experiments in the ocean.

The shallow-water environment has a channel of approximately constant depth of 145 m out to a range of up to 10 km. The sound speed of the water channel varied from 1508 to 1512 m/s as shown in Fig. 1. A sedimentary layer of approximately 5 m separates the water channel and the bedrock half space. The data are summarized in Table I.

The transfer functions of the system described in Table I were calculated using the fast-field modeling package OASES,[44] which is an upgraded version of SAFARI Version 3.0 distributed by SACLANTCEN.[45] The corresponding time series were then calculated using Fourier synthesis.

IV. SIMULATION RESULTS

All simulations have been conducted on the data from the numerical model described in Sec. III. The SRA consisted of 16 equispaced elements, with a spacing of 5 m from a depth of 50 to 125 m. Two receiver ranges were investigated, namely 2 and 5 km. The frequency range was dc to 1 kHz. Since one of the advantages of inverse filtering is the ability to remove amplitude aberrations, all the simulations have been performed using an impulse with a duration of 0.5 ms, as this provides energy over the entire bandwidth of the transfer functions under investigation.

A. Optimization

The following section discusses the issues associated with the calculation of an ''optimal'' regularization coefficient, β_{opt}, so that the inverse filters may then be calculated using Eq. (7).

When using the frequency-domain optimization[46] with an SRA of two or more elements, it is theoretically possible to determine an optimal value of β which will result in a perfect transfer function in the frequency domain (provided no common zeros exist in the plant matrix[47]). For example, the composite frequency response between the outer two elements of the SRA and receiver at a range of 2 km using frequency-domain optimization is illustrated in Fig. 6.

As discussed in Sec. II B, frequency-domain optimization does not consider the time-domain filter response and often leads to aliasing of the inverse filter IRFs in the time domain. This is seen in the received signal shown in Fig. 6. Comparing Fig. 6 against Fig. 7 with an optimal β, it can be seen that in addition to the desired impulse at approximately $t=2$ s there are other spurious impulses at approximately $t=1.4$ s and $t=3.4$ s, all of which are associated with temporal aliasing.

Therefore, although the calculation of the inverse filters is done in the frequency domain because of computation efficiency, when optimizing the filters the Fourier-synthesised time-domain representation must be considered. Since the inverse filters are optimal in a least-squares sense, it is logical that the mean-squared error (MSE) in the reproduced signal is the desired cost function by which to opti-

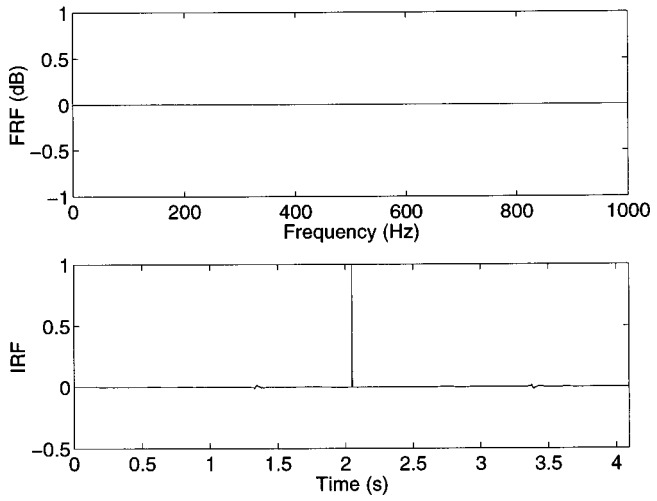


FIG. 6. Frequency response function and equivalent impulse response function showing the effect of a small regularization coefficient, $\beta \ll \beta_{\text{opt}}$.

mize regularization coefficient, i.e., consistent with the design of a Weiner filter.[48] The normalized MSE is defined as

$$\mathrm{MSE}=\frac{\Sigma_1^{n_w}(u(t+t_0)-w(t))^2}{\Sigma_1^{n_u}(u(t))^2}, \tag{9}$$

where n_u and n_w are the number of samples in the desired and received signals, respectively, and t_0 is the appropriate propagation delay. Equation (9) requires prior knowledge of the desired signal before the optimal regularization coefficient may be determined.

The optimization of β cannot be solved for directly but must be determined using an appropriate search routine. Starting with the knowledge that the singular values of $\mathbf{C}^H\mathbf{C}$ are directly related to the length of the inverse filters, it is then possible to use the singular values to provide a range of regularization coefficients from which an optimal may be chosen. The singular values of $\mathbf{C}^H\mathbf{C}+\beta\mathbf{I}$ when using Tikhonov regularization are given by

$$\hat{\sigma}_i(\omega)=\sigma_i(\omega)+\beta_i(\omega), \tag{10}$$

where σ_i is the ith singular value of the Hermitian matrix $\mathbf{C}^H\mathbf{C}$ and $\beta_i=\beta$ is the ith singular value of the Hermitian matrix $\beta\mathbf{I}$. It should be noted that Eq. (10) is a particular property of Hermitian matrices and does not hold for all matrices.

For optimal filtering in a least-squares sense, the regularization coefficient must clearly be less than the norm (maximum singular value) without regularization, $\sigma_{\max}$, over the entire frequency range of interest, i.e., $\beta_{\mathrm{upper}}=\sigma_{\max}$. Therefore, the norm defines the upper bound of the search space. However, the lower bound is less well defined. It is both a function of the number of time samples, the amount of gain required to overcome the zeros in the plant transfer function matrix, and also the signal-to-noise ratio (SNR). The expression for the optimal filter matrix in the presence of noise is derived in the Appendix, and is given by

$$\mathbf{H}=[\mathbf{C}^H\mathbf{C}+E[\mathbf{N}^H\mathbf{N}]+\beta\mathbf{I}]^{-1}\mathbf{C}^H, \tag{11}$$

where $\mathbf{C}$ is the plant matrix at the receiver in the absence of noise and $\mathbf{N}$ is the associated stochastic noise matrix. The additional noise term $E[\mathbf{N}^H\mathbf{N}]$, which incidentally is diagonal when the individual elements on $\mathbf{N}$ are uncorrelated, acts to regularize the matrix to be inverted and may be used to define the lower bounds of the search space. In practice, the SNR of most transfer functions are not better than 40 dB (1 to 100), and therefore the optimal regularization coefficient need not be less than one ten-thousandth of the norm, i.e., $\beta_{\mathrm{lower}}=(1/100)^2\sigma_{\max}=10^{-4}\sigma_{\max}$.

Consider the system described in Sec. III. Figure 7 shows the MSE between the transmitted signal (an impulse) and received signal plotted against regularization coefficient for a receiver at a depth of 85 m and a range of 2 km. It can be seen that the MSE between the reproduced signal and the desired signal is at a minimum when the normalized regularization coefficient, $\hat{\beta}=\beta/\sigma_{\max}$, is approximately 1%. It was found that for all the cases considered the normalized regularization coefficient was always of the order of 1%. One interesting feature of regularization is that the peak amplitude of the transmitted pulse is slightly reduced, as can be seen by comparing the signal transmission without regularization (Fig. 6) and with regularization (Fig. 7). This arises because there is insufficient gain in the inverse filters, particularly at low frequencies, to completely compensate for the plant matrix. Consequently the composite frequency response function suffers from dips, which leads to a reduced peak amplitude in the impulse.

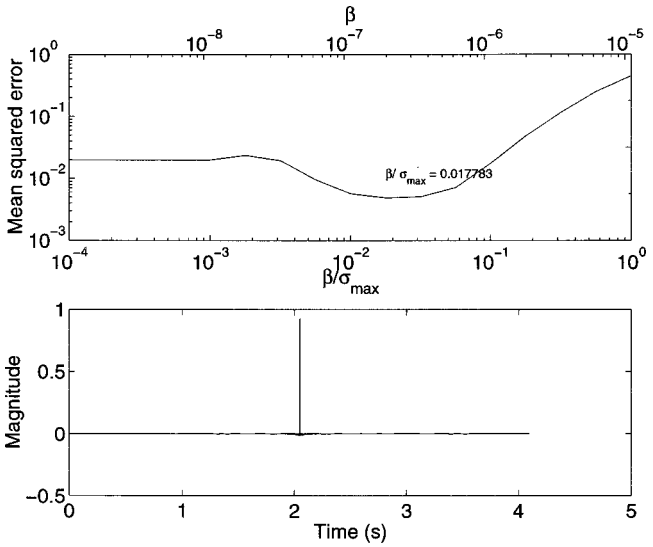


FIG. 7. Mean-squared error vs regularization coefficient and optimized reproduced signal using Tikhonov regularization.

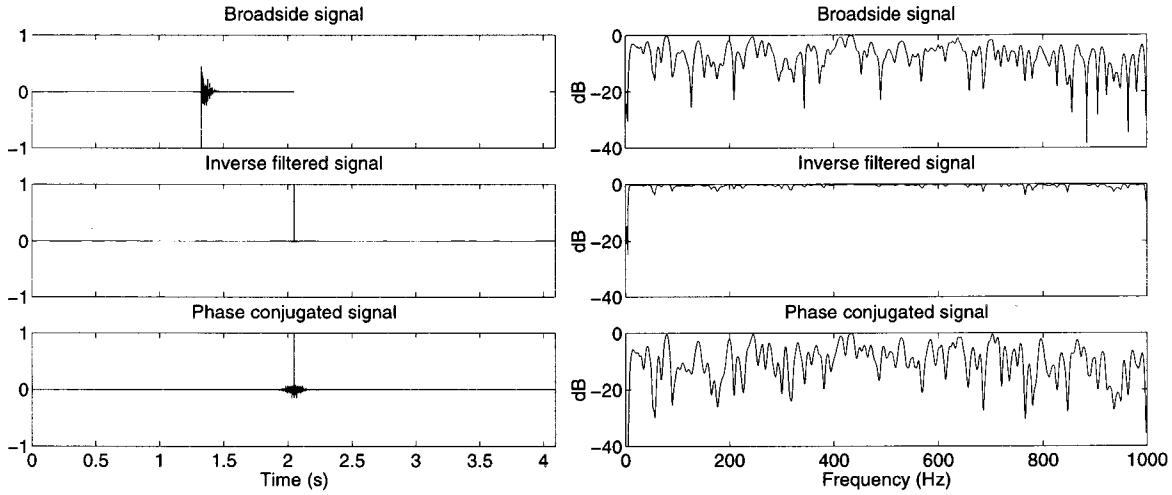
All the following simulations have used both optimization of β and cyclic shifting.

B. Temporal focusing

Consider the transmission problem discussed previously where only the two outer elements in the SRA are active and a single receiver at a depth of 85 m and a range of 2 km. Figure 8(a) shows the impulse response function at the desired location for a directly transmitted (broadside beamformed) signal, a phase-conjugated signal, and an optimally filtered signal. The respective MSEs of the three techniques are 252%, 35.4%, and 0.9%. It should be noted that when calculating the MSE using Eq. (9) it was necessary to provide equalization gain for the phase-conjugate and broadside transmissions, so therefore the maximum signal strength was normalized to unity. It becomes immediately apparent that both phase conjugation and inverse filtering improve the quality of the impulse response function and greatly enhance temporal focusing, particularly when compared to the directly transmitted signal. However, the MSE for time reversal (TR) with only two elements is quite high.

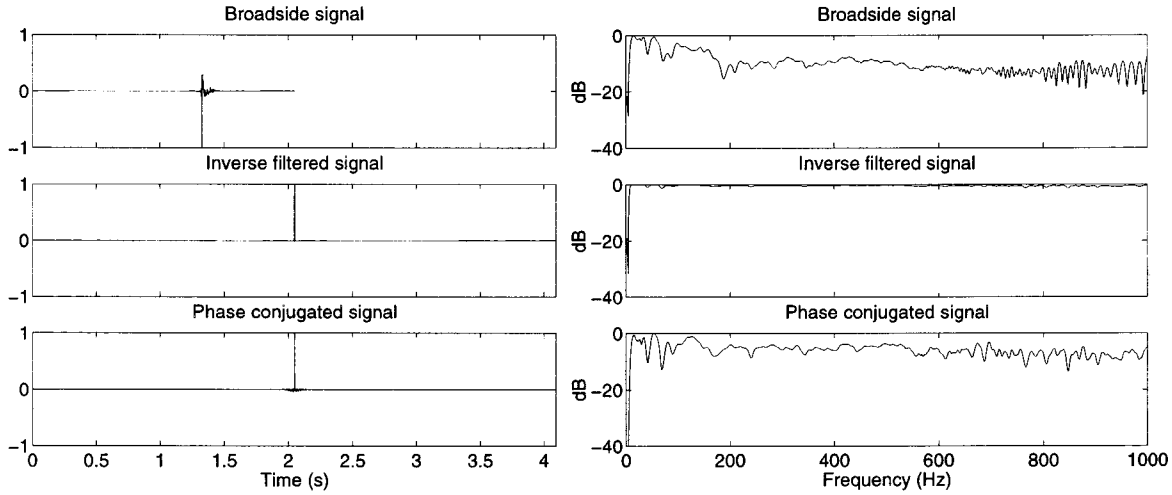
Figure 8(b) is a plot of the frequency response functions corresponding to the impulse response functions in Fig. 8(a). Although phase conjugation has produced what looks like an impulse, it does not have a uniform magnitude response, unlike the inverse filter. Note that with only two elements the magnitude response with time reversal is worse than that of the broadside transmission as predicted.

When all 16 elements in the SRA are active, then the difference in performance between TR and IF is small, as



(a) 2 elements : IRF's

(b) 2 elements : FRF's



(c) 16 elements : IRF's

(d) 16 elements : FRF's

FIG. 8. Impulse response functions showing temporal focusing and corresponding frequency response functions when using either 2 SRA elements or all 16.

seen in Figure 8(c). The respective MSE of the three techniques is now 239%, 6.9%, and 0.6%. The corresponding frequency response functions are plotted in Fig. 8(d). As the number of elements are increased the amplitude aberration is reduced as predicted in Sec. II A, with the aberration for TR now less than that of the broadside transmission. It should be noted that even with only two elements in the SRA, transmission using IF results in a lower MSE than a TR array with 16 elements. Therefore, IF can offer a substantial reduction in the hardware complexity and costs of a practical system when compared to TR.

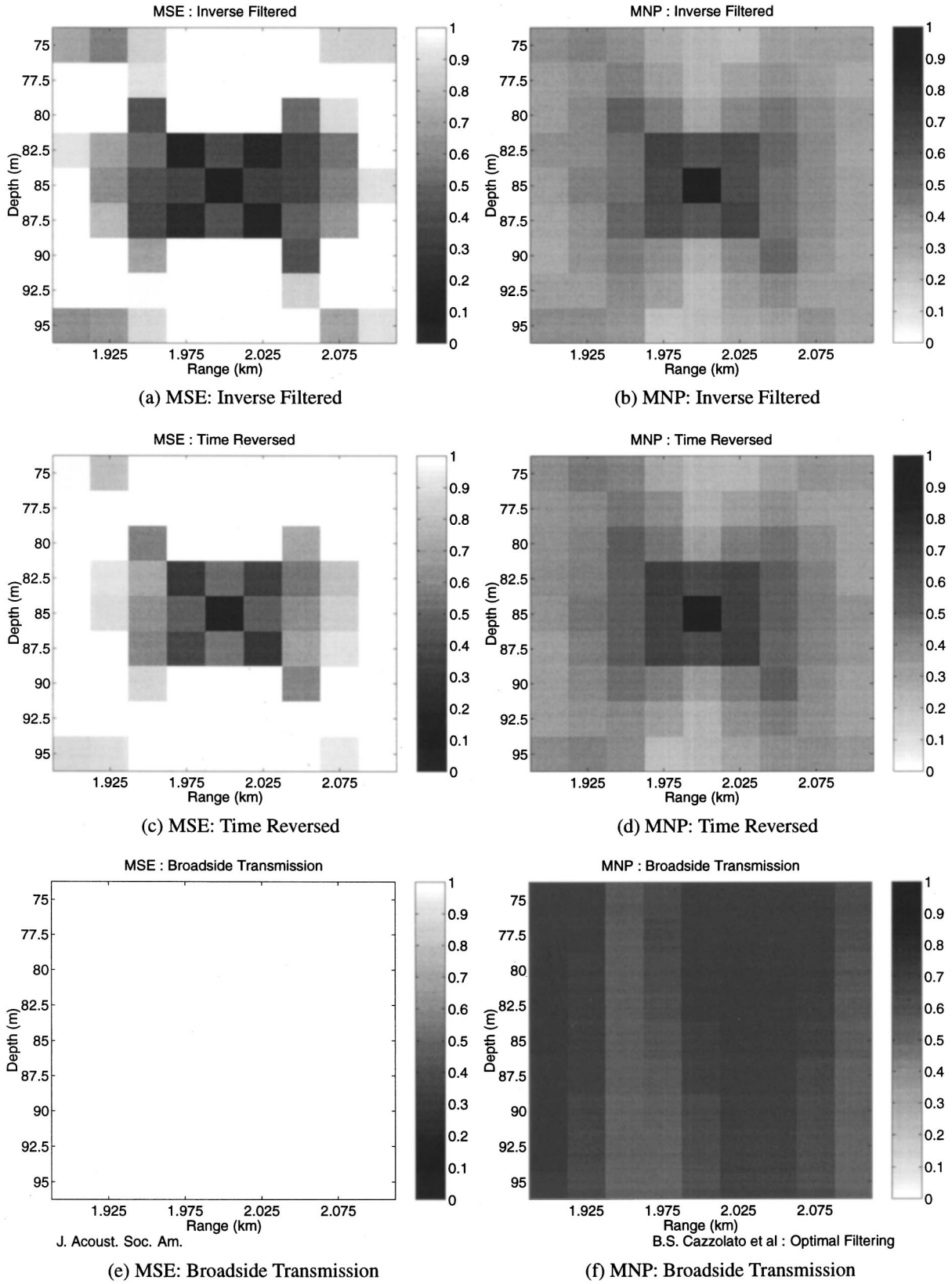
The energy loss, ΔE, between the SRA and receiver is also of interest. The ratio of energy between coherent and noncoherent addition from an array of sources is approximately given by $10\log_{10}(N)$, where N is the number of elements in the array. Since the broadside transmission is only partially coherent, the gain between time-reversal or inverse filtering and broadside transmission should be something less than a fully coherent transmission gain (12 dB for 16 elements). In fact, $10\log_{10}(\Delta E_{\text{inverse}}/\Delta E_{\text{broadside}})=5.6\,\text{dB}$ and $10\log_{10}(\Delta E_{\text{time reversal}}/\Delta E_{\text{broadside}})=6.3\,\text{dB}$ for all 16 elements. The reason that the inverse filter gain is slightly less than the time-reversal gain is because inverse filtering attempts to compensate for frequencies where transmission is poor, whereas the opposite occurs with time reversal. This difference grows with fewer elements in the SRA.

C. Spatial focusing

1. Background

It is important to know how sensitive the inverse filtering (and time-reversal) technique is to changes in the locations of the transducers from when the transfer function is estimated and when the signal is retransmitted. It will be seen that in fact the zone of focus for the two techniques is the same for practical purposes.

Cassereau and Fink[49] state that at a given frequency, for a free-field environment with sensors completely surrounding the receiver, the maximum available resolution for the



(a) MSE: Inverse Filtered

(b) MNP: Inverse Filtered

(c) MSE: Time Reversed

(d) MNP: Time Reversed

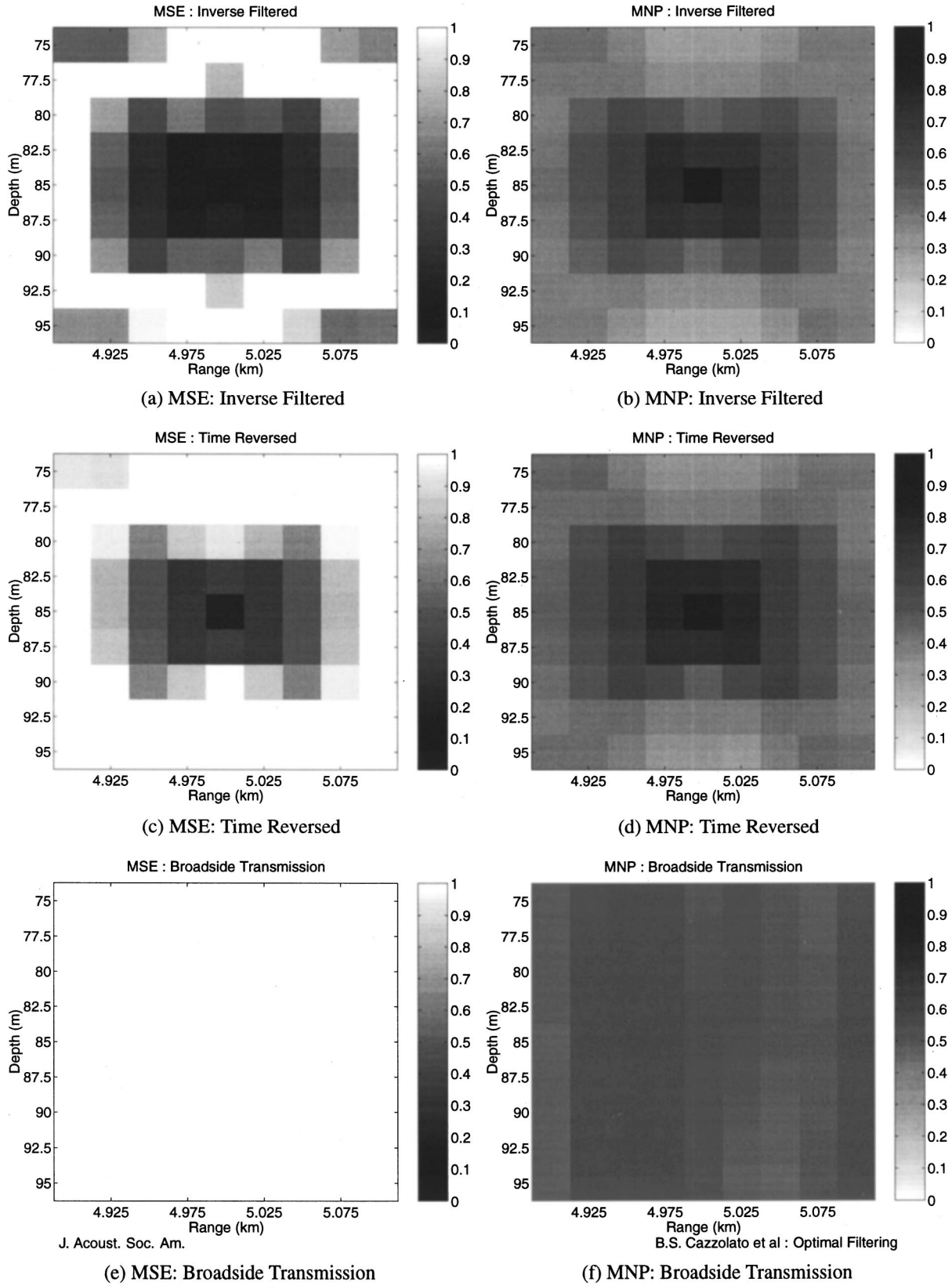
(e) MSE: Broadside Transmission

(f) MNP: Broadside Transmission

FIG. 9. MSE and MNP for a receiver at a range of 2 km.

focusing process by time reversal is $\lambda/2$. When a planar array is used in a free-field environment, the zone of the focusing increases along the axis between the source plane and the receiver as the aperture shrinks.[49]

In a highly reverberant environment the situation is slightly different because the transfer functions between two points are defined by the modal response of the system. Tolstoy and Clay[33] showed that the zone of focus for time reversal is related to the cross-correlation function between the impulse response function of the desired focus point and



(a) MSE: Inverse Filtered

(b) MNP: Inverse Filtered

(c) MSE: Time Reversed

(d) MNP: Time Reversed

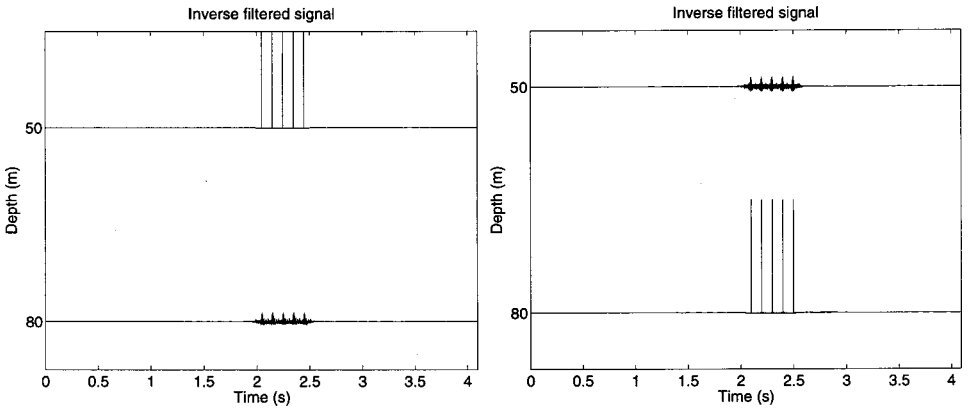
(e) MSE: Broadside Transmission

(f) MNP: Broadside Transmission

FIG. 10. MSE and MNP for a receiver at a range of 5 km.

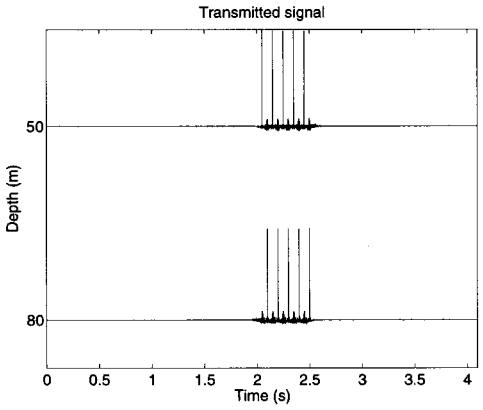
neighboring points. This cross correlation of acoustical pressures is a function of the vertical wave number, γ, and the horizontal wave number, κ, where $\gamma^2+\kappa^2=k^2=\omega^2/c_0^2$.

In fact, the cross-correlation function for an isotropic medium with modes m and n (known as the *mode-interference term*) is proportional to $\sin(\gamma_m z)\sin(\gamma_n z)\cos((\kappa_m-\kappa_n)r)$, where the subscripts refer to the mode numbers, z is the depth of the receiver, and r is the range.[33] So, clearly a change in receiver or source depths sufficient to increase either $\gamma_m z$ or $\gamma_n z$ by $\pi/2$, or greater, causes large changes in



(a) Single receiver at 50m

(b) Single receiver at 80m



(c) Receivers at 50m and 80m

FIG. 11. Two-receiver transmission, which each element acting independently and cross talk present.

the amplitudes and phases of the mode-interference terms. Tolstoy and Clay[33] show that since $(\kappa_m - \kappa_n)$ is generally small compared to k at long ranges, then the term $\cos((\kappa_m - \kappa_n)r)$ requires large changes in range for a change in the mode-interference term.

The tighter vertical focusing compared to radial focusing with time-reversal systems in acoustic waveguides has been confirmed experimentally for a single-element SRA[4,35] and for a multiple-element SRA.[12]

Losses due to absorption and scattering are detrimental to phase conjugation since they cause attenuation of higher-order modes, yielding a less-tight focus than would be possible without loss.[12] Furthermore, this blurring will increase as the range between the source and the array increases owing to the strong range and mode number dependence of the attenuation. This will be seen to be the case with inverse filtering as well.

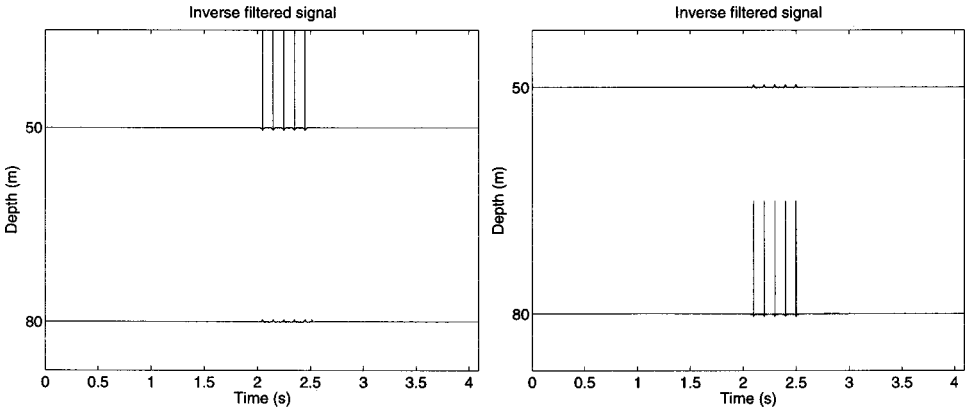
2. Results

Using the numerical model of the shallow-water waveguide described in Sec. III, the zone of focus has been measured both vertically and horizontally for two ranges; one at a range of 2 km, the other at 5 km. The 16 elements in the SRA were active, with the active receiver at a depth of 85 m.

Two performance metrics have been used. The first is the MSE, which provides a measure of the quality of the signal. The second is the maximum normalized pressure (MNP), which is the ratio between the amplitude of the signal at some location to the amplitude of the signal at the desired location and is a quantitative metric for time reversal used by the majority of authors.[9,11,15,34,49–51] It should be noted, however, that since the application of the time reversal for these authors was predominantly medical ultrasound, the objective was to provide the greatest peak amplitude at the receiver, and signal reproduction quality was of secondary importance. For communication systems it is necessary to consider both of these factors.

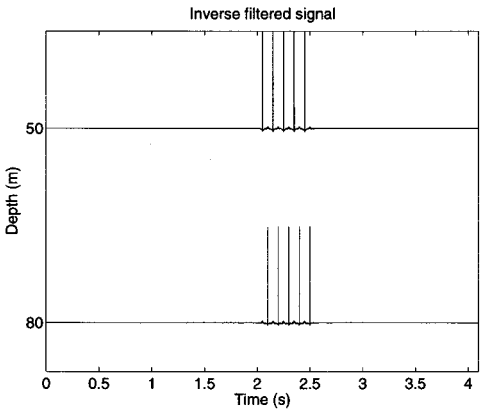
Figures 9 and 10 show the MSE and MNP for the three transmission filters for a range of 2 and 5 km. The MSE was calculated over the entire duration of the transmission, which was 4 and 10 seconds for the 2- and 5-km ranges, respectively. It can be seen that in terms of pressure amplitude (MNP), time reversal and inverse filtering are essentially equivalent, although the inverse filtering exhibits slightly tighter focusing. The location of the peak amplitude with broadside transmission is purely dependent on the modal response of the waveguide. As expected, the size of the zone of coherence increases with range as modal stripping removes the high-order modes with large grazing angles.

The size of the low MSE zone is closely linked to the size of the high-amplitude zone. Furthermore, although the MNP plots when using time reversal are almost identical to those when using inverse filtering, the MSE are markedly different, with the zone of low MSE being much larger for



(a) Single receiver at 50m

(b) Single receiver at 80m



(c) Receivers at 50m and 80m

FIG. 12. Two-receiver transmission, with the elements acting together and cross talk removed.

the latter. The MSE for broadside transmission is poor at 2 and 5 km.

It is not at all surprising that both phase conjugation and inverse filtering offer a similar level of spatial focusing since both techniques attempt to remove the phase response and associated time delays from the plant matrix.

D. Multichannel transmission

The desire in communication systems is always to increase bandwidth and reduce error rates. Since it has been shown that in theory it is possible to transmit different signals to individual receivers using inverse filters, then one may use this technique to either increase the bandwidth by effectively creating different channels, or reduce the error rates through increased redundancy.

Figure 11(a) shows five 1-ms pulses with a 100-ms duty cycle transmitted from the SRA to a receiver at a depth of 50 m using optimal filtering. Figure 11(b) shows the same five 1-ms pulses delayed by 50 ms and then transmitted from the SRA to a receiver at 80 m using a new set of optimal filters. If these two filter sets are added together (along with the 50-ms group delay on receiver 2), then it is possible to transmit a signal to both receivers simultaneously. The received signal is then equal to the superposition of the two independent transmissions, as seen in Fig. 11(c). It is clear there is some cross talk between the receivers, particularly at the low frequencies. Were less active sources used in the SRA, then cross talk would be greater than shown here with 16 elements. This simple example illustrates the limitation of designing filters for multiple receivers using the single-channel approach.

1. Cross-talk cancellation networks

When cross talk between ''channels'' is neglected, as was done previously, then at low frequencies when the two receivers are poorly separated, contamination from one channel to the other occurs. This may be overcome by integrating cross-talk cancellation networks into the inverse filters.

Kirkeby *et al.*[38] implemented a technique to remove cross talk for binaural systems and it is straightforward to show that this is achieved when $\mathbf{H}(\omega)$ is given by Eq. (5). Therefore, to generate the cross-talk filters, the receivers cannot be considered independently, as was done in the previous section but must be considered as a whole.

Repeating the two receiver simulations of Fig. 11 but this time calculating the cross-talk cancellation network, inverse filtering has greatly reduced the cross talk between the elements, as seen in Fig. 12.

Difficulties arise when the $\mathbf{C}^H\mathbf{C}$ matrix becomes close to singular. Physically this represents two or more receivers with similar array–receiver transfer functions and will arise when the receiver elements are closer than the limits imposed in Sec. IV C. Similar limitations are experienced when using cross-talk cancellation networks with binaural sound

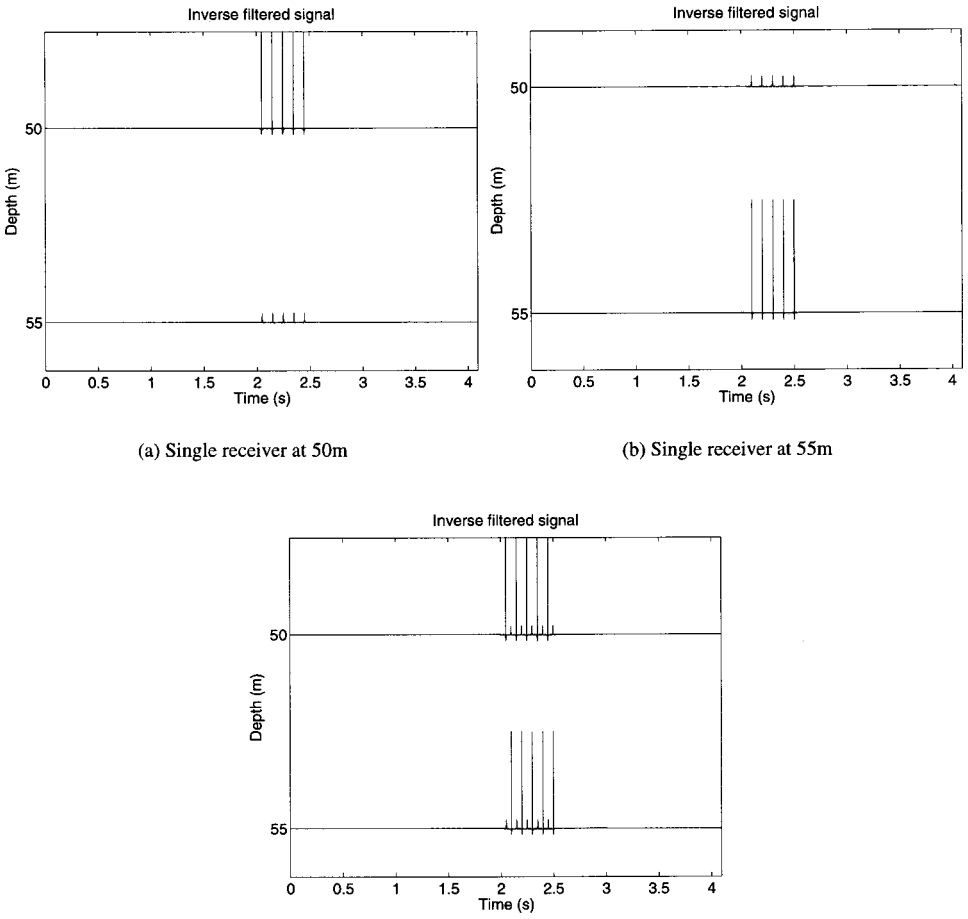


FIG. 13. Two-receiver transmission, with the elements acting together and cross talk removed.

reproduction systems.[38] Figure 13 shows the difficulties associated with cancelling cross talk between two adjacent receivers, one at a depth of 50 m, the other at 55 m. It is evident that there is some cross talk between the receivers.

E. Number of bits in A/D D/A

In practice, digital input–output systems are limited in resolution by the number of bits from the analog to digital (A/D) and digital to analog conversion (D/A). The cost of such hardware is typically proportional to the number of bits available. Therefore, as practical systems are often a compromise between the desired dynamic range and cost, it is very important to have an understanding of what constitutes an adequate dynamic range for a particular application.

Derode *et al.*[52] compared the performance of a 1-bit and 9-bit phase conjugate system when transmitting pulses from a 128-element array through a complex multiple-path environment. They found that when the signal was digitized over 1 bit (either 1 or -1), both temporal and spatial resolutions remained unchanged, signal levels were slightly higher, and the signal to noise ratio was slightly lower.

It is also of interest to see the impact that quantization has on the inverse filters. Since quantization is a nonlinear and noninvertible process,[48] the operator must be applied to the time-domain signals to emulate what happens in practice. Using a one's complement bipolar quantization of a signal, x, assuming saturation at full scale, the digitized signal is given by[48]

$$\hat{x} = \frac{\max(|x|)}{2^{(b-1)}}\,\mathrm{round}\left(\frac{x}{\max(|x|)}\,2^{(b-1)}\right), \tag{12}$$

where b is the number of quantization bits.

Consider the waveguide used previously, with a 16-element vertical line array transmitting a 0.5-ms pulse to an active receiver at a depth of 85 m and a range of 2 km. In Fig. 14 the time series of the filtered signals for a variety of quantization levels on the A/D stage are plotted; 1 bit, 2 bit, 4 bit, and 8 bit. Although not shown here, machine precision (53 bit) and 16 bit (MSE=1.04%) are essentially identical and both show very little change from 8 bit. The 1-bit quantization (with permissible signal levels of $+1$ or -1) has produced extremely poor results and is significantly worse than seen by Derode *et al.*[52] This is probably due to the differences in the physical systems under consideration, where the number of transmission paths for the system considered by Derode *et al.*[52] was several orders of magnitude greater than for the waveguide considered here. The addition of another bit (permissible signal levels -1, 0, and 1) has greatly improved the results both qualitatively and quantitatively. With 4 bits the MSE is only twice that without any quantization, and 8-bit quantization is essentially equivalent to no quantization, and obviously adequate for this system and signal type.

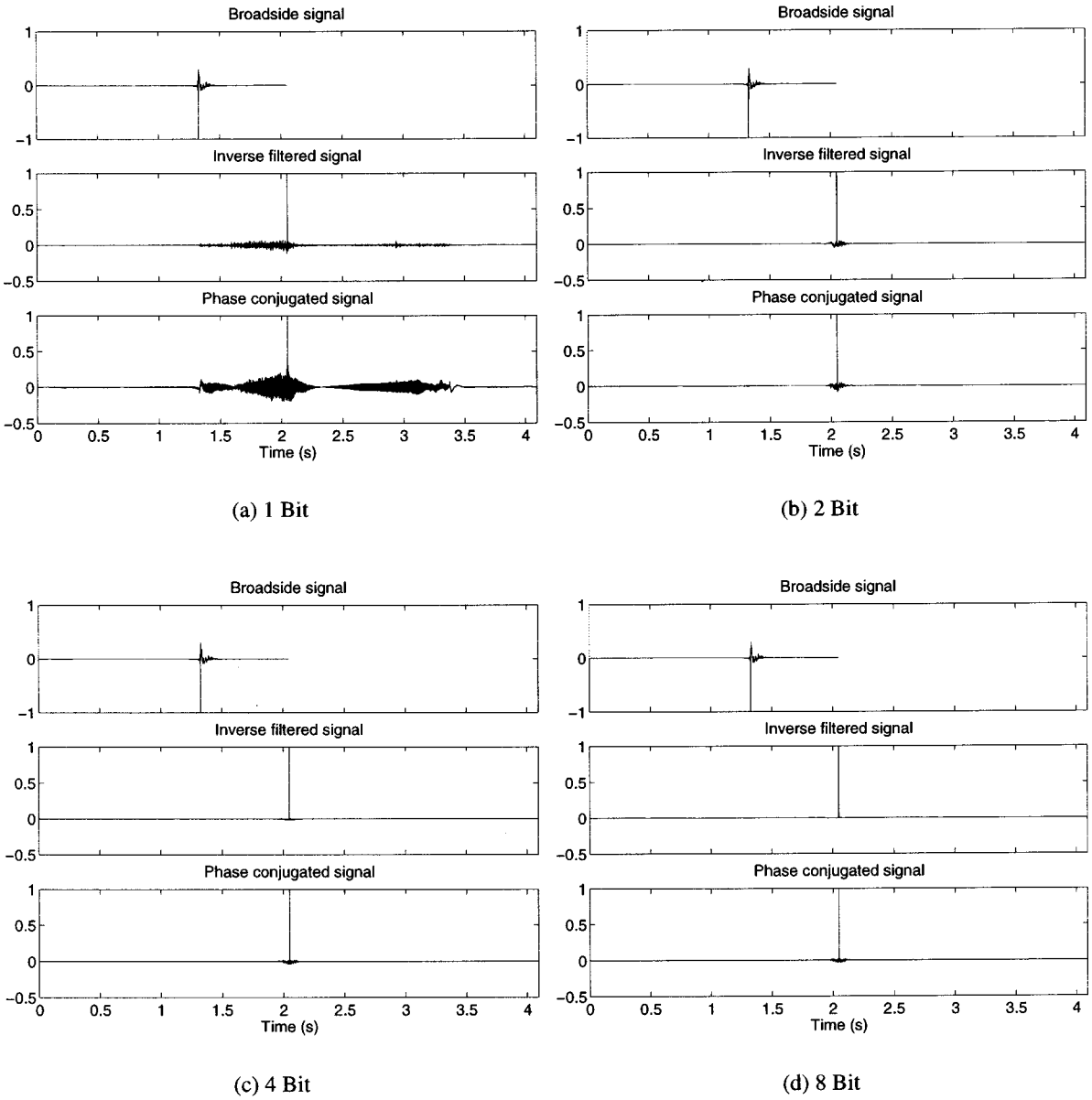


FIG. 14. The effect of A/D (input) quantization on the transmission of a 1-ms pulse.

Quantization was then applied only to the output stage of the system to emulate the D/A conversion. The results are shown in Fig. 15. It can be seen that the order of magnitude of the quantization error is similar for both D/A and A/D. The results for quantization of the phase-conjugated signal are equal for both D/A and A/D due to the symmetric nature of the transmission.

The results of the quantization simulations are summarized in Table II for the three transmission filters; broadside beamforming, phase conjugation, and optimal filtering.

The use of a pulse to test the effect of quantization may not be ideal since a pulse is well represented by low-resolution quantization, even 1 bit. A more challenging and practical signal is a chirp. The tests were repeated for a chirp spanning the entire frequency range (0–1 kHz), where it was found that the MSEs of the time-reversed and inverse-filtered signals were approximately the same as for the impulse.

F. Effects of noise

It appears that the inverse filters are inherently robust to contaminating noise. The expression for the optimal filter matrix in the presence of noise (in the Appendix) shows that uncorrelated noise has very little impact on the inverse filters when the SNR is high. It was found here (and elsewhere[43]) that low levels of correlated or uncorrelated broadband noise improve the invertibility of the plant matrix. However, when the SNR of the plant transfer matrix **C** approaches unity, the noise simply acts to mask the transfer function at these frequencies. The gain of the inverse filters at these frequencies is subsequently very low, and therefore the noise is effectively removed. A detailed analysis on the effects of noise on signal quality is well beyond the scope of the current paper.

V. CONCLUSIONS

Two signal-compression techniques have been investigated, namely phase conjugation and optimal inverse filter-

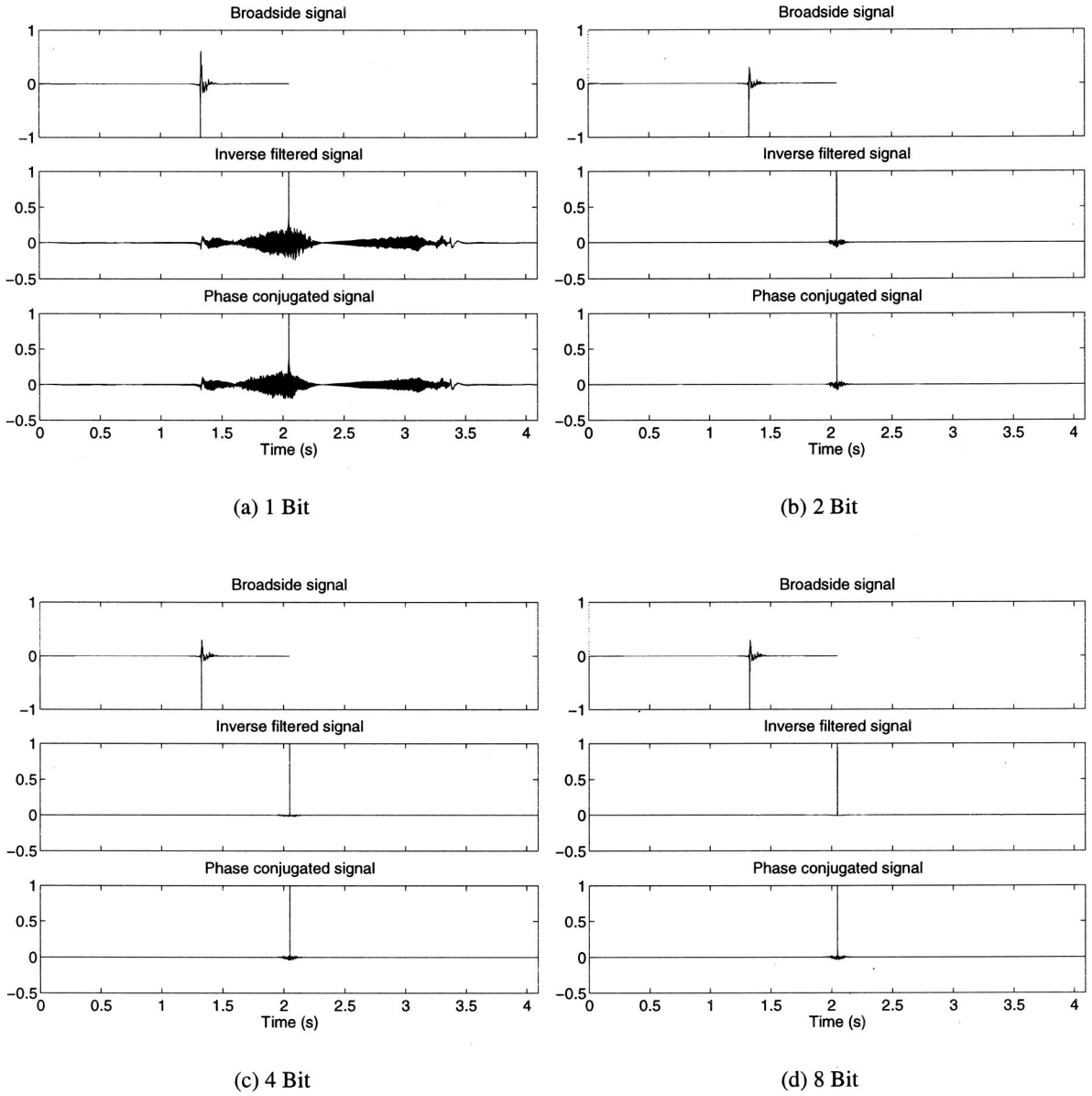


FIG. 15. The effect of D/A (output) quantization on the transmission of a 1-ms pulse.

ing. Both of the techniques remove the phase aberrations from the composite transfer function. However, only inverse filtering acts to remove the magnitude aberration from the composite transfer function. With only a single source, phase conjugation doubles the original magnitude aberration, but as the number of elements in the source array are increased the amplitude aberration is reduced.

Inverse filtering allows significantly fewer elements in the SRA compared to TR for the same reproduced signal quality.

The time-reversal technique has the advantage that it is extremely simple to implement and is inherently stable, since the process simply involves recording the impulse response function and then retransmitting in reverse. In this respect it

TABLE II. The effect of quantization on the mean-square error (%) when transmitting a 1-ms pulse.

Stage	Transmission technique	Number of bits				
		1	2	4	8	∞
Input	Broadside transmission	239	239	239	239	239
	Phase conjugation	1767	13.9	7.1	6.9	6.9
	Inverse filtering	72.9	11.2	1.0	0.6	0.6
Output	Broadside transmission	588	239	239	239	239
	Phase conjugation	1767	13.9	7.1	6.9	6.9
	Inverse filtering	1787	10.9	1.6	0.6	0.6

is very different from the inverse filtering technique.

Regularization of the inverse filter operator using Tikhonov regularization was investigated and found to be effective at not only improving the condition number (invertibility) of the operator but also in reducing temporal aliasing.

From the simulations it appears that the optimal regularization coefficient is relatively independent of the physical configuration of the waveguide and source/receiver arrangement as well as the FFT size. As a first approximation, the optimal regularization coefficient β_{opt} should be about one-hundredth of the norm of the $\mathbf{C}^H\mathbf{C}$ matrix. This is of course equal to the sum of the squared transfer functions for a single receiver.

It was shown that inverse filtering facilitates multichannel data transmission, which may be used to either increase bandwidth or increase system redundancy. The singular values from a singular value decomposition of the $\mathbf{C}^H\mathbf{C}$ matrix provide information about how well multiple receivers are separated and also provide a quantifiable measure of the difficulty in transmitting to multiple receivers.

The effect that the dynamic range (number of bits) of the SRA has on the sound reproduction was investigated. It was found that a 16-bit system essentially gave a perfect reproduction, with an 8-bit system showing only slight reduction in quality. Surprisingly, a 4-bit IF system provided better data quality compared to either a 16-bit broadside transmission or a 16-bit TR transmission.

Clearly, further work is required to determine how robust the technique is in practice, in particular to changes to the plant matrix with time. This will explored in the forthcoming sea trials planned in late 2001.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support of the Royal Academy of Engineering and DERA. The authors would also like to thank Henrik Schmidt for making OASES freely available and his invaluable help with the package. This work was carried out as part of Technology Group 1 of the MOD Corporate Research Program.

APPENDIX: CALCULATION OF THE OPTIMAL FILTER MATRIX IN THE PRESENCE OF NOISE

For the system shown in Fig. A1 the desired cost function to be minimized is[53]

$$J = E[\mathbf{e}^H\mathbf{e}] + \beta\mathbf{v}^H\mathbf{v}, \tag{A1}$$

where E denotes the expectation operator which is necessary for stochastic signals. The error $\mathbf{e}$ is given by

$$\mathbf{e} = \mathbf{u}e^{j\omega t_0} - \hat{\mathbf{w}}, \tag{A2}$$

where $\mathbf{u}e^{j\omega t_0}$ is the desired signal delayed by some time t_0 to ensure causality is not violated, $\hat{\mathbf{w}} = \mathbf{w} + \mathbf{n}$ is the column vector of reproduced signals contaminated by the background noise $\mathbf{n}$, $\hat{\mathbf{v}} = \mathbf{v} + \mathbf{m}$ is the column vector of source input signals contaminated by the source noise $\mathbf{m}$, and $\hat{\mathbf{C}} = \mathbf{C} + \mathbf{N}$ is the plant matrix estimate contaminated by the plant noise matrix $\mathbf{N}$. Making the appropriate substitutions into Eq. (A1)

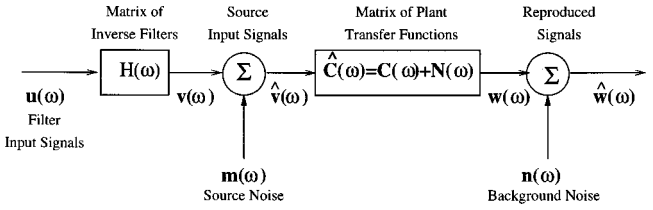


FIG. A1. System diagram with contaminating noise.

$$J = E[[\mathbf{u}e^{j\omega t_0} - \hat{\mathbf{C}}\mathbf{v} - \hat{\mathbf{C}}\mathbf{m} - \mathbf{n}]^H[\mathbf{u}e^{j\omega t_0} - \hat{\mathbf{C}}\mathbf{v} - \hat{\mathbf{C}}\mathbf{m} - \mathbf{n}]] + \beta\mathbf{v}^H\mathbf{v}, \tag{A3}$$

and expanding

$$\begin{aligned} J = E[&|\mathbf{u}e^{j\omega t_0}|^2 - e^{j\omega t_0}\mathbf{u}^H\hat{\mathbf{C}}\mathbf{v} - e^{j\omega t_0}\mathbf{u}^H\hat{\mathbf{C}}\mathbf{m} - e^{j\omega t_0}\mathbf{u}^H\mathbf{n} \\ &- e^{j\omega t_0}\mathbf{v}^H\hat{\mathbf{C}}^H\mathbf{u} + \mathbf{v}^H\hat{\mathbf{C}}^H\hat{\mathbf{C}}\mathbf{v} + \mathbf{v}^H\hat{\mathbf{C}}^H\hat{\mathbf{C}}\mathbf{m} + \mathbf{v}^H\hat{\mathbf{C}}^H\mathbf{n} \\ &- e^{j\omega t_0}\mathbf{m}^H\hat{\mathbf{C}}^H\mathbf{u} + \mathbf{m}^H\hat{\mathbf{C}}^H\hat{\mathbf{C}}\mathbf{v} + \mathbf{m}^H\hat{\mathbf{C}}^H\hat{\mathbf{C}}\mathbf{m} + \mathbf{m}^H\hat{\mathbf{C}}^H\mathbf{n} \\ &- e^{j\omega t_0}\mathbf{n}^H\mathbf{u} + \mathbf{n}^H\hat{\mathbf{C}}\mathbf{v} + \mathbf{n}^H\hat{\mathbf{C}}\mathbf{m} + |\mathbf{n}|^2] + \beta\mathbf{v}^H\mathbf{v}. \end{aligned} \tag{A4}$$

Given that $\mathbf{u}$ and $\mathbf{v}$ are deterministic then these may be taken outside the expectation operator

$$\begin{aligned} J = &|\mathbf{u}e^{j\omega t_0}|^2 - e^{j\omega t_0}\mathbf{u}^H E[\hat{\mathbf{C}}]\mathbf{v} - e^{j\omega t_0}\mathbf{u}^H E[\hat{\mathbf{C}}\mathbf{m}] \\ &- e^{j\omega t_0}\mathbf{u}^H E[\mathbf{n}] - e^{j\omega t_0}\mathbf{v}^H E[\hat{\mathbf{C}}^H]\mathbf{u} + \mathbf{v}^H E[\hat{\mathbf{C}}^H\hat{\mathbf{C}}]\mathbf{v} \\ &+ \mathbf{v}^H E[\hat{\mathbf{C}}^H\hat{\mathbf{C}}\mathbf{m}] + \mathbf{v}^H E[\hat{\mathbf{C}}^H\mathbf{n}] - e^{j\omega t_0}E[\mathbf{m}^H\hat{\mathbf{C}}^H]\mathbf{u} \\ &+ E[\mathbf{m}^H\hat{\mathbf{C}}^H\hat{\mathbf{C}}]\mathbf{v} + E[\mathbf{m}^H\hat{\mathbf{C}}^H\hat{\mathbf{C}}\mathbf{m}] + E[\mathbf{m}^H\hat{\mathbf{C}}^H\mathbf{n}] \\ &- e^{j\omega t_0}E[\mathbf{n}^H]\mathbf{u} + E[\mathbf{n}^H\hat{\mathbf{C}}]\mathbf{v} + E[\mathbf{n}^H\hat{\mathbf{C}}\mathbf{m}] + E[|\mathbf{n}|^2] \\ &+ \beta\mathbf{v}^H\mathbf{v}. \end{aligned} \tag{A5}$$

This can be represented by the quadratic function

$$J = \mathbf{v}^H\mathbf{A}\mathbf{v} + \mathbf{v}^H\mathbf{b} + \mathbf{b}^H\mathbf{v} + c, \tag{A6}$$

which is minimized by[53]

$$\mathbf{v}_{\text{opt}} = -\mathbf{A}^{-1}\mathbf{b} \tag{A7}$$

$$= e^{j\omega t_0}[E[\hat{\mathbf{C}}^H\hat{\mathbf{C}}] + \beta\mathbf{I}]^{-1}E[\hat{\mathbf{C}}^H[\mathbf{u} - \hat{\mathbf{C}}\mathbf{m} - \mathbf{n}]]. \tag{A8}$$

Given that $\mathbf{v} = \mathbf{H}\mathbf{u}$, postmultiplying Eq. (20) with the Hermitian transpose of the filter input vector $\mathbf{u}$ and taking the expected value gives

$$\begin{aligned} \mathbf{H}_{\text{opt}}E[\mathbf{u}\mathbf{u}^H] = e^{j\omega t_0}[E[\hat{\mathbf{C}}^H\hat{\mathbf{C}}] + \beta\mathbf{I}]^{-1}[E[\hat{\mathbf{C}}^H\mathbf{u}\mathbf{u}^H] \\ - E[\hat{\mathbf{C}}^H\hat{\mathbf{C}}\mathbf{m}\mathbf{u}^H] - E[\hat{\mathbf{C}}^H\mathbf{n}\mathbf{u}^H]]. \end{aligned} \tag{A9}$$

Since $\mathbf{u}$ is uncorrelated with the stochastic signals $\mathbf{m}$ and $\mathbf{n}$, then $E[\mathbf{m}\mathbf{u}^H]$ and $E[\mathbf{n}\mathbf{u}^H]$ are null matrices. Also, if the plant noise matrix $\mathbf{N}$ is uncorrelated with the plant matrix $\mathbf{C}$ then the following hold: $E[\mathbf{N}] = 0$, $E[\hat{\mathbf{C}}] = \mathbf{C}$ and $E[\hat{\mathbf{C}}^H\hat{\mathbf{C}}] = \mathbf{C}^H\mathbf{C} + E[\mathbf{N}^H\mathbf{N}]$. Therefore

$$\mathbf{H}_{\text{opt}}\mathbf{S}_{\mathbf{uu}} = e^{j\omega t_0}[E[\hat{\mathbf{C}}^H\hat{\mathbf{C}}] + \beta\mathbf{I}]^{-1}\mathbf{C}^H\mathbf{S}_{\mathbf{uu}}, \tag{A10}$$

where $\mathbf{S}_{\mathbf{uu}} = E[\mathbf{u}\mathbf{u}^H]$ is the autospectral matrix of the filter inputs. Finally, for the above expression to hold for all $\mathbf{S}_{\mathbf{uu}}$,

the estimate of the optimal filter matrix is given by

$$\mathbf{H}_{\text{opt}} = e^{j\omega t}[\mathbf{C}^H\mathbf{C} + E[\mathbf{N}^H\mathbf{N}] + \beta\mathbf{I}]^{-1}\mathbf{C}^H. \quad \text{(A11)}$$

This is the same as Eq. (7) but with the additional term for the plant noise matrix. Therefore, only the uncorrelated noise that is convolved with the input signals affects the estimate of the optimal filter matrix.

[1] M. Stojanovic, J. Catipovic, and J. G. Proakis, ''Adaptive multichannel combining and equalization for under water acoustic communications,'' J. Acoust. Soc. Am. **94**, 1621–1631 (1993).

[2] M. Stojanovic, J. A. Catipovic, and J. G. Proakis, ''Reduced-complexity spatial and temporal processing of underwater acoustic communication signals,'' J. Acoust. Soc. Am. **98**, 961–972 (1995).

[3] M. Stojanovic, J. G. Proakis, and J. A. Catipovic, ''Performance of high-rate adaptive equalization on a shallow water acoustic channel,'' J. Acoust. Soc. Am. **100**, 2213–2219 (1996).

[4] A. Parvulescu and C. S. Clay, ''Reproducibility of signal transmissions in the ocean,'' Radio Electron. Eng. **29**, 223–228 (1965).

[5] D. R. Jackson and D. R. Dowling, ''Phase conjugation in underwater acoustics,'' J. Acoust. Soc. Am. **89**, 171–181 (1991).

[6] D. R. Dowling and D. R. Jackson, ''Narrow-band performance of phase-conjugate arrays in dynamic random media,'' J. Acoust. Soc. Am. **91**, 3257–3277 (1992).

[7] D. R. Dowling, ''Phase-conjugate array focusing in a moving medium,'' J. Acoust. Soc. Am. **94**, 1716–1718 (1993).

[8] D. R. Dowling, ''Acoustic pulse compression using phase-conjugate processing,'' J. Acoust. Soc. Am. **95**, 1450–1458 (1994).

[9] C. Prada, S. Manneville, D. Spoliansky, and M. Fink, ''Decomposition of the time reversal operator. Detection and selective focusing on two scatterers,'' J. Acoust. Soc. Am. **99**, 2067–2076 (1996).

[10] P. Roux, B. Roman, and M. Fink, ''Time-reversal in an ultrasonic waveguide,'' Appl. Phys. Lett. **70**(14), 1811–1813 (1997).

[11] C. Prada and M. Fink, ''Separation of interfering acoustic scattered signals using the invariants of the time-reversal operator: Application to Lamb waves characterization,'' J. Acoust. Soc. Am. **104**, 801–807 (1998).

[12] W. A. Kuperman, W. S. Hodgkiss, H. C. Song, T. Akal, C. Ferla, and D. R. Jackson, ''Phase conjugation in the ocean: Experimental demonstration of an acoustic time-reversal mirror,'' J. Acoust. Soc. Am. **103**, 25–40 (1998).

[13] S. R. Khosla and D. R. Dowling, ''Time-reversing array retrofocusing in simple dynamic underwater environments,'' J. Acoust. Soc. Am. **104**, 3339–3350 (1998).

[14] H. C. Song, W. A. Kuperman, and W. S. Hodgkiss, ''A time-reversal mirror with variable range focusing,'' J. Acoust. Soc. Am. **103**, 3234–3240 (1998).

[15] N. Mordant, C. Prada, and M. Fink, ''Highly resolved detection and selective focusing in a waveguide using the D.O.R.T. method,'' J. Acoust. Soc. Am. **105**, 2634–2642 (1999).

[16] M. Tanter, J.-L. Thomas, and M. Fink, ''Focusing and steering through absorbing and aberrating layers: Application to ultrasonic propagation through the skull,'' J. Acoust. Soc. Am. **103**, 2403–2410 (1998).

[17] M. Tanter, J.-L. Thomas, and M. Fink, ''Time reversal and the inverse filter,'' J. Acoust. Soc. Am. **108**, 223–234 (2000).

[18] J. L. Flanagan and R. C. Lummis, ''Signal processing to reduce multipath distortion in small rooms,'' J. Acoust. Soc. Am. **47**, 1475–1481 (1970).

[19] J. B. Allen, D. A. Berkley, and J. Blauert, ''Multi-microphone signal-processing technique to remove room reverberation from speech signals,'' J. Acoust. Soc. Am. **62**, 912–915 (1977).

[20] S. T. Neely and J. B. Allen, ''Invertibility of a room impulse response,'' J. Acoust. Soc. Am. **66**, 165–169 (1979).

[21] J. Mourjopoulos, ''On the variation and invertibility of room impulse response functions,'' J. Sound Vib. **102**(2), 217–228 (1985).

[22] M. Tohyama, R. H. Lyon, and T. Koike, ''Pulse waveform recovery in a reverberant condition,'' J. Acoust. Soc. Am. **91**(5), 2805–2812 (1992).

[23] O. Kirkeby and P. A. Nelson, ''Reproduction of plane wave sound fields,'' J. Acoust. Soc. Am. **94**, 2992–3000 (1993).

[24] O. Kirkeby and P. A. Nelson, ''Properties of least squares inverse filters used for multi-channel sound reproduction,'' in *Active 95*, pp. 1259–1270 (1995).

[25] O. Kirkeby, P. A. Nelson, H. Hamada, and F. Orduna-Bustamante, ''Fast deconvolution and multi-channel systems using regularization,'' ISVR Technical Report 255, University of Southampton, 1996.

[26] O. Kirkeby, P. A. Nelson, H. Hamada, and F. Orduna-Bustamante, ''Fast deconvolution and multichannel systems using regularization,'' IEEE Trans. Speech Audio Process. **6**(2), 189–194 (1998).

[27] M. Siderius and J.-P. Hermand, ''Yellow Shark Spring 1995: Inversion results from sparse broadband acoustic measurements over a highly range-dependent soft clay layer,'' J. Acoust. Soc. Am. **106**, 637–651 (1999).

[28] H. Schmidt and F. B. Jensen, ''A full wave solution for the propagation in multilayered viscoelastic media with application to Gaussian beam reflection at fluid–solid interfaces,'' J. Acoust. Soc. Am. **77**, 813–825 (1985).

[29] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1994).

[30] C. S. Clay, ''Waveguides, arrays, and filters,'' Geophysics **31**, 501–506 (1966).

[31] C. S. Clay, ''Use of arrays for acoustic transmission in a noisy ocean,'' Rev. Geophys. **4**, 475–507 (1966).

[32] C. S. Clay, ''Optimum time domain transmission and source location in a waveguide,'' J. Acoust. Soc. Am. **81**, 660–664 (1987).

[33] I. Tolstoy and C. S. Clay, *Ocean Acoustics—Theory and Experiment in Underwater Sound*, 2nd ed. (American Institute of Physics, New York, 1987).

[34] C. Dorme and M. Fink, ''Focusing in transmit–receive mode through inhomogeneous media: The time reversal matched filter approach,'' J. Acoust. Soc. Am. **98**, 1155–1162 (1995).

[35] A. Parvulescu, ''Matched signal (M.E.S.S.) processing by the ocean,'' J. Acoust. Soc. Am. **98**, 943–960 (1995).

[36] P. A. Nelson, F. Orduna-Bustamante, and H. Hamada, ''Inverse filter design and equalization zones in multichannel sound reproduction,'' IEEE Trans. Speech Audio Process. **3**(3), 185–192 (1995).

[37] P. A. Nelson, O. Kirkeby, T. Takeuchi, and H. Hamada, ''Sound fields for the production of virtual acoustic images,'' J. Sound Vib. **204(2)**, 386–396 (1997).

[38] O. Kirkeby, P. Rubak, P. A. Nelson, and A. Farina, ''Digital cross-talk cancellation networks,'' J. Audio Eng. Soc. (submitted).

[39] O. Kirkeby, P. Rubak, P. A. Nelson, and A. Farina, ''Design of cross-talk cancellation networks by using fast deconvolution,'' in AES Convention, Munich, April 1999.

[40] P. A. Nelson, ''A tutorial introduction to the design of multichannel inverse filters for audio signal processing,'' (Lecture notes presented to the Benelux Section of the IEEE Signal Processing Society, Eindhoven), December 1997.

[41] Y. Kahana, P. A. Nelson, and S. Yoon, ''Experiments on the synthesis of virtual acoustic sources in automotive interiors,'' in AES 16th International Conference on Spatial Sound Reproduction, 1999.

[42] H. Irisawa, S. Shimada, and H. Hokari, ''Study of a fast method to calculate inverse filters,'' J. Audio Eng. Soc. **46**(7/8), 611–620 (1998).

[43] A. Neumaier, ''Solving ill-conditioned and singular linear systems: A tutorial on regularization,'' SIAM Rev. **40**, 636–666 (1998).

[44] H. Schmidt, OASES *Version 2.1: User Guide and Reference Manual*, Department of Ocean Engineering, MIT, 2.1 edition, 1997.

[45] H. Schmidt, SAFARI: Seismo-acoustic fast field algorithm for range independent environments. User's guide. Technical Report SR 113, SACLANT, ASW Research Centre, La Spezia, Italy, 1987.

[46] S. H. Yoon and P. A. Nelson, ''Estimation of acoustic source strength by inverse methods: Part II: Methods for choosing regularization parameters. ISVR Technical Report 279, University of Southampton, October 1998.

[47] M. Miyoshi and Y. Kaneda, ''Inverse filtering of room acoustics,'' IEEE Trans. Acoust., Speech, Signal Process. **36**(2), 145–152 (1988).

[48] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing—Principles, Algorithms and Applications*, 3rd ed. (Prentice-Hall International, 1996).

[49] D. Cassereau and M. Fink, ''Focusing with plane time-reversal mirrors: An efficient alternative to closed cavities,'' J. Acoust. Soc. Am. **94**, 2373–2386 (1993).

[50] H. C. Song, W. A. Kuperman, W. S. Hodgkiss, T. Akal, and C. Ferla, ''Iterative time reversal in the ocean,'' J. Acoust. Soc. Am. **105**, 3176–3184 (1999).

[51] C. Draeger and M. Fink, ''One-channel time-reversal in chaotic cavities: Theoretical limits,'' J. Acoust. Soc. Am. **105**, 611–617 (1999).

[52] A. Derode, A. Tourin, and M. Fink, ''Ultrasonic pulse compression with one-bit time reversal through multiple scattering,'' J. Appl. Phys. **85**(9), 6343–6352 (1999).

[53] P. A. Nelson, ''Some inverse problems in acoustics,'' in Proceedings of the 6th International Congress on Sound and Vibration, July 1999.

Guided waves energy velocity in absorbing and non-absorbing plates

A. Bernard[a] and M. J. S. Lowe
Department of Mechanical Engineering, Imperial College of Science, Technology and Medicine, Exhibition Road, London SW7 2BX, United Kingdom

M. Deschamps
Laboratoire de Mécanique Physique, Université Bordeaux 1, UMR 5469 C.N.R.S, Cours de la Libération, 33405 Talence Cedex, France

(Received 1 March 2000; revised 13 March 2001; accepted 4 April 2001)

This paper presents a study of the velocity of the propagation of energy in guided waves in plates. The motivation of the work comes from the practical observation that the conventional approach to predicting the velocities of pulses or wave packets, that is, the simple group velocity calculation, breaks down when the guided waves are attenuative. The conventional approach is therefore not valid for guided waves in absorbing materials or for leaky waves. The paper presents a theoretical derivation of an expression to predict the energy velocity of guided waves in an isotropic plate, based on the integration of the Poynting energy vectors. When applied to modes with no attenuation, it is shown analytically from this expression that the energy velocity is always identical to the group velocity. On the other hand, when applied to attenuative modes, numerical integration of the expression to yield the true energy velocity shows that this can differ quite significantly from the group velocity. Experimental validation of the expression is achieved by measuring the velocity of wave packets in an absorbing plate, under such conditions when the energy velocity differs substantially from the group velocity. Excellent agreement is found between the predictions and the measurements. The paper also shows the Poynting vectors in the various model studies, and some interesting phenomena relating to their directions. [DOI: 10.1121/1.1375845]

PACS numbers: 43.35.Cg, 43.35.Pt [SGK]

I. INTRODUCTION

A fundamental feature of waves is their capability of carrying energy and information over long distances. A particular interest of the authors is the use of guided waves for the non-destructive testing (NDT) of structures and the characterization of materials. The guided waves in plate, bar or pipe structures are described by the dispersion curves which show how the wave numbers and velocities of the waves vary with the frequency. The velocities which are of interest are the phase velocity, the group velocity and the energy velocity. The phase velocity is the velocity at which the wave fronts or crests travel, the group velocity is the derivative of the frequency wave-number dispersion relation, $Vg = \partial\omega/\partial k_x$ (where ω is the angular frequency, and k_x is the wave number in the direction of propagation), and the energy velocity is the velocity at which the wave carries its potential and kinetic energy along the structure. Long range testing usually makes use of finite tone bursts or wave packets and optimally exploits waves at frequencies where there is little dispersion, thus the different frequency components within the wave packet propagate at the same velocity and so the wave packet retains its shape as it travels. Naturally we would expect the energy to be transported at the speed of travel of the wave packet, and typically in practice it is true to take this to be equal to the group velocity. This is in fact a most useful property to consider in the context of long range propagation because in such work the focus tends to be on the behavior of the wave packet rather than the phase information within it, and the wave packet velocity is very readily measured simply by recording the arrival times of the packet.

It is interesting in such applications that the energy appears to propagate at a speed given so simply by the group velocity derivative. However such a relationship does not always hold, a clear example being attenuating waves. If an attenuating wave is described, as is conventional, by a complex wave number, then the group velocity calculation yields non-physical solutions such as infinite velocities at some locations on the dispersion curves. It was this anomaly which motivated the work presented in this paper to perform a strict derivation for the energy velocity in attenuating waves and then compare the results with the simple group velocity expression.

A great deal of work has already been published concerning group velocity and energy velocity. In 1951, Broer[1] considered when and why the rate of energy propagation of waves, in a one-dimensional conservative system without dissipation, equals the group velocity. Using the method of stationary phase he demonstrated that everywhere in the wave system the energy is propagated with the group velocity corresponding to the local wave number. Biot,[2] in 1957,

[a]Permanent address: Laboratoire de Mécanique Physique, Université Bordeaux I, UMR 5469 C.N.R.S, Cours de la Libération, 33405 Talence Cedex, France.

showed that there is a rigorous identity between the group velocity and the velocity of energy transport in non-homogeneous media with or without anomalous dispersion. After a survey of the theory of group velocity for one-dimensional and three-dimensional, isotropic and anisotropic, homogeneous and inhomogeneous, conservative and dissipative, linear and non-linear systems exhibiting wave propagation under free and forced motion condition from Lighthill,[3] Hayes[4] showed in 1977 in a simpler way that the energy flux velocity vector for a single infinite train of elliptically polarized harmonic small amplitude plane waves propagating in a homogeneous conservative, dispersive system is equal to the group velocity. Then it was in 1979 that Hayes and Musgrave[5] explained that the equivalence between group velocity and energy velocity does not hold true in general for inhomogeneous waves. This was confirmed for example by Poirée[6] in 1984, who demonstrated that the energy velocity of the spatially dispersive (that is, the dispersion is caused by geometric effects) plane evanescent wave solution of the linear acoustic equation in a perfect fluid is equal to the phase velocity instead of the group velocity. It was also shown by Borejko[7] in 1987 and by Mainardy[8] who postulated that when the energy is not conserved, the identification of energy velocity with the kinetic concept of group velocity is not valid. Finally, Deschamps, Poirée, and Poncelet[9] showed the energy velocity of complex harmonic plane waves, characterized by a complex wave vector and a complex frequency, may be interpreted as the phase velocity in the direction of the real part of the slowness bivector.

This paper examines the calculation of the energy velocity in guided waves by integrating the Poynting vectors through the depth of the plate and over a temporal cycle, as set out for example by Auld.[10] Results using this approach, which is general in its applicability to guided waves, are then obtained for three specific guided wave cases and compared to the group velocity solutions. First the Lamb waves in a non-absorbing isotropic plate in vacuum are studied in order to demonstrate the agreement, in this lossless case, of the energy velocity with the group velocity. Then a non-absorbing isotropic plate immersed in water is studied, revealing the departure of the energy velocity solutions from the group velocities. Finally an absorbing plate in vacuum is studied, again showing separate energy velocity and group velocity solutions. Experimental results for one of the modes in the absorbing plate are also shown, confirming that a wave packet does indeed travel at the predicted energy velocity.

II. WAVE FIELDS IN A PLATE

We start by presenting the expressions for the fields of guided waves in a plate; these will be used later for the energy derivations. Let us consider an infinite plane parallel absorbing plate of thickness $2h$, density ρ, longitudinal velocity $^*C_L$, and shear velocity $^*C_T$ immersed in a fluid. The simpler examples of an absorbing (or not) plate in vacuum (or not) will easily be deduced from this general case.

The attenuation in the material (Kelvin–Voigt spring-dashpot linear model) is assumed to be linearly proportional to the frequency; this is a reasonably representative model for most homogeneous structural materials. The complex velocities are defined by

$$^*C_{L,T} = c_{L,T}\left(1 - i\frac{\alpha_{L,T}}{2\pi}\right)^{-1}, \tag{1}$$

where $c_{L,T}$ represent the longitudinal and shear bulk velocities of the material and $\alpha_{L,T}$ the longitudinal and shear attenuation in the material, expressed in Nepers per wavelength. The symbol * on the top left-hand side of the different following quantities indicates that they are complex.

The general symmetric displacements inside the plate are given by[11]

$$\begin{aligned} ^*\mu_x &= ^*f(y)\exp[i(\omega t - ^*k_x x)], \\ ^*u_y &= ^*g(y)\exp[i(\omega t - ^*k_x x)], \end{aligned} \tag{2}$$

where ω is the circular frequency (real and positive because only harmonic waves are considered in this paper), $^*k_x = k_x' + ik_x''$ is the complex wave number, x is the coordinate along the plate, and y is the coordinate normal to the plate and

$$\begin{aligned} ^*f(y) &= -i\,^*k_x A_2 \cos(^*py) + ^*qB_1 \cos(^*qy), \\ ^*g(y) &= -^*pA_2 \sin(^*py) + i\,^*k_x B_1 \sin(^*qy), \end{aligned} \tag{3}$$

in which A_2 stands for the amplitude of the two longitudinal partial waves, B_1 defines the amplitude of the two shear partial waves, and

$$^*p = \sqrt{\frac{\omega^2}{^*C_L^2} - ^*k_x^2}, \tag{4}$$

$$^*q = \sqrt{\frac{\omega^2}{^*C_T^2} - ^*k_x^2}. \tag{5}$$

The well-known dispersion equation of the symmetric Lamb wave function is, for example, given by[12]

$$\begin{aligned} Cs(\omega, ^*k_x) = (\omega^2 - 2\,^*C_T^{2\,*}k_x^2)^2 \cot(^*ph) \\ + 4\,^*C_T^{4\,*}k_x^{2\,*}p\,^*q \cot(^*qh) + i\,^*\tau = 0 \end{aligned} \tag{6}$$

with

$$^*\tau = \frac{\rho_F^*p}{p_S m}, \tag{7}$$

where ρ_S stands for the density of the solid, ρ_F for the density of the fluid, $^*m = \sqrt{(\omega^2/c_F^2) - ^*k_x^2}$ and c_F is the longitudinal bulk velocity in the fluid.

The dispersion equation of the anti-symmetric Lamb wave function is obtained by changing the cotangents into tangents and by taking $^*\tau = -^*\tau$.

Let us now consider the real part of the displacements in the plate at a given y_0 position given by Eq. (2), denoted by u_x and u_y. They can be expressed as follows:

$$\begin{aligned} u_x &= f' \cos(\varphi) - f'' \sin(\varphi), \\ u_y &= g' \cos(\varphi) - g'' \sin(\varphi), \end{aligned} \tag{8}$$

where f' and g' stand for the real part of $^*f(y_0)$ and $^*g(y_0)$, respectively, f'' and g'' for their imaginary parts and

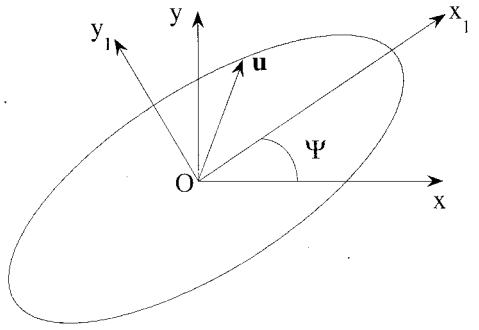


FIG. 1. Lamb wave displacements over a time period where O describes the position over the thickness of the plate.

φ for the common phase term $(\omega t - k_x' x)$. The amplitude decay term $\exp(-k_x'' x)$, being common to u_x and u_y, has voluntarily been omitted. After several manipulations of Eq. (8) it can be shown easily that any Lamb wave displacements describe an ellipse whose large axis is rotated by an angle ψ from the reference coordinate system. This is illustrated in Fig. 1, in which O is the coordinate origin, x is the direction along the plate and y is the direction normal to the plate. The locus of the displacement over one cycle of a wave describes an ellipse such as the one shown in the figure. The axes of the ellipse, $x_1 y_1$, are rotated by the angle ψ with respect to the reference system xy. In fact, the displacements of the four partial waves which constitute a Lamb wave, and their superposition to form the Lamb wave, are all elliptical.

The angle of the axis of the ellipse is such that

$$\tan(\Psi) = \frac{g''}{f''} = -\frac{f'}{g'}. \tag{9}$$

It is clear from this equation that for the general case when $^*k_x$ is complex, i.e., f', f'', g', g'' are not null, $\tan(\psi)$ is neither zero nor infinite. Moreover, the simple case of a non-absorbing plate in vacuum ($\alpha_{L,T}=0$, $f''=0$, and $g'=0$) is the only example for which ψ is always $\pi/2$. In that case the axis of the ellipse, for each different y position inside the plate, are collinear to the plate axis.

III. NON-ABSORBING PLATE IN VACUUM

Here we consider the particular case of Lamb waves when there is no attenuation. This is included here to confirm the agreement of the group velocity and energy velocity in such cases.

Now, since $\alpha_{L,T}=0$ and $\tau=0$, the non-standing but propagating solutions of the symmetric and antisymmetric Lamb wave dispersion equations given by Eq. (6) are real ($^*k_x = k_x$).

A. Group velocity

This is the velocity of a modulated wave which is constructed by taking two waves with slightly different values of ω and k_x.[10]

The propagation velocity of the carrier is the phase velocity defined by

$$V_\phi = \frac{\omega}{k_x} \tag{10}$$

and the propagation velocity of the envelope is the group velocity:

$$V_g = \frac{\partial \omega}{\partial k_x}. \tag{11}$$

With this definition, any dispersion relation linking the frequency ω and the wave number k_x is suitable to calculate the group velocity. Thus the symmetric Lamb function given in Eq. (6) is considered and as $dCs(\omega,k_x)=0$ therefore the group velocity can be expressed as

$$V_g = -\frac{\partial Cs(\omega,k_x)/\partial k_x}{\partial Cs(\omega,k_x)/\partial \omega}. \tag{12}$$

Then, expanding this gives

$$V_g = \frac{k_x c_L^2}{\omega}\left(\frac{A}{B}\right) \tag{13}$$

with

$$A = -8pqc_T^2(\omega^2 - 2c_T^2 k_x^2)\cot(ph) - 4c_T^4(k_x^2(p^2+q^2) - 2p^2q^2)\cot(qh) + hq\left(\frac{(\omega^2 - 2c_T^2k_x^2)^2}{\sin^2(ph)} + \frac{4c_T^4 k_x^2 p^2}{\sin^2(qh)}\right) \tag{14}$$

and

$$B = -4pqc_L^2(\omega^2 - 2c_T^2k_x^2)\cot(ph) - 4c_T^2k_x^2(c_L^2p^2 + c_T^2q^2)\cot(qh) + hq\left(\frac{(\omega^2 - 2c_T^2k_x^2)^2}{\sin^2(ph)} + \frac{4c_T^2c_L^2k_x^2p^2}{\sin^2(qh)}\right). \tag{15}$$

B. Energy velocity

The energy velocity vector is defined at a given y position in the plate by[10]

$$\mathbf{V}_e(y) = \frac{\langle \mathbf{P} \rangle}{\langle E \rangle}. \tag{16}$$

where $\langle(.)\rangle$ denotes the time average over one period:

$$\langle(.)\rangle = \frac{1}{T}\int_0^T (.)\,dt \tag{17}$$

and where $\mathbf{P}$ stands for the Poynting vector and E for the total energy in the system (potential and kinetic). Those quantities are real and defined as follows:

$$\mathbf{P} = -\bar{\bar{\sigma}}\,\dot{\mathbf{u}}. \tag{18}$$

$$E = \tfrac{1}{2}\rho \dot{\mathbf{u}}\cdot\dot{\mathbf{u}} + \tfrac{1}{2}\lambda\theta^2 + \mu\bar{\bar{\varepsilon}}:\bar{\bar{\varepsilon}}, \tag{19}$$

where $\dot{\mathbf{u}}$ is the particle velocity vector, λ and μ are the Lamé coefficients, $\bar{\bar{\sigma}}$ and $\bar{\bar{\varepsilon}}$ are the stress and the strain tensor, respectively, and $\theta = \varepsilon_{11}+\varepsilon_{22}+\varepsilon_{33}$ is the dilatation. In the calculation of $\mathbf{P}$ and E all the different quantities have been calculated using the real part of the displacements given in Eq. (2).

However, as we are dealing with Lamb waves, that is waves confined within a plate, our interest is in the energy velocity flux in the whole plate. This requires additionally an integral through the thickness of the plate and is defined by[11]

$$\overline{\mathbf{V}}_{\mathbf{e}}=\frac{\langle\langle \mathbf{P}\rangle\rangle}{\langle\langle E\rangle\rangle}, \tag{20}$$

where $\langle\langle(.)\rangle\rangle$ denotes an average over a time period as well as over the thickness of the plate:

$$\langle\langle(.)\rangle\rangle=\frac{1}{2h}\int_{-h}^{h}\langle(.)\rangle dy. \tag{21}$$

After long and non-trivial calculations, it is found that, when $V_\phi > c_L$ (that is p and q are pure real), the component of the Poynting vector at any location following the y direction is always null and the component of the flux in the whole plate following the **x** direction is

$$\begin{aligned}\langle\langle P_x\rangle\rangle=&\frac{\rho\omega^3 k_x \sin^2(qh)\sin^2(ph)}{q}\\ &\times\left\{-4c_T^2p^2(\omega^2-3c_T^2q^2)\cot(qh)\right.\\ &+\frac{q}{p}(\omega^2-2c_T^2q^2)(\omega^2-2c_T^2q^2+8c_T^2p^2)\cot(ph)\\ &\left.+hq\left(\frac{(\omega^2-2c_T^2k_x^2)^2}{\sin^2(ph)}+\frac{4c_T^4k_x^2p^2}{\sin^2(qh)}\right)\right\}.\end{aligned} \tag{22}$$

Also

$$\begin{aligned}\langle\langle E\rangle\rangle=&\frac{\rho\omega^4\sin^2(ph)\sin^2(qh)}{qc_L^2}\\ &\times\left\{-\frac{4c_T^2p^2(\omega^2-c_T^2q^2)(\omega^2-3c_T^2q^2)}{(\omega^2-c_T^2q^2+c_T^2p^2)}\cot(qh)\right.\\ &+\frac{q(\omega^2-c_T^2q^2)(\omega^2-2c_T^2q^2)(\omega^2-c_T^2q^2+8c_T^2p^2)}{p(\omega^2-c_T^2q^2+c_T^2p^2)}\\ &\left.\times\cot(ph)+hq\left(\frac{(\omega^2-2c_T^2k_x^2)^2}{\sin^2(ph)}+\frac{4c_T^2c_L^2k_x^2p^2}{\sin^2(qh)}\right)\right\}.\end{aligned} \tag{23}$$

The energy velocity vector, expanded to its direction components, is thus

$$\overline{\mathbf{V}}_{\mathbf{e}}=\begin{bmatrix}V_e\\0\end{bmatrix}. \tag{24}$$

Making use again of the abbreviations A and B of Eqs. (14) and (15), it is found that

$$V_e=\frac{k_xc_L^2}{\omega B}\left(A+\frac{q}{p}Cs\right). \tag{25}$$

But by definition [Eq. (6)], $Cs=0$, so by comparing Eq. (25) with Eq. (13), it follows that $V_e=V_g$.

The situations when $c_T<V_\phi<c_L$ and $V_\phi<c_T$ have of course been explored too. They also lead by the same process of analysis to the equality between energy and group velocities but for brevity they are not presented. The only alteration in the algebra is that hyperbolic functions appear in place of the trigonometric functions because p and/or q become purely imaginary. The same demonstration can be made for the antisymmetric Lamb modes.

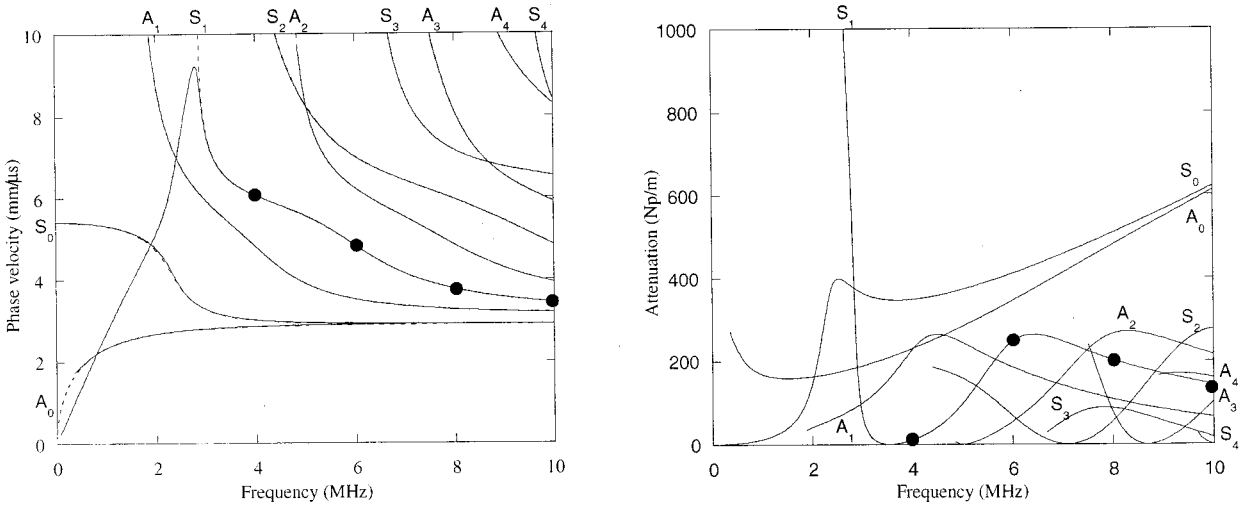


FIG. 2. Dispersion curves comparison between a 1 mm thick aluminum plate in vacuum (dotted line) and in water (solid line). Part (a) phase velocity (ω/k_x') and part (b) attenuation (k_x'' for the leaky case) both versus the frequency. The marks ● indicate the frequencies for the energy plots in Fig. 3.

IV. NON-ABSORBING PLATE IN A FLUID

Let us now study the non-absorbing plate immersed in a fluid for which the equality between the group velocity and the energy velocity is known not always to hold true. In this configuration, solutions of symmetric and antisymmetric Lamb functions are now complex for most locations on the dispersion curves, which means that the *x* wave number becomes a complex quantity. The imaginary part of the wave number describes the attenuation of the guided wave due to leakage of energy by radiation into the fluid.

If a calculation of the group velocity is now attempted according to Eqs. (11) and (12), it becomes necessary to consider the complex wave number, so *Cs* must be differentiated with respect to $*k_x$. This would be possible mathematically, but it no longer makes physical sense.

Consideration of the energy velocity calculation also leads to difficulty, in this case not by a physical constraint but by algebraic complication. The relations given by Eqs. (22) and (23) would now have to contain complex wave numbers (neither purely real nor purely imaginary), and so would be a linear combination of trigonometric and hyperbolic functions. It was decided realistically therefore that an analytical result for the energy velocity vectors was unworkable, and so numerical computations were undertaken instead.

Figure 2 shows some computed dispersion curves for a 1 mm thick aluminum plate, in vacuum (dotted lines) or immersed in water (solid lines). Part (a) shows the phase velocity, indicating very little difference between the curves for the vacuum case and the immersed case. Part (b) shows the attenuation of the immersed case, expressed in Nepers/m; of course the attenuation of the vacuum case is zero. The properties of the materials are $c_L = 6.37\,\text{mm}/\mu\text{s}$, $c_T = 3.1\,\text{mm}/\mu\text{s}$, $\rho = 2800\,\text{kg/m}^3$ for the aluminum and $c_L = 1.5\,\text{mm}/\mu\text{s}$, $\rho = 1000\,\text{kg/m}^3$ for the water.

An obvious difference between the vacuum and immersion cases is with the mode S_1. In the immersion case this mode exhibits a cutoff in phase velocity which has already

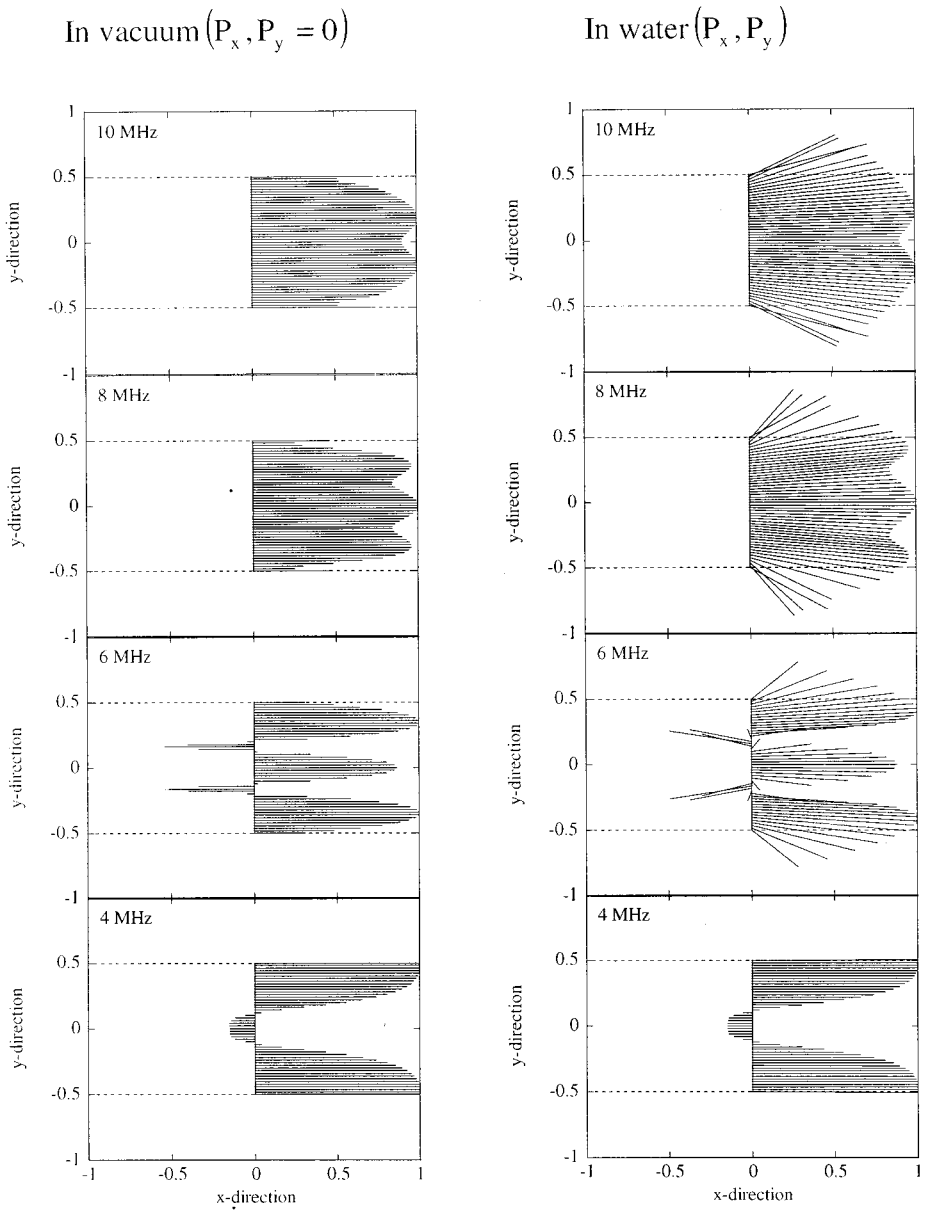


FIG. 3. Local energy velocity vectors comparison between the aluminum plate in vacuum (left) and in water (right) calculated at different *y* positions for different frequencies labeled ● in Fig. 2.

been examined by Lenoir *et al.*,[13] for example. This behavior is accompanied by an increase of attenuation up to 4000 Np/m as the frequency tends to zero [for clarity, not shown in part (b) of Fig. 2]. This very high attenuation reduces the phase velocity of S_1 instead of its rise to infinity as for a plate in vacuum. Such cutoffs are also observed if the problem is solved not in real frequency and complex slowness but in complex frequency and real slowness.[14–16] In that case, again, the imaginary part of the frequency tends to infinity and cutoffs in frequency instead of phase velocity appear for the modes commonly labeled A_0 and S_0, confirming the idea that the high attenuation leads to cutoff either in frequency or in phase velocity. We will consider now some more detailed results for this mode, with the aim of studying the influence of the fluid loading.

A. Energy velocity vectors comparison

Since we know that the imaginary part of the wave number indicates the leakage of energy into the fluid, we should expect the energy velocity vectors to show some components in the direction normal to the plate. This is of course in contrast with the plate in vacuum for which it was shown earlier that the component in the normal direction is zero. We have calculated the energy velocity vectors at four different frequencies (10, 8, 6, 4 MHz), marked by filled circles in Fig. 2. The vectors, defined by Eq. (16), are shown over a range of different y positions in Fig. 3. The plots on the left-hand side of the figure are for the aluminum plate in vacuum and on the right-hand side for the aluminum plate immersed in water. In the vertical direction each plot represents one of the four chosen locations on the dispersion curve. In each plot the top and bottom of the plate are identified by dotted lines and the energy velocity vectors are shown as solid lines. The vectors have directions and lengths according to their actual directions and amplitudes. The amplitude scale is arbitrarily normalized to the maximum in each plot.

Let us first consider the left column of Fig. 3. As explained in Sec. III, the solutions of Lamb wave functions being real for the case in vacuum, the energy velocity vectors only have a component along the **x** direction: $\langle\langle P_y \rangle\rangle$ is always null. However, regardless of the direction of the global propagation of energy along the plate, it is permissible according to the derivation that locally an individual vector may lie in either the positive or the negative **x** direction. It is interesting to see that in fact vectors with negative **x** directions are actually predicted at the 4 MHz and 6 MHz locations when clearly the global energy direction is positive. There are also circumstances when the global energy propagation, that is the sum of these vectors, is in the negative direction. For example, this occurs for the S_1 mode when the phase velocity is greater than 11.44 mm/μs in the so-called backward propagating or negative group velocity region. The negative group velocity phenomenon has been examined in detail by Meitzler,[17] demonstrated experimentally by Wolf *et al.*,[18] and used to explain observed phenomena of backscattering from curved shells by Kaduchak *et al.*[19,20] The energy velocity vectors for one position on the curve in this case are presented in Fig. 4, showing a profile similar to that plotted by Meitzler.

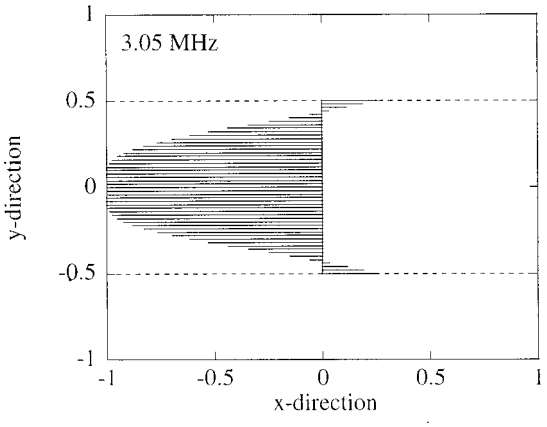


FIG. 4. Local energy velocity vectors for the mode S_1 corresponding to a negative group velocity and to a negative global energy velocity.

This raises the interesting observation that energy can travel in the negative direction even though the mode is modelled by the superposition of four partial plane waves (two longitudinal and two shear) whose wave number projections on the **x** direction are positive. Clearly it is necessary to be very careful when dealing with the energy in waves, and issues of the superposition of waves. Lamb waves exist because of constructive or destructive interferences between those partial waves. Therefore, what is seen is a consequence of those interferences but does not express the exact contribution of each phenomenon. In fact, we acknowledge that this issue is not really understood yet.

In contrast, the leaky case generally exhibits energy velocity components on the **y** direction. It is important to note that, in this second example, the orientations of the energy velocity vectors are a consequence of the nature of k_x which has become a complex quantity. As for the elliptical displacements presented in Sec. II, an angle of rotation function of the y position in the plate is observable, and the energy seems to move towards the fluid. However, there is no evidence that these angles correspond to the ellipse angles. The global shape of the energy velocity vectors has been conserved and as in vacuum some components are oriented toward the negative **x** direction. We also see that the extent to which the vectors point in the **y** direction corresponds broadly to the strength of the leakage. For example, at 4 MHz, the attenuation (the imaginary part of $*k_x$), is almost null, and correspondingly the components of the energy velocity vectors along the **y** axis are rather insignificant compared to the components along the **x** axis, and the vectors in general closely resemble those for the plate in vacuum.

B. Energy velocity and group velocity comparison

The motivations which drove us to study the energy velocity concept instead of the group velocity arose because of the discontinuities and the unacceptable values (greater than the longitudinal bulk velocity of the material) obtained when calculating the group velocity of attenuative guided waves. One simple, but obviously approximate, approach in such cases is to use the definition: $V_g = [\partial\omega/\partial\,\mathrm{Re}(*k_x)]$. This ex-

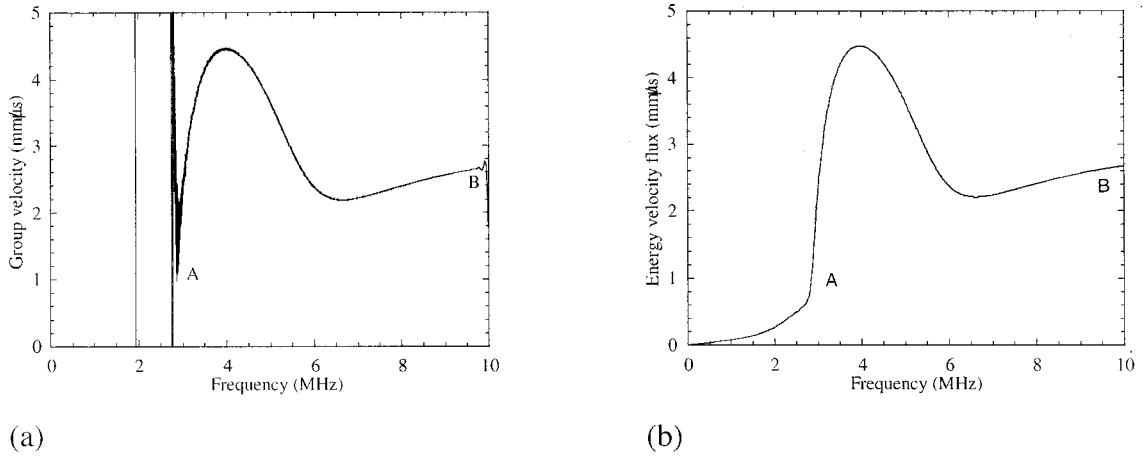


FIG. 5. (a) Group velocity for S_1, (b) modulus of the energy velocity flux $\tilde{\mathbf{V}}_\mathbf{e}$ for S_1. The curves in parts (a) and (b) of the figure are identical between the locations labeled A and B.

pression is of course correct in the limit of reducing attenuation. Figure 5(a) shows the predicted group velocity for the S_1 mode for the aluminum plate immersed in water, made using this expression. The discontinuities which arise at around 3 MHz down to 0 MHz correspond to the location on Fig. 2 where the phase velocity is decreasing back to zero, the region of the curve where the attenuation is extremely large. They are in fact due to infinite slopes when representing the solutions on a graph of frequency versus real wave number, these slopes being a genuine feature of this curve. Indeed the group velocity is negative at all frequencies between the two vertical lines on the graph, that is between frequencies of about 1.95 MHz and 2.75 MHz.

Figure 5(b) shows the energy velocity prediction calculated numerically using Eq. (20). It shows that the energy velocity and the group velocity are in good agreement in the region where the attenuation is low, that is in the region from the label A to the label B in the figure. In fact if the two graphs were plotted on the same figure then the lines would overlay in this region. However the agreement is not good where the attenuation is high, that is at frequencies below location A. Here the energy velocity curve is smooth and does not present any unacceptable values such as excursions above the longitudinal bulk velocity of the material. We can also observe that it converges towards zero as the frequency decreases and the attenuation increases.

V. ABSORBING PLATE IN VACUUM

We now consider another attenuative system, but in this case the attenuation is due to material damping rather than leakage. The system is a 12.7 mm thick plastic plate in vacuum. The plastic is a High Performance PolyEthylene (HPPE). This particular material was chosen because it has been fully studied by Chan[21,22] who measured the following properties: $c_L = 2.344\,\text{mm}/\mu\text{s}$, $\alpha_L = 0.055\,\text{Np/wavelength}$, $c_T = 0.953\,\text{mm}/\mu\text{s}$, $\alpha_T = 0.286\,\text{Np/wavelength}$, $\rho = 953\ \text{kg/m}^3$. His dispersion curves, some of which are reproduced here in Fig. 6. were calculated using a general purpose model developed by Lowe[23] and Pavlakovic, Lowe *et al.*[24]

A. Theoretical curves

The modes, shown in Fig. 6, are divided into two categories. The first, called shear modes, have phase velocities

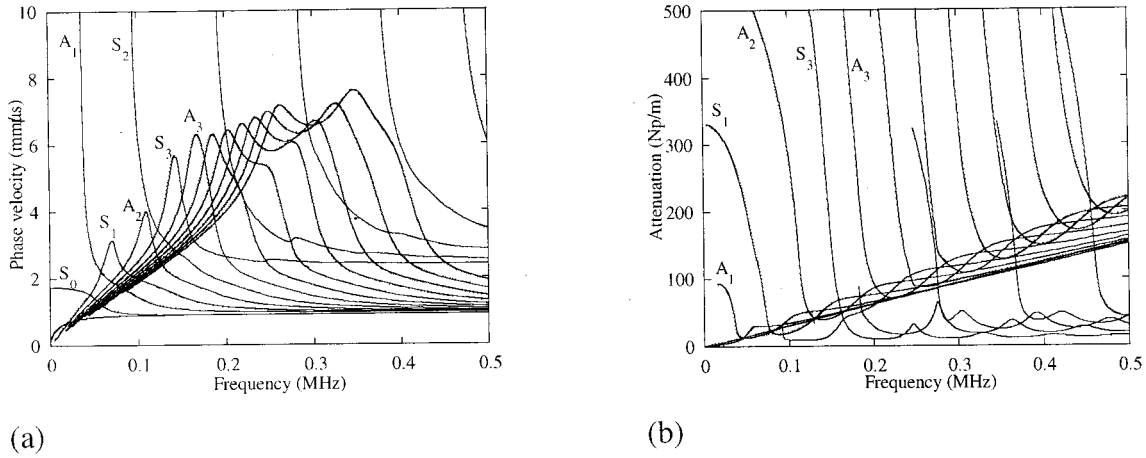


FIG. 6. Dispersion curves for a 12.7 mm HPPE plate in vacuum (Refs. 21 and 22). Part (a) shows the phase velocity and part (b) shows the attenuation, both versus the frequency.

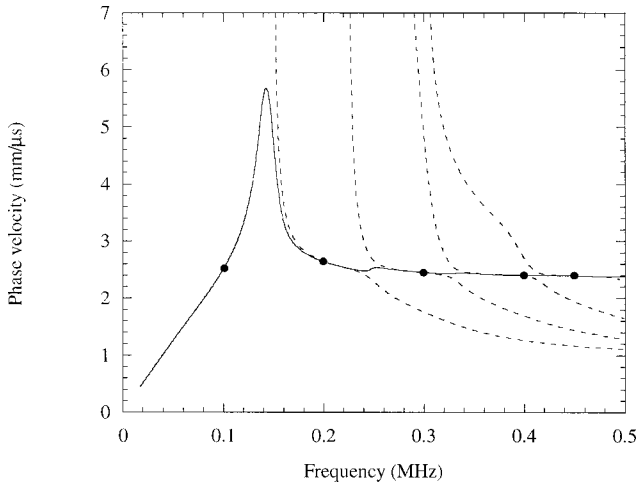


FIG. 7. Comparison between calculated S_3 dispersion curve (solid line) in HPPE and dispersion curves in HPPE without material damping (dotted lines).

which tend towards the shear bulk velocity of HPPE at high frequency and the second, called longitudinal modes, have phase velocities which tend towards the longitudinal bulk velocity of HPPE. It is interesting to see in this example too that cut-offs in phase velocity occur, in this case for nearly all of the modes. Again these correspond to locations where the attenuation is very large [for clarity, not shown in Fig. 6(b)]. The modes which have a lower attenuation correspond in fact to the longitudinal modes, all of the shear modes having an attenuation higher than that of S_0 and A_0 [oblique lower line in Fig. 6(b)].

B. Energy velocity vectors comparison

In this section we calculate the local energy velocity vectors $\mathbf{V_e(y)}$ at different depths through the thickness of the HPPE and compare them with the hypothetical vectors for the same plate if the damping was not present. This should be a fundamentally different study than was performed earlier for the immersed plate because in the present case the energy is lost within the plate and there can be no radiation from the surfaces of the plate.

The mode chosen to do the comparison is S_3. This mode, which has been studied in detail by Chan,[21] has relatively low attenuation and so was identified as being the most amenable for practical work. We will present some experimental results using it in Sec. VI. Figure 7 shows a comparison between the dispersion curve of S_3 in HPPE (solid line) and some of the modes for the same plate but without any material damping (dotted lines). As explained by Chan, when material damping is absent plateauing regions exist. By joining these regions together, it is possible to visualize a new set of dispersion curves which tend towards the longitudinal bulk velocity of HPPE and these represent the curves when the damping is included. In fact, S_3, like all the other longitudinal modes, can be understood to be a result of these regroupings between several different modes, in this case the four dotted curves in the figure.

Energy vector plots for five locations on the dispersion curves are shown in Fig. 8. The locations are identified by filled circles in Fig. 7. The arrangement for the plot follows that which was explained in the earlier comparison in Sec. IV A. The left-hand side shows the vector plots for the HPPE plate without material damping and the right-hand side shows them for the HPPE plate with material damping. At 0.1 MHz there is no equivalent propagating mode in the undamped case and so no plot can be shown on the left-hand side for this frequency. Although, S_3 results from an association of four different modes, the comparison between the energy velocity vectors for the two configurations looks close, with the exception perhaps at 0.4 MHz where some of the energy has negative components for the HPPE without material damping but not for the HPPE with material damping. However, the most interesting feature of this study is evident at 0.1 MHz. Here, where the attenuation is relatively large, it can be seen that some of the vectors have components in the **y** direction, despite the fact that the plate is in vacuum and there is no leakage. But the vectors at the surfaces of the plate are parallel to the **x** direction and so there is only internal transfer of energy. Closer examination at other frequencies shows that non-zero **y** components exist also at higher frequencies but are very small.

VI. EXPERIMENTS

Finally, we present some results from an experimental study which was performed in order to validate the energy velocity calculations. There is a fundamental difficulty in designing a suitable experiment because the energy velocity and the group velocity differ significantly only when the attenuation is relatively large, in which case the waves decay rapidly as they propagate. The S_3 mode in the HPPE plate presents an interesting possibility because there are some locations where the group velocity rises while the energy velocity falls, yet the attenuation is generally not excessive. Figures 9 shows the predicted group and global energy velocities of the S_3 mode in solid and dotted lines, respectively; Figure 10 shows the phase velocity. The sharp rise of the group velocity at 0.25 MHz is typical of the singularity associated with an infinite slope of the wave number, as discussed earlier. The regions where the group and energy velocities differ most are also, as expected, the regions where the attenuation is highest.

The experiments were conducted using two wide band contact transducers (Panametrics), one at each end of a HPPE plate, in a through-transmission arrangement. The transducers were clamped to the ends of the plate with their axes aligned with the center line of the plate. Extensional motion of the faces of the transducers thus coupled directly with the in-plane extensional motion of the S_3 mode. Initially the length of the plate between the transducers was 170 mm, but measurements were also made over a shorter length of about 130 mm. The length in the other dimension was very much larger so that there was no possibility of receiving reflections from other boundaries. A narrow band signal consisting of 50 cycles at a chosen frequency in a Hanning window was applied to the emitter, using a Wavemaker Duet (Macro Design, Ltd.) signal generator. The received signal was captured on a digital oscilloscope (LeCroy 9400), 300

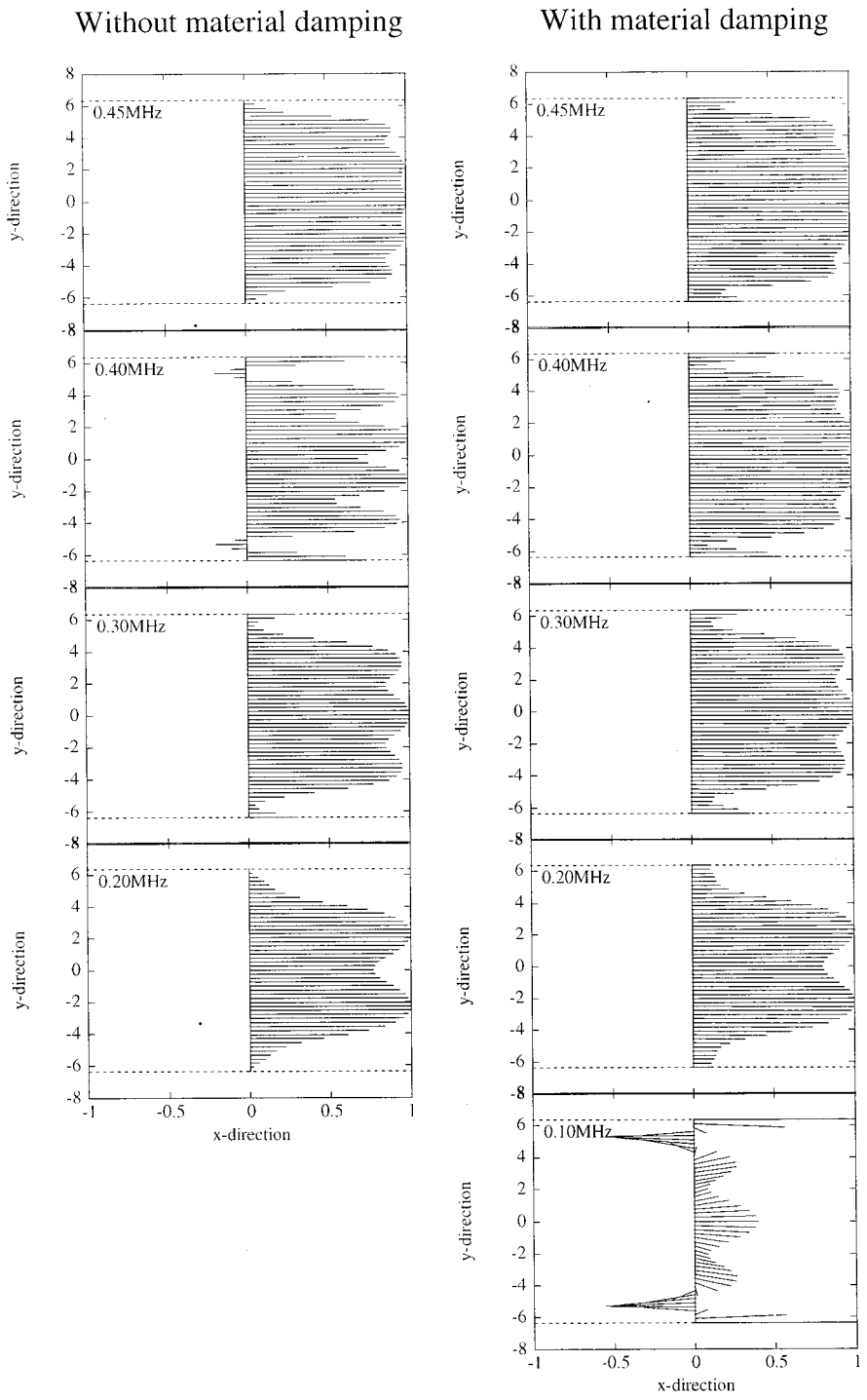


FIG. 8. Energy velocity vectors comparison between HPPE without material damping (left) and with material damping (right) calculated at different y positions.

averages taken, and stored on a PC. The measurements were repeated for a range of frequencies from 200 to 400 kHz, and for the two propagation distances.

An issue with this type of measurement is the need to separate the desired mode from any others which may be excited at a chosen frequency. In this setup the strong attenuation and the nature of the transducer coupling both contributed to ensuring that only the S_3 mode was received with any significant magnitude. In Fig. 6(b) it can be seen that the majority of the modes have very much higher attenuation than the S_3 mode within the range of frequency. In fact, the attenuation difference is such that any of these unwanted modes would be attenuated while crossing the plate by at least 50 dB more than the S_3 mode and so could comfortably be ignored. The only remaining mode with moderate attenuation is fundamentally different from S_3 because it is antisymmetric. It was therefore possible to avoid both exciting and receiving it by careful positioning of the two transducers on the axis of the plate so that both the geometry and the transduction were symmetric.

The global energy velocity and the phase velocity were calculated separately at each test frequency. The energy velocity was measured by overlaying the envelopes of the received wave packets for the two different lengths of plate, and measuring the difference between their arrival times. Of course this was only possible when the shape of the wave

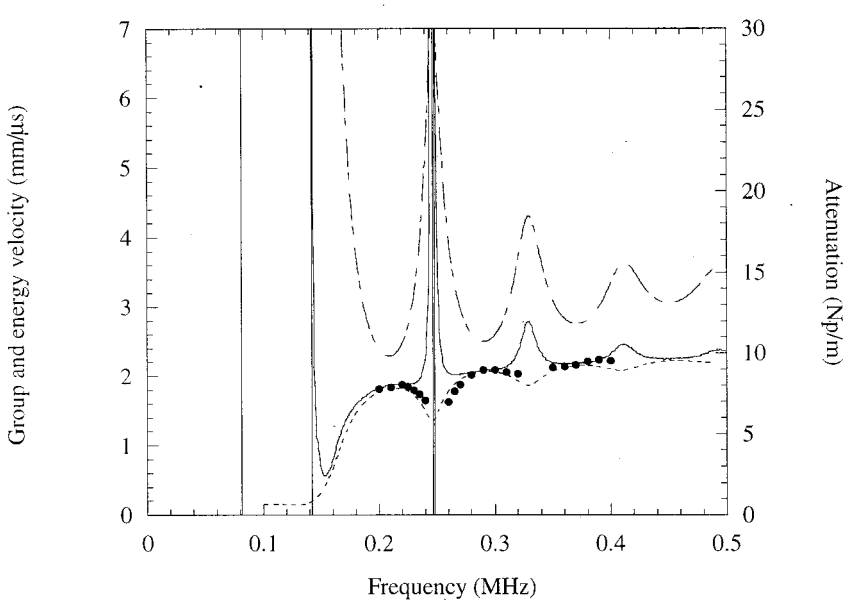


FIG. 9. Comparison between group velocity (solid line), global energy velocity [dotted line (···)], and experimental measurements (solid circles) for the mode S_3. The attenuation is represented with a dashed line (-–-).

packet remained consistent so that an accurate overlay was possible. The signal was sufficiently narrow band that in fact the envelopes retained reasonable shapes at most frequencies, so that the whole of each envelope could be matched sensibly with its counterpart when overlaying.

However there remain some gaps in the attenuation results at the locations where the attenuation was high and the signals were correspondingly too weak and distorted. The phase velocity was calculated in a similar manner, but comparing instead the arrival times of the wave cycles within the envelope. In this case the distortion of the shape of the envelope did not matter and so it was possible to obtain measurements at all of the chosen frequencies. However these measurements did require particular care in order to ensure that the correct wave cycle was compared; several additional measurements using intermediate propagation distances were used to confirm this.

The results are shown by filled circles in Figs. 9 and 10.

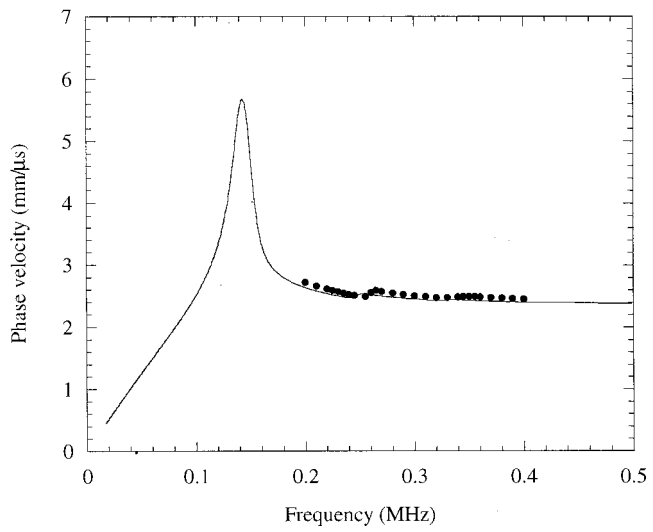


FIG. 10. Comparison between phase velocity (solid line) and experimental measurements (solid circles) for the mode S_3.

The experimental measurements can be seen to agree very well with the global energy velocity prediction. The good agreement between experiment and prediction of the phase velocities, in Fig. 10, serves also to confirm that the measurements of the S_3 mode were obtained correctly.

VII. CONCLUSIONS

The energy velocity of guided waves in flat plates has been calculated from the Poynting vector functions. In the case of the lossless Lamb modes in a vacuum-bounded non-absorbing plate, the energy velocity integral has been shown analytically to equate to the group velocity. In the cases of a non-absorbing plate immersed in water and of an absorbing plate in vacuum it has been shown numerically that the energy velocity can differ substantially from the group velocity, especially at locations on the dispersion curves where the attenuation is high. In such cases the group velocity can show discontinuities and unbelievably high values, but such features are not evident with the energy velocity. An experimental study of the S_3 mode in an absorbing plate has demonstrated good agreement of the velocity of a wave packet with the predicted energy velocity at such locations where there is divergence from the group velocity. The energy velocity should therefore be preferred as the correct measure of velocity when predicting the propagation of wave packets of attenuative modes. However, the group velocity is very much quicker to calculate and is likely to be acceptably accurate unless the attenuation is strong.

The Poynting vectors within the plate have also been plotted for these cases and demonstrate some interesting phenomena. In the lossless plate the vectors never have components in the direction normal to the direction of propagation; however the vectors in some locations in the plate can sometimes point in the opposite direction to those at other locations. In the immersed plate the vectors give clear evidence of the leakage of energy into the fluid, while in the absorbing plate the vectors indicate transfers of energy across parts of the plate but of course not across the surface boundaries.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support from the British Council for their collaboration in this research, via a British Council fellowship for A. Bernard.

[1] L. J. F. Broer, ''On the propagation of energy in linear conservative waves,'' Appl. Sci. Res., Sect. A **2**, 329–344 (1951).
[2] M. A. Biot, ''General theorems on the equivalence of group velocity and energy transport,'' Phys. Rev. **105**, 1129–1137 (1957).
[3] M. J. Lighthill, ''Group velocity,'' J. Inst. Math. Appl. **1**, 1–28 (1965).
[4] M. Hayes, ''A note on group velocity,'' Proc. R. Soc. London, Ser. A **354**, 533–535 (1977).
[5] M. Hayes and M. J. P. Musgrave, ''On energy flux and group velocity,'' Wave Motion **1**, 75–82 (1979).
[6] B. Poirée, ''Vitesse de propagation de l'énergie de l'onde plane évanescente acoustique,'' Rev. Cethedec Signal **21**(79), 103–111 (1984).
[7] P. Borejko, ''Inhomogeneous plane waves in a constrained elastic body,'' Q. J. Mech. Appl. Math. **40**, 70–87 (1987).
[8] F. Mainardi, ''Energy velocity for hyperbolic dispersive waves,'' Wave Motion **9**, 201–208 (1987).
[9] M. Deschamps, B. Poirée, and O. Poncelet, ''Energy velocity of complex harmonic plane waves in viscous fluids,'' Wave Motion **25**, 51–60 (1997).
[10] B. A. Auld, *Acoustic Fields and Waves in Solids* (Wiley Interscience, New York, 1973), Vol. 1.
[11] J. D. Achenbach, *Wave Propagation in Elastic Solids* (North-Holland, Amsterdam, 1973).
[12] R. Fiorito, W. Madigosky, and H. Uberall, ''Resonance theory of acoustic waves interacting with an elastic plate,'' J. Acoust. Soc. Am. **66**, 1857–1866 (1979).
[13] O. Lenoir, J. Duclos, J. M. Conoir, J. L. Izbicki, ''Study of Lamb waves based upon the frequency and angular derivatives of the phase of the reflection coefficient,'' J. Acoust. Soc. Am. **94**, 330–343 (1993).
[14] A. Bernard, M. Deschamps, and M. J. S. Lowe, ''Comparison between the dispersion curves calculated in complex frequency and the minima of the reflection coefficients for an embedded layer,'' J. Acoust. Soc. Am. **107**, 793–800 (2000).
[15] M. Deschamps and O. Poncelet, ''Transient Lamb waves: comparison between theory and experiments,'' J. Acoust. Soc. Am. **107**, 3120–3129 (2000).
[16] O. Poncelet and M. Deschamps, ''Lamb waves generated by complex harmonic inhomogeneous plane waves,'' J. Acoust. Soc. Am. **102**, 292–300 (1997).
[17] A. H. Meitzler, ''Backward-wave transmission of stress pulses in elastic cylinders and plates,'' J. Acoust. Soc. Am. **38**, 835–842 (1965).
[18] J. Wolf, T. D. K. Ngoc, R. Kille, and W. G. Mayer, ''Investigation of Lamb waves having a negative group velocity,'' J. Acoust. Soc. Am. **83**, 122–126 (1988).
[19] G. Kaduchak, D. H. Hughes, and P. L. Marston, ''Enhancement of the backscattering of high-frequency tone bursts by thin spherical shells associated with a backwards wave: observations and ray approximation,'' J. Acoust. Soc. Am. **96**, 3704–3714 (1994).
[20] G. Kaduchak and P. L. Marston, ''Traveling-wave decomposition of surface displacements associated with scattering by a cylindrical shell: numerical evaluation displaying guided forward and backward wave properties,'' J. Acoust. Soc. Am. **98**, 3501–3507 (1995).
[21] C. W. Chan and P. Cawley, ''Lamb waves in highly attenuative plastic plates.'' J. Acoust. Soc. Am. **104**, 874–881 (1998).
[22] C. W. Chan and P. Cawley, ''The existence of low loss Lamb modes in highly attenuative media,'' in *Review of Progress in Quantitative NDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1995), Vol. 12, pp. 193–200.
[23] M. Lowe, ''Matrix techniques for modeling ultrasonic waves in multilayered media,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42**, 525–542 (1995).
[24] B. Pavlakovic, M. J. S. Lowe, D. N. Alleyne, and P. Cawley, ''Disperse: a general purpose program for creating dispersion curves,'' in *Review of Progress in Quantitative NDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1997), Vol. 16, pp. 185–192.

Quasistatic coupling coefficients for electrostrictive ceramics

Jean C. Piquette[a)]
Naval Undersea Warfare Center, Division Newport, 1176 Howell Street, Newport, Rhode Island 02841

(Received 12 August 2000; accepted for publication 13 April 2001)

A generalized definition of the coupling coefficient, useful for the evaluation of transducers that incorporate an electrostrictive active element, is introduced. The definition is expressed under quasistatic conditions and becomes zero when no bias is applied (assuming that the effects of remanence are negligible), and remains zero under zero bias even when a significant prestress is present. This reflects a property of the piezoelectric coupling coefficient, which vanishes when the ceramic becomes depoled. The behavior of this definition thus differs from that of another definition, introduced elsewhere, which produces a significant nonzero result at zero bias. [See C. Hom *et al.*, ''Calculation of quasi-static electromechanical coupling coefficients for electrostrictive ceramic materials,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **41**, 542–551 (1994).] The present definition also leads in a natural way to a coupling coefficient for biased piezoelectric ceramics, and an equation is given for that case. Moreover, in the case of a biased electrostrictive ceramic it is found that a coupling coefficient derived from an equivalent circuit [J. C. Piquette and S. E. Forsythe, ''Generalized material model for lead magnesium niobate (PMN) and an associated electromechanical equivalent circuit,'' J. Acoust. Soc. Am. **104**, 2763–2772 (1998)] gives an excellent approximation to the exact value, and is found to be accurate to within a few percent for drive amplitudes as high as 75% of the bias. It is shown that maximizing the coupling coefficient automatically discriminates against transducer designs (and operating conditions) that would produce significant harmonic distortion. [DOI: 10.1121/1.1377291]

PACS numbers: 43.38.Ar, 43.20.Px, 43.30.Yj, 43.58.Vb [SLE]

I. INTRODUCTION

Electrostrictive ceramics [such as lead magnesium niobate (PMN)] have gained widely in interest in recent years. The increased interest derives from the fact that these materials exhibit significantly larger strains for a given electric field strength than those exhibited by the standard piezoelectric ceramics [such as lead zirconate titanate (PZT)]. Since the coupling coefficient is widely used as a metric for distinguishing between candidate transducer materials,[1] it is important that a useful definition be formulated for the case of electrostrictive ceramics. (The metric also subsumes the operating conditions, such as the bias.)

A definition of the coupling coefficient for electrostrictive materials was proposed in Ref. 2. That definition is based on a generalization of the IEEE Standard on Piezoelectricity,[3] which was originally developed for linear piezoelectric materials. The definition advanced in Ref. 2 produces significant nonzero values when applied to unbiased electrostrictive materials. Thus that definition may not be suitable for comparing candidate electrostrictive transducer configurations, since it does not reject the unbiased transducer as unsuitable and might even produce a higher value for an unbiased transducer than for a biased one. [An unbiased electrostrictive transducer must be rejected as unsuitable, at least if the first harmonic transducer response is of primary interest (as it is in the current work), owing to the fact that its output would be dominated by harmonics of the drive frequency. That is, harmonic distortion of the output would be extremely large, and the total harmonic distortion could easily exceed 100%.]

In electrostrictive materials, where only a negligible remanent polarization is typically present, the bias serves the role of providing an ''effective'' remnant. That is, a biased electrostrictive ceramic behaves much like an ordinary piezoelectric ceramic. Thus, at least insofar as its performance in a transducer is concerned, an electrostrictive ceramic that is unbiased would be expected to behave much like a piezoelectric ceramic that has become depoled. A fully depoled piezoelectric ceramic would have a null piezoelectric d constant and, in view of the direct dependence of the coupling coefficient on this parameter,[4] the coupling coefficient would vanish in such a circumstance. Hence, a primary goal of the work reported here was to develop a definition of the coupling coefficient that produces a zero value for an unbiased electrostrictive ceramic (assuming remanence is negligible), in order that it reflect this property of piezoelectric ceramics. Of course to be useful, such a generalization must also recover all the known characteristics of the standard coupling coefficient.

The definition of the coupling coefficient introduced here generalizes a definition advanced by Berlincourt *et al.*[5] The approach of Ref. 5 is based on evaluating the change in the energy stored in an active material, which has both elastic and dielectric properties, during a process where the material does mechanical work on its surroundings while an electric field is applied. The energy change is calculated through the use of an expression presented by Mason.[6] The resulting expression is separated into elastic, dielectric, and

a)Electronic mail: piquettejc@npt.nuwc.navy.mil

mutual energy terms. These terms are then combined in a particular way to produce the definition.

The definition can be generalized to electrostrictive materials by applying the approach of Ref. 5 to the theory of Refs. 7 and 8, with the zeroes of the physical variables redefined suitably. The coupling coefficient defined in this way not only has the desired property of vanishing for an unbiased electrostrictive material, but also accurately reproduces, for remarkably high drive levels, the behavior of a coupling coefficient based on a linearized version of the theory (as incorporated in an equivalent circuit[8]). Also recovered is the ordinary piezoelectric coupling coefficient.[4] Moreover, an expression suitable for application to biased piezoelectric ceramics is obtained. (The use of biased piezoelectric ceramics has recently gained in interest.[9])

In Sec. II the methodology used to generalize the coupling-coefficient definition is initially developed by considering linear piezoelectricity. Although a nonzero bias is considered in this section, the results obtained do not truly represent the behavior of a piezoelectric ceramic under the application of a bias, since such materials are actually nonlinear. The methods developed in Sec. II are applied to electrostrictive ceramics in Sec. III. Consistency between the piezoelectric and electrostrictive theories is achieved by redefining the zeroes of the physical variables that describe system behavior. In Sec. IV, three applications are considered. First, the case of a realistic biased piezoelectric ceramic, viewed as an electrostrictive ceramic that retains a significant remanent polarization, is treated in detail. An explicit expression for the coupling coefficient is derived for this case. Next an unbiased electrostrictive ceramic, which has a negligible remnant, is considered. It is shown that the definition introduced in Sec. III indeed produces a coupling coefficient that is identically zero for this case. Finally, the biased electrostrictive ceramic is considered. It is shown that the relatively simple formula for the coupling coefficient resulting from an equivalent circuit for an electrostrictive transducer is fully consistent with the general formula developed in Sec. III. The simple equivalent-circuit formula is also shown to give a very accurate estimate (to within a few percent of the exact result), even for drives as high as 75% of the bias. A discussion is given in Sec. V, and a summary and conclusion are given in Sec. VI.

II. BIASED *LINEAR* PIEZOELECTRIC CERAMIC

The case of the end-electroded length expander bar[10,11] is used to derive the results presented here and in the following sections. Extensions to other geometries would be relatively straightforward.

The equations relevant to the end-electroded length expander bar, for the case of *linear* piezoelectricity,[10] are

$$S_3 = s_{33}^D T_3 + g_{33} D_3 \tag{1}$$

and

$$E_3 = -g_{33} T_3 + \beta_{33}^T D_3 . \tag{2}$$

(The usual piezoelectric notations, as defined in Ref. 5, are used here, with the ''3'' direction denoting the long axis of the bar.) The energy differential for an active material exhibiting both elastic and dielectric properties (neglecting thermal and magnetic effects) is[6]

$$dU = T_3 dS_3 + E_3 \, dD_3 . \tag{3}$$

(The expression given in Ref. 6 for dU has been simplified for the case of the bar, with the relevant boundary conditions[10] applied.) Suppose now that a biasing electric field E_{BIAS} and a monofrequency ac electric field of amplitude E_{drive} are applied to the material such that

$$E_3 = E_{\mathrm{bias}} - E_{\mathrm{drive}} \cos \omega t . \tag{4}$$

Here, the phasing of the ac component has been chosen for convenience. The quantity ω denotes the angular frequency, where $\omega = 2\pi/\tau$, and τ is the period. (All constants throughout this article are assumed positive.)

The bias field E_{bias} will produce a strain in the material. A new set of variables is now introduced to eliminate the explicit appearance of such nonzero (additive) constant quantities from Eqs. (1) and (2), while still retaining the structure of the original equations. In particular, introduce new variables with carets $\hat{E}_3$, $\hat{D}_3$, and $\hat{S}_3$ such that

$$\hat{E}_3 \equiv E_3 - E_{\mathrm{bias}} + E_{\mathrm{drive}} , \tag{5}$$

$$\hat{D}_3 \equiv D_3 - \left(\frac{E_{\mathrm{bias}} - E_{\mathrm{drive}}}{\beta_{33}^T} \right), \tag{6}$$

and

$$\hat{S}_3 \equiv S_3 - g_{33} \left(\frac{E_{\mathrm{bias}} - E_{\mathrm{drive}}}{\beta_{33}^T} \right). \tag{7}$$

[The original stress variable T_3 is retained in the new variable set. However, had a nonzero prestress been considered, a new stress variable would also have to be introduced. The definitions of Eqs. (5)–(7) would also be affected by the presence of such a prestress.] With these definitions solved for the original variables E_3, D_3, and S_3; and with the results substituted into Eqs. (1) and (2), we obtain

$$\hat{S}_3 = s_{33}^D T_3 + g_{33} \hat{D}_3 \tag{8}$$

and

$$\hat{E}_3 = -g_{33} T_3 + \beta_{33}^T \hat{D}_3 , \tag{9}$$

which obviously have the same structure as Eqs. (1) and (2). From this point onward, the formulation of the coupling coefficient will be effected in terms of the new variables defined in Eqs. (5)–(7), and in terms of the new equation set expressed in Eqs. (8) and (9). Consistent with this point of view, the physical variables E_3, D_3, and S_3 of Eq. (3) are to be replaced by their hatted counterparts.

To deduce a coupling coefficient for this system, we will adopt the approach of Berlincourt *et al.*[5] Specifically, we will compute the change in system energy under the application of an electric field, in a circumstance in which the piezoelectric material performs mechanical work. The coupling coefficient is then determined through the identification of elastic, dielectric, and mutual energy terms, denoted U_e, U_d, and U_m, respectively. We consider the calculation of the changes in system energy that occur when the electric field

of Eq. (4) is applied over the half-cycle occurring between $t=0$ and $t=\tau/2$. (The use of a half cycle is for calculational convenience only, and has no special significance. However, it should be noted that over this half cycle the applied field E_3 rises from its minimum value of $E_{\text{bias}}-E_{\text{drive}}$ to its maximum value of $E_{\text{bias}}+E_{\text{drive}}$.)

Using Eqs. (4) and (5), the new electric field variable $\hat{E}_3$ can be reexpressed as

$$\hat{E}_3=E_{\text{drive}}-E_{\text{drive}}\cos\omega t. \tag{10}$$

From Eq. (10) it is seen that when $t=0$, $\hat{E}_3=0$; and when $t=\tau/2$, $\hat{E}_3=2E_{\text{drive}}$. Also, assuming no applied external stress ($T_3=0$) at $t=0$, it can be seen from Eqs. (8) and (9) that $\hat{D}_3=0$ and $\hat{S}_3=0$ at $t=0$ as well. Thus, in integrating the differential of energy given by Eq. (3) (with the original variables replaced by the new variables with carets) over the half-cycle being considered, the lower limits of integration are each zero. [It was in order that such zero lower limits would obtain that the variable redefinitions given in Eqs. (5)–(7) were introduced.] Thus, the total change in system energy over this half cycle is

$$\Delta U=\int_0^{\hat{S}_F}T_3\,d\hat{S}_3+\int_0^{\hat{D}_F}\hat{E}_3\,d\hat{D}_3, \tag{11}$$

where the upper limits of integration $\hat{S}_F$ and $\hat{D}_F$ represent the ''final'' values of the strain with variables and displacement variables, respectively, occurring at $t=\tau/2$.

In order that the piezoelectric material perform nonzero mechanical work during the application of the half-cycle drive currently under consideration, a condition will now be imposed. For this condition to be satisfied, a nonzero external stress must act on the material. (However, this stress is zero at $t=0$.) During the application of the half-cycle of electric field, it is assumed that an external stress is applied to the material that causes the strain to change in direct proportion to the displacement field. That is, it is assumed that, owing to the application of an external stress, the strain and displacement field are related as

$$\hat{S}_3=-\gamma\,\hat{D}_3. \tag{12}$$

Here, γ is a constant of proportionality, with suitable dimensions, assumed to be positive. [The explicit minus sign of Eq. (12) is introduced for consistency with the conventional definition of the coupling coefficient.] Again, Eq. (12) is introduced simply for calculational convenience, and has no other special significance. Notice, however, that Eq. (12) is consistent with the requirement that the variables $\hat{D}_3$ and $\hat{S}_3$ both be zero at $t=0$, so that the lower limits of integration in Eq. (11) are not affected by imposing this condition.

Prior to carrying out the integrations of Eq. (11), it is helpful to first solve Eq. (8) for T_3, giving

$$T_3=\frac{\hat{S}_3}{s_{33}^D}-\frac{g_{33}}{s_{33}^D}\hat{D}_3. \tag{13}$$

[It is worthwhile noting here that the external stress required to enforce Eq. (12) can be determined by substituting Eq. (12) into Eq. (13).] Substituting Eq. (13) into Eq. (9) produces

$$\hat{E}_3=-\frac{g_{33}}{s_{33}^D}\hat{S}_3+\left(\frac{g_{33}^2}{s_{33}^D}+\beta_{33}^T\right)\hat{D}_3. \tag{14}$$

Substituting Eqs. (13) and (14) into Eq. (11) gives

$$\Delta U=\frac{1}{s_{33}^D}\int_0^{\hat{S}_F}\hat{S}_3\,d\hat{S}_3-\frac{g_{33}}{s_{33}^D}\int_0^{\hat{S}_F}\hat{D}_3\,d\hat{S}_3-\frac{g_{33}}{s_{33}^D}\int_0^{\hat{D}_F}\hat{S}_3\,d\hat{D}_3+\left(\frac{g_{33}^2}{s_{33}^D}+\beta_{33}^T\right)\int_0^{\hat{D}_F}\hat{D}_3\,d\hat{D}_3. \tag{15}$$

In the form of Eq. (15), the identification of elastic, dielectric, and mutual energy terms is straightforward. The elastic energy is associated with purely mechanical processes, and so involves integrals containing only the strain variable [the first term of Eq. (15)]. The dielectric energy is associated with purely electrical processes, and so involves integrals containing only the displacement field variable [the last term of Eq. (15)]. Finally, the mutual energy is associated with mechanical and electrical interactions, and so involves integrals containing products of the strain and displacement field variables [the second and third terms of Eq. (15)]. Therefore we have

$$U_e=\frac{1}{s_{33}^D}\int_0^{\hat{S}_F}\hat{S}_3\,d\hat{S}_3\quad\text{(elastic energy)}, \tag{16}$$

$$U_d=\left(\frac{g_{33}^2}{s_{33}^D}+\beta_{33}^T\right)\int_0^{\hat{D}_F}\hat{D}_3\,d\hat{D}_3\quad\text{(dielectric energy)}, \tag{17}$$

and

$$2U_m=-\frac{g_{33}}{s_{33}^D}\int_0^{\hat{S}_F}\hat{D}_3\,d\hat{S}_3-\frac{g_{33}}{s_{33}^D}\int_0^{\hat{D}_F}\hat{S}_3\,d\hat{D}_3\quad\text{(mutual energy)}. \tag{18}$$

[The factor of 2 appearing on the left-hand side of Eq. (18) follows a convention introduced in Ref. 5, Eq. (29).]

Carrying out the integrals of Eqs. (16)–(18) and eliminating $\hat{S}_F$ in favor of $\hat{D}_F$ by using Eq. (12) gives

$$U_e=\frac{1}{2}\gamma^2\frac{1}{s_{33}^D}\hat{D}_F^2, \tag{19}$$

$$U_d=\frac{1}{2}\left(\frac{g_{33}^2}{s_{33}^D}+\beta_{33}^T\right)\hat{D}_F^2, \tag{20}$$

and

$$U_m=\frac{1}{2}\frac{g_{33}}{s_{33}^D}\gamma\,\hat{D}_F^2. \tag{21}$$

Next we use the formula[12] for the coupling coefficient given in Ref. 5, viz.

$$k_{33}=\frac{U_m}{\sqrt{U_eU_d}}. \tag{22}$$

Substituting Eqs. (19)–(21) into Eq. (22) gives

$$k_{33}=\frac{g_{33}}{\sqrt{g_{33}^2+\beta_{33}^T s_{33}^D}}. \tag{23}$$

By using the formulas[13–15]

$$g_{33}=\beta_{33}^T d_{33}, \tag{24}$$

$$\beta_{33}^T=\frac{1}{\epsilon_0 \epsilon_{33}^T}, \tag{25}$$

and

$$s_{33}^E=s_{33}^D+\frac{g_{33}^2}{\beta_{33}^T}, \tag{26}$$

it is not too difficult to put Eq. (23) into the standard form,[16]

$$k_{33}=\frac{d_{33}}{\sqrt{\epsilon_0\ \epsilon_{33}^T\ s_{33}^E}}. \tag{27}$$

The above-mentioned derivation was based on generalizing the theory of a *linear* piezoelectric material to account for the presence of a bias voltage. As can be seen from Eq. (27), the conclusion is that the coupling coefficient for such a material is identical to that for the unbiased case. [This conclusion can also be seen to be reasonable by examining Ref. 3, p. 39, Fig. 12(b). For a linear piezoelectric material, the slopes of the lines in Fig. 12(b), which represent the permittivities at constant stress and at constant strain, would be unaffected by the application of a bias. For consistency with the treatment of the coupling coefficient given in Ref. 3, as well as the treatment given here, the zeroes of the physical variables would have to be redefined such that the geometry of Fig. 12(b) is unaffected. Thus, the shaded areas shown in the figure would not be modified by a bias. It follows that the coupling coefficient remains unchanged.] However, piezoelectric materials are not truly linear, so this conclusion does not apply to them. Piezoelectric materials can be treated as electrostrictive materials that retain a large remanent pole[17] (and they will be so treated here). The consideration of realistic biased piezoelectric materials will therefore be based on the theory of electrostriction, and this matter is taken up in Sec III. The *linear* theory of piezoelectric materials was investigated in the present section not to obtain a realistic expression for the coupling coefficient for that case, but rather to establish the methodology. A realistic coupling coefficient for biased piezoelectric ceramics is derived in Sec. IV.

III. GENERALIZATION TO ELECTROSTRICTIVE CERAMICS

We now use the calculations of Sec. II as a guide in producing a procedure suitable for calculation of the coupling coefficient of an electrostrictive active element. (Since the piezoelectric transducer can be considered to be a special case of the electrostrictive transducer, one in which a large remanent polarization is present, we expect the piezoelectric case, biased as well as unbiased, to be subsumed by this treatment.)

We use for the basis of the generalized calculation the model of electrostriction presented in Refs. 7 and 8. The basic (one-dimensional) equations produced by that model for the end-electroded length expander bar are

$$E_3=\frac{D_3-P_0}{\sqrt{(\epsilon_0\epsilon_{33}^T)^2-a(D_3-P_0)^2}}-2Q_{33}T_3D_3 \tag{28}$$

and

$$S_3=s_{33}^D T_3+Q_{33}\ D_3^2. \tag{29}$$

Here Q_{33} is the electrostriction constant, ϵ_{33}^T is the low-field dielectric permittivity, P_0 is the remanent polarization (retained for generality, so that piezoelectricity is subsumed), and a is the ''saturation parameter,''[7] a measure of the rate at which the polarization saturates. The remaining quantities are standard piezoelectric notations. (cf. Refs. 7 and 8 for a complete description).

As in the piezoelectric case considered in Sec. II, we will again consider a case in which the electric field applied to the material is of the form of that given in Eq. (4). Also as before, it will be necessary to introduce a modified variable set to effect the calculation of the coupling coefficient in terms of quantities that are zero at the beginning of the drive cycle. For generality we consider the prestress T_0 to be nonzero. For convenience in what follows, we introduce the notation

$$E_{\text{ref}}\equiv E_{\text{bias}}-E_{\text{drive}}, \tag{30}$$

where E_{ref} will be termed the ''reference'' electric field. It denotes the minimum of the electric field of Eq. (4). In terms of E_{ref}, Eq. (5) becomes

$$\hat{E}_3=E_3-E_{\text{ref}}. \tag{31}$$

It is convenient to also introduce a reference displacement field D_{ref}, defined to be the solution of

$$E_{\text{ref}}=\frac{D_{\text{ref}}}{\sqrt{(\epsilon_0\epsilon_{33}^T)^2-aD_{\text{ref}}^2}}-2Q_{33}T_0(P_0+D_{\text{ref}}). \tag{32}$$

This equation results when $E_3=E_{\text{ref}}$, $T=T_0$, and $D_3=P_0+D_{\text{ref}}$ are substituted into Eq. (28). Equation (32) can readily be reexpressed in the form of a fourth-degree polynomial in D_{ref}. If $T_0=0$, Eq. (32) reduces to a quadratic equation, and we find the simple solution $D_{\text{ref}}^{T_a=0}=\epsilon_0\epsilon_{33}^T E_{\text{ref}}/\sqrt{1+aE_{\text{ref}}^2}$.

In the spirit of the treatment presented for the piezoelectric case we now introduce a new variable set, which yields zero values of the variables at $t=0$ for the drive of Eq. (4) or, equivalently, the drive of Eq. (10). In particular, we introduce the hatted variables

$$\hat{D}_3\equiv D_3-(P_0+D_{\text{ref}}), \tag{33}$$

$$\hat{S}_3\equiv S_3-s_{33}^D\ T_0-Q_{33}(P_0+D_{\text{ref}})^2, \tag{34}$$

and,

$$\hat{T}_3\equiv T_3-T_0. \tag{35}$$

Substituting Eqs. (31) and (33)–(35) into Eqs. (28) and (29) produces the new equation set

$$\hat{E}_3 = -E_{\text{ref}} + \frac{\hat{D}_3 + D_{\text{ref}}}{\sqrt{(\epsilon_0 \epsilon_{33}^T)^2 - a(\hat{D}_3 + D_{\text{ref}})^2}} - 2Q_{33}\hat{T}_3\hat{D}_3$$
$$- 2Q_{33}(P_0 + D_{\text{ref}})\hat{T}_3 - 2Q_{33}T_0\hat{D}_3$$
$$- 2Q_{33}T_0(P_0 + D_{\text{ref}}) \tag{36}$$

and

$$\hat{S}_3 = s_{33}^D \hat{T}_3 + Q_{33}\hat{D}_3^2 + 2Q_{33}(P_0 + D_{\text{ref}})\hat{D}_3. \tag{37}$$

When $\hat{E}_3 = 0$ and $\hat{T}_3 = 0$ (note that these values apply at $t = 0$ for the half-cycle drive under consideration), Eq. (36) can be used to show that $\hat{D}_3 = 0$ also at $t = 0$. [This is easily seen from the fact that the simultaneous vanishing of $\hat{E}_3$, $\hat{T}_3$, and $\hat{D}_3$ causes Eq. (36) to reduce to the definition of D_{ref} given by Eq. (32), so that the satisfaction of the resulting equation is guaranteed by the manner in which D_{ref} has been defined.] Moreover when $\hat{D}_3 = 0$ and $\hat{T}_3 = 0$ are both substituted into Eq. (37), it is seen that $\hat{S}_3 = 0$ as well. This establishes the fact that the new variable set given by Eqs. (31) and (33)–(35) indeed provides zero values of the variables at $t = 0$, as required for the application of Eq. (11), with T_3 replaced by $\hat{T}_3$, to the present case.

As before, we next rearrange the new equation set in order to render the identification of the elastic, dielectric, and mutual energy terms a straightforward process. We begin by solving Eq. (37) for $\hat{T}_3$, giving

$$\hat{T}_3 = \frac{\hat{S}_3}{s_{33}^D} - \frac{Q_{33}}{s_{33}^D}\hat{D}_3^2 - 2\frac{Q_{33}}{s_{33}^D}(P_0 + D_{\text{ref}})\hat{D}_3. \tag{38}$$

Next Eq. (38) is substituted into Eq. (36) to obtain

$$\hat{E}_3 = -E_{\text{ref}} + \frac{\hat{D}_3 + D_{\text{ref}}}{\sqrt{(\epsilon_0 \epsilon_{33}^T)^2 - a(\hat{D}_3 + D_{\text{ref}})^2}} - \frac{2Q_{33}\hat{D}_3\hat{S}_3}{s_{33}^D}$$
$$+ \frac{2Q_{33}^2\hat{D}_3^3}{s_{33}^D} + \frac{6Q_{33}^2}{s_{33}^D}(P_0 + D_{\text{ref}})\hat{D}_3^2$$
$$- \frac{2Q_{33}}{s_{33}^D}(P_0 + D_{\text{ref}})\hat{S}_3 + \frac{4Q_{33}^2}{s_{33}^D}(P_0 + D_{\text{ref}})^2\hat{D}_3$$
$$- 2Q_{33}T_0\hat{D}_3 - 2Q_{33}T_0(P_0 + D_{\text{ref}}). \tag{39}$$

Once again as in the piezoelectric case, we substitute Eqs. (38) and (39) into Eq. (11), substituting $\hat{T}_3$ for T_3. We proceed immediately to break the resulting expression into appropriate individual energy terms according to whether only the strain variable appears (elastic), only the displacement field variable appears (dielectric), or both variables appear (mutual) in each integrand. The results are

$$U_e = \frac{1}{s_{33}^D}\int_0^{\hat{S}_F} \hat{S}_3\, d\hat{S}_3, \quad \text{(elastic energy)} \tag{40}$$

$$U_d = \int_0^{\hat{D}_F} -[E_{\text{ref}} + 2Q_{33}T_0(P_0 + D_{\text{ref}})]d\hat{D}_3$$
$$+ \int_0^{\hat{D}_F} \frac{\hat{D}_3 + D_{\text{ref}}}{\sqrt{(\epsilon_0 \epsilon_{33}^T)^2 - a(\hat{D}_3 + D_{\text{ref}})^2}} d\hat{D}_3$$
$$+ \frac{2Q_{33}^2}{s_{33}^D}\int_0^{\hat{D}_F} \hat{D}_3^3\, d\hat{D}_3 + \frac{6Q_{33}^2(P_0 + D_{\text{ref}})}{s_{33}^D}$$
$$\times \int_0^{\hat{D}_F} \hat{D}_3^2\, d\hat{D}_3 + \left[\frac{4Q_{33}^2}{s_{33}^D}(P_0 + D_{\text{ref}})^2 - 2Q_{33}T_0\right]$$
$$\times \int_0^{\hat{D}_F} \hat{D}_3\, d\hat{D}_3, \quad \text{(dielectric energy)} \tag{41}$$

and

$$2U_m = \frac{-Q_{33}}{s_{33}^D}\int_0^{\hat{S}_F} \hat{D}_3^2\, d\hat{S}_3 - \frac{2Q_{33}}{s_{33}^D}(P_0 + D_{\text{ref}})\int_0^{\hat{S}_F} \hat{D}_3\, d\hat{S}_3$$
$$- \frac{2Q_{33}}{s_{33}^D}\int_0^{\hat{D}_F} \hat{D}_3\hat{S}_3\, d\hat{D}_3 - \frac{2Q_{33}}{s_{33}^D}$$
$$\times (P_0 + D_{\text{ref}})\int_0^{\hat{D}_F} \hat{S}_3\, d\hat{D}_3, \quad \text{(mutual energy)}. \tag{42}$$

[Once again the conventional factor of 2 from Ref. 5, Eq. (29), appears in Eq. (42).]

To complete the calculation of the coupling coefficient, Eq. (12) is imposed and the integrals of Eqs. (40)–(42) are carried out. One subtlety that arises in the calculation for the present case that did not arise in the piezoelectric case, however, is that a direct substitution of the results of these integrations into Eq. (22) produces an expression that depends weakly on the parameter γ. The difference in mathematical behavior from the piezoelectric case arises from the underlying electrostrictive model, which is nonlinear. To eliminate this dependence of the result upon γ, we modify Eq. (22) slightly so that

$$k_{33} \equiv \lim_{\gamma \to 0} \frac{U_m}{\sqrt{U_e U_d}}. \tag{43}$$

Referring back to Eq. (12), we note that the limit introduced in Eq. (43) suggests an infinitesimal change in strain. This infinitesimal change in strain, despite the presence of what might be a very large driving electric field, results from the influence of the applied external stress required to enforce Eq. (12). However the appearance of an infinitesimal strain does not mean that the present definition does not apply to circumstances in which large strains occur, a case of significant interest here. (In a similar way, the appearance of a specific drive cycle within the IEEE Standard on Piezoelectricity[3] does not limit the applicability of the resulting coupling coefficient definition to circumstances in which that specific drive cycle arises.) As noted previously the dependence of k_{33} upon the parameter γ prior to imposing the

limit used in Eq. (43) is extremely weak, and the weakness of that dependence is important. Indeed the dependence is so weak that, instead of imposing the given limit, a numerical value of γ, which when substituted into Eq. (12) would yield a strain that *exceeds* that produced by the applied field alone, could have been imposed instead without the resulting value of k_{33} being significantly affected. This issue will be taken up again and explored in some detail in Sec. IV C, where a specific numerical example is considered.

Carrying out the integrals of Eqs. (40)–(42) and applying Eq. (43) (which constitutes the definition of the coupling coefficient being advanced here), gives

$$k_{33}^2 = \frac{aQ_{33}^2 \hat{D}_F^2 [\hat{D}_F + 2(P_0 + D_{\text{ref}})]^2}{2s_{33}^D A + a\hat{D}_F B}, \tag{44}$$

where

$$A \equiv \sqrt{(\epsilon_0 \epsilon_{33}^T)^2 - aD_{\text{ref}}^2} - \sqrt{(\epsilon_0 \epsilon_{33}^T)^2 - a(\hat{D}_F + D_{\text{ref}})^2},$$

and

$$\equiv Q_{33}[\hat{D}_F + 2(P_0 + D_{\text{ref}})] \times \{Q_{33}\hat{D}_F[\hat{D}_F + 2(P_0 + D_{\text{ref}}) - 2s_{33}^D T_0]\} - 2s_{33}^D E_{\text{ref}}.$$

Here $\hat{D}_F$ is the solution of Eq. (39) for $\hat{D}_3$ with $\hat{E}_3$ set equal to $2E_{\text{drive}}$ [i.e., the maximum hatted E field achieved under the drive given by Eq. (10)], imposing the condition of Eq. (12), with E_{ref} given by Eq. (30) and D_{ref} the solution of Eq. (32). When Eq. (39) is evaluated in this way, it can be put into the form of an eighth-degree polynomial, and is readily solved by standard numerical methods. Equation (44) is the desired coupling coefficient for an electrostrictive end-electroded length expander bar, and constitutes the primary result of the present paper.

IV. APPLICATIONS

A. Biased piezoelectric ceramic

Biased operation of piezoelectric materials has been considered in the past,[18,19] as well as more recently.[9] It is therefore worthwhile to deduce a theoretical expression for this case. (For the purposes of the following calculations we return to assuming negligible prestress, i.e., $T_0 = 0$.)

The formulas given for the elastic, dielectric, and mutual energies for the piezoelectric case [Eqs. (16)–(18)] and those given for electrostriction [Eqs. (40)–(42)] can be compared to establish a relationship between the two material types. In doing so, we assume the drive applied to the biased piezoelectric ceramic is such that no significant harmonics above the first are produced in either the strain or the displacement field.

It will be helpful to recall certain relations between the usual piezoelectric constants and the constants of the electrostrictive theory,[8] viz.

$$\beta_{33}^T d_{33} = g_{33} = 2Q_{33}P_0. \tag{45}$$

Consider now Eq. (17), the dielectric energy for the *linear* case. Focusing on the integral appearing there, we note it contains the product $\hat{D}_3 d\hat{D}_3$. It should be understood that the quantity $\hat{D}_3$ can have only a constant and a purely first-harmonic Fourier component, in view of the assumed linearity of the piezoelectric system for which Eq. (17) was developed. Obviously then, the differential $d\hat{D}_3$ can have only a purely first harmonic component. The nature of this integral can be used as a guide in simplifying Eqs. (40)–(42) for application to a realistic biased piezoelectric ceramic (which is being treated here as an electrostrictive ceramic that retains a significant remnant),[17] where we are concerned with the linear response. We therefore seek to simplify these equations such that the surviving expressions contain essentially purely first-harmonic integrands (possibly including constant components) that multiply essentially purely first-harmonic differentials, in accordance with the nature of Eq. (17). [Equations (16) and (18) clearly also have the same harmonic character as Eq. (17), so that similar analogies can be made with the corresponding integrals in the electrostrictive case.] Since it is assumed that the material is driven in such a way that no significant harmonic contributions above the first appear in $\hat{S}_3$ and $\hat{D}_3$, it is clear that the differentials appearing in Eqs. (40)–(42) already have the required character. We may thus focus our attention on identifying terms in Eqs. (40)–(42) that significantly affect the first-harmonic components *of the integrands*.

Typically, when a biased piezoelectric material is driven so that no significant harmonics above the first appear in the strain or the displacement field, the constant component of the hatted D field is of the same order of magnitude as the hatted first harmonic. (This was verified in numerical tests.) This behavior reflects the equality of the constant and first-harmonic components of the hatted E field, defined in Eq. (10). However, both the constant and first harmonic components of both $\hat{S}_3$ and $\hat{D}_3$ are much smaller than the constants that have been subtracted out in the definitions given in Eqs. (33) and (34). Thus, it is the numerical value of the combination $P_0 + D_{\text{ref}}$ that controls the preponderance of the constant contributions to the strain and the displacement field. (Recall it is being assumed here that $T_0 = 0$.)

Under the given circumstances, the hatted D field may be considered to have the approximate form $\hat{D}_3 = \xi[c_0 + c_1 \sin(\omega t) + c_2 \cos(\omega t)]$, where c_0, c_1, c_2 are (dimensionless) constants, and $\xi/(P_0 + D_{\text{ref}}) \ll 1$. [Since this inequality is used in what follows as the basis for simplifying Eqs. (40)–(42), its satisfaction should be considered to be the quantitative criterion for the applicability of the simplified formulas developed in the current section. If this criterion is not satisfied, Eq. (44) should be preferred for determining the coupling coefficient. The quantities c_0 and $(c_1^2 + c_2^2)^{1/2}$ are of order unity, so that $\hat{D}_3$ is of order ξ.] This expression for $\hat{D}_3$ can be used to compare the three terms in Eq. (41) whose integrands consist of a simple power of $\hat{D}_3$. Ignoring the purely numerical constants, but accounting for the $(P_0 + D_{\text{ref}})$ factors in the coefficients, it can be seen that the term containing the $\hat{D}_3^2$ integrand is smaller than that containing only $\hat{D}_3$ by a factor of $\xi/(P_0 + D_{\text{ref}})$. (Assuming again that $T_0 = 0$.) Similarly, the term containing $\hat{D}_3^3$ is smaller than that containing only $\hat{D}_3$ by a factor of $[\xi/(P_0 + D_{\text{ref}})]^2$. Thus, the

terms containing the $\hat{D}_3^2$ and $\hat{D}_3^3$ integrands may both safely be ignored relative to that containing just $\hat{D}_3$, under the assumed criterion and negligible prestress.

Arguments similar to those just given can be used to show, once again in the circumstances currently under consideration, that the terms in Eq. (42) whose integrands contain $\hat{D}_3^2$ or the product $\hat{D}_3\hat{S}_3$ are negligible compared with the terms whose integrands contain only $\hat{D}_3$ or only $\hat{S}_3$. Finally, expanding the term containing the radical in Eq. (41) in a Maclaurin series in $\hat{D}_3$ truncated at the linear term, and rewriting Eqs. (40)–(42) neglecting terms involving nonunity powers or products of $\hat{S}_3$ and $\hat{D}_3$, results in the simplified set

$$U_e^{(1)} = \frac{1}{s_{33}^D}\int_0^{\hat{S}_F} \hat{S}_3\, d\hat{S}_3, \tag{46}$$

$$U_d^{(1)} = \left[\frac{(g_{33}^{\text{effective}})^2}{s_{33}^D} + \beta_{33}^{T(\text{effective})}\right]\int_0^{\hat{D}_F} \hat{D}_3\, d\hat{D}_3, \tag{47}$$

and

$$2U_m^{(1)} = -\frac{g_{33}^{\text{effective}}}{s_{33}^D}\int_0^{\hat{S}_F} \hat{D}_3\, d\hat{S}_3 - \frac{g_{33}^{\text{effective}}}{s_{33}^D}\int_0^{\hat{D}_F} \hat{S}_3\, d\hat{D}_3. \tag{48}$$

Here the ''1'' superscript on the energy variables is used to denote the fact that only the dominant first harmonic contributions to the energy, resulting from the calculations outlined previously, have been retained. Also, the notations

$$\beta_{33}^{T(\text{effective})} \equiv \frac{(1+aE_{\text{ref}}^2)^{3/2}}{\epsilon_0 \epsilon_{33}^T} \tag{49}$$

[which is suggestive of an ''effective'' dielectric impermeability β at bias; cf. Eq. (25)] and

$$g_{33}^{\text{effective}} \equiv 2Q_{33}P_0 + 2Q_{33}D_{\text{ref}} = g_{33} + 2Q_{33}D_{\text{ref}} \tag{50}$$

[which is suggestive of an effective piezoelectric g constant at bias; cf. Eq. (45)] have been introduced. [Equation (49) for the effective impermeability at bias is also easily deduced directly from the theory of electrostriction.]

Equation (50) shows that a biased piezoelectric can be considered to have an enhanced g constant. (That is, the effective g_{33} constant at bias exceeds that of the unbiased ceramic, owing to the presence of the D_{ref} term.) Such enhancement has been observed experimentally (cf. Ref. 18, Fig. 19), at least in the case of g_{31}. The expression for the effective g constant at bias given by Eq. (50) can be further simplified by neglecting the contribution to D_{ref} from E_{drive}, which is assumed small when only first-harmonic contributions are of interest and prestress is negligible [cf. Eq. (30) and the simplified expression for D_{ref} at zero prestress given following Eq. (32)]. This gives

$$g_{33}^{\text{effective}} \approx g_{33} + 2Q_{33}\frac{\epsilon_0 \epsilon_{33}^T E_{\text{bias}}}{\sqrt{1+aE_{\text{bias}}^2}}. \tag{51}$$

[Equation (49) can be simplified in a similar way.] Equation (51) predicts that g_{33} will approach a limiting value, as $E_{\text{bias}} \to \infty$, of $g_{33}^{(\infty)} \to g_{33} + 2Q_{33}\epsilon_0\epsilon_{33}^T/\sqrt{a}$. Such high-bias limiting behavior is clearly consistent with the experimental result shown in Ref. 18, Fig. 19.

Note that the given transformations have rendered Eqs. (46)–(48) similar to Eqs. (16)–(18) of the piezoelectric case. One consequence of this similarity is that the coupling coefficient of Eq. (23) still applies to the biased piezoelectric, with

$$k_{33}^{\text{biased piezoelectric}} = \frac{g_{33}^{\text{effective}}}{\sqrt{(g_{33}^{\text{effective}})^2 + \beta_{33}^{T(\text{effective})} s_{33}^D}}, \tag{52}$$

where $\beta_{33}^{T(\text{effective})}$ and $g_{33}^{\text{effective}}$ are given by Eqs. (49) and (50), respectively. [In general, the compliance term s_{33}^D in Eq. (52) will also vary with bias, although that variation is usually much less in piezoelectric ceramics, such as lead zirconate titanate (PZT), than in electrostrictive ceramics, such as lead magnesium niobate (PMN). However, the theory of Refs. 7 and 8 treats s_{33}^D as a measurable, and does not explicitly predict its value as a function of bias. Nonetheless, it should be understood that the value of s_{33}^D *at bias* should be used in Eq. (52).]

Equations (49), (51), and (52) show that a bias field applied to a piezoelectric ceramic will modify the coupling coefficient. It is worthwhile noting that numerical tests show that the effective coupling coefficient of Eq. (52) has a maximum, when this quantity is viewed as a function of bias. Thus, these tests suggest that a biased piezoelectric ceramic can exhibit an enhanced coupling coefficient relative to that of an unbiased piezoelectric. Such an enhancement has been observed experimentally in barium titanate,[20] cf. Ref. 18, Figs. 16 and 17. However, the amount of enhancement obviously depends on the specific properties of the material, and the required quantities are not generally available. Calculation of the optimum bias requires a knowledge of s_{33}^D, g_{33}, ϵ_{33}^T, Q_{33}, and a. One does not generally know all five of these quantities. However, Eqs. (49) and (51) suggest they might be determined by measuring the effective impermeability [or the reciprocal of the permittivity; cf. Eq. (25)] and g constant as functions of the bias applied to a piezoelectric ceramic. Least-squares fitting of these equations to the measurements would yield g_{33}, ϵ_{33}^T, Q_{33}, and a. A separate measurement of s_{33}^D, again at bias, would also be required.

The fact that $g_{33}^{\text{effective}}$ approaches a finite limiting value as $E_{\text{bias}} \to \infty$, together with the fact that the effective impermeability increases monotonically with increasing bias, as can be seen from Eq. (49), explains why there is a peak value in $k_{33}^{\text{biased piezoelectric}}$, at least if any variation of s_{33}^D with bias is ignored [cf. Eq. (52)]. The predicted increase of the impermeability with bias is consistent with the decrease in permittivity noted in Ref. 18, p. 710.

B. Unbiased electrostrictive ceramic

One goal of the present work was to derive an electrostrictive coupling coefficient that vanishes at zero bias (assuming no significant remanent pole is present), reflecting the behavior of a depoled piezoelectric ceramic. In the present section, we verify that the coupling coefficient defined here exhibits the desired property.

To verify the desired behavior it is sufficient to show that the mutual energy vanishes [cf. Eq. (43)]. Assuming $P_0=0$; imposing the condition of Eq. (12); and carrying out the integrals of Eq. (42), gives

$$2U_m=\frac{\gamma Q_{33}}{s_{33}^D}\hat{D}_F^2(\hat{D}_F+2D_{\text{ref}}). \tag{53}$$

[Although Eq. (53) includes a coefficient of γ, and although Eq. (43) involves a limit in which γ approaches zero, this does not establish the desired vanishing of k_{33}, because a similar factor of γ also appears within the terms comprised by the denominator of Eq. (43). Thus, the zeroing of U_m for zero bias must be established by other means.]

Recall that the quantity $\hat{D}_F$ appearing in Eq. (53) denotes the value of $\hat{D}_3$ when the applied electric field reaches its maximum value. This maximum is reached at zero bias when $E_3=E_{\text{drive}}$, or $\hat{E}_3=2E_{\text{drive}}$. Evaluating $\hat{D}_3$ under these conditions [cf. Eq. (39)]; imposing $E_{\text{bias}}=0$ and $P_0=0$; and applying Eq. (12) gives

$$\begin{aligned}E_{\text{drive}}=&\frac{\hat{D}_F+D_{\text{ref}}}{\sqrt{(\epsilon_0\epsilon_{33}^T)^2-a(\hat{D}_F+D_{\text{ref}})^2}}+\frac{2Q_{33}}{s_{33}^D}\gamma\hat{D}_F^2\\&+\frac{2Q_{33}^2}{s_{33}^D}\hat{D}_F^3+\frac{6Q_{33}^2}{s_{33}^D}D_{\text{ref}}\hat{D}_F^2+2\frac{Q_{33}}{s_{33}^D}\gamma D_{\text{ref}}\hat{D}_F\\&+\frac{4Q_{33}^2}{s_{33}^D}D_{\text{ref}}^2\hat{D}_F-2Q_{33}T_0\hat{D}_F-2Q_{33}T_0D_{\text{ref}}.\end{aligned} \tag{54}$$

In view of Eq. (32), evaluated with $E_{\text{bias}}=0$, a root of Eq. (54) in the limit of $\gamma\to 0$ is seen to be

$$\hat{D}_F=-2D_{\text{ref}}\quad(\text{zero bias}), \tag{55}$$

as can be verified by direct substitution. Substituting $\hat{D}_F=-2D_{\text{ref}}$ as given by Eq. (55) into Eq. (53) establishes the fact that U_m vanishes identically for the conditions of zero bias and remanence. Equation (55) can also be substituted directly into Eq. (44) together with $P_0=0$ to show that k_{33} vanishes identically for zero bias and remanence, as desired.

C. Biased electrostrictive ceramic

We now consider Eq. (44), the coupling coefficient for the general case. In what follows we consider some typical numerical values for the parameters of the theory, when applied to PMN. In particular, it is assumed that $Q_{33}=0.0146\,\text{m}^4/\text{C}^2$, $s_{33}^D=1.1\times10^{-11}\,\text{m}^2/\text{N}$, $\epsilon_{33}^T=28\,737$ (dimensionless), $a=7.386\times10^{-13}\,\text{m}^2/\text{V}^2$, $T_0=0$, and $P_0=0$. (These parameters were determined in a least-squares fit of measurements made on a sample of interest, except for s_{33}^D and P_0. The values used for those quantities are characteristic of PMN, however.) In view of the requirement that the expression for the coupling coefficient be evaluated as γ vanishes [cf. Eq. (43)], a value of $\gamma=1.0\times10^{-20}\,\text{m}^2/\text{C}$ was used for that parameter in the equations as they appear prior to that limit being imposed. However, it was found that the result obtained for the coupling coefficient was not significantly affected by the chosen value of γ unless a value in excess of about $1.0\times10^{-2}\,\text{m}^2/\text{C}$ was used.

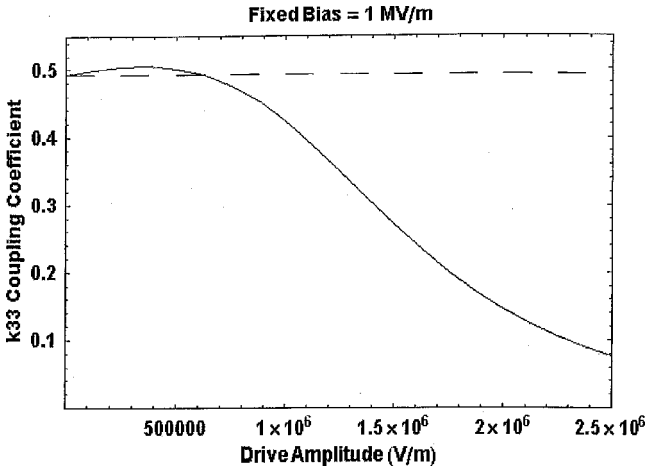


FIG. 1. The k_{33} coupling coefficient, at a fixed bias, plotted as a function of the amplitude of the alternating, or driving, electric field (solid line). The fixed bias field is 1 MV/m, a value that is close to the ''optimum'' determined by Eq. (57), which is approximately 0.95 MV/m. The horizontal dashed line represents the numerical value of the coupling coefficient computed from Eq. (56), a formula based on an equivalent circuit for an electrostrictive transducer (Ref. 8). [The numerical value produced by Eq. (56) for the case considered is $k_{33}=0.4926$.] As expected, the equivalent-circuit coupling coefficient is seen to be most accurate for lower-level drives, but gives an approximate value (accurate to within a few percent of the exact value) for drives up to 75% of the bias.

In Ref. 8 a coupling coefficient for PMN, based on the topology of an equivalent circuit, was given. It is therefore interesting to see how numerical results produced by that equation compare with those produced by Eq. (44). The equivalent-circuit coupling coefficient formula, expressed in terms of the standard notations, is

$$k_{33}^2=\frac{4Q_{33}^2(\epsilon_0\epsilon_{33}^T)^2E_{\text{bias}}^2}{s_{33}^D(1+aE_{\text{bias}}^2)^{5/2}+4Q_{33}^2(\epsilon_0\epsilon_{33}^T)^3E_{\text{bias}}^2}. \tag{56}$$

It is easy to see that this expression vanishes when $E_{\text{bias}}=0$, in conformity with the results given in the preceding section. In numerical tests comparing Eq. (44) with Eq. (56), in which E_{bias} was varied from 1.0×10^5 to 2.5×10^6 V/m, the results were found to agree to four significant figures. In these tests, E_{drive} was taken to be 100 V/m, a small fraction of the bias, in order that harmonic generation would remain insignificant. [The reader who wishes to see a plot of the coupling coefficient produced by Eq. (56) as a function of bias should see Ref. 8, Fig. 2.] These tests demonstrate the consistency of the present definition with Eq. (56).

Next, tests were undertaken in which E_{drive} was allowed to vary substantially. The results of one such test are depicted graphically in Fig. 1. In the calculations shown here, a fixed bias of 1.0 MV/m was used and Eq. (44) was evaluated for values of E_{drive} ranging from 0 to 2.5 MV/m. The result is plotted as a solid-line curve in Fig. 1. The value obtained from Eq. (56) is plotted as a dashed-line curve. As can be seen, the two curves give identical values at the zero-drive level (the left end of the graph).

It is worthwhile noting that the value derived from the expression for the coupling coefficient based on the equivalent circuit remains accurate to fairly high drive amplitudes, as can be seen by examining Fig. 1. For a drive amplitude equal to 75% of the bias, the value of the coupling coefficient based on the circuit is only in error by about 3.6% with respect to that based on the exact expression. However, it can be seen in Fig. 1 that for drives much above 75% of the bias, the circuit-based value becomes increasingly inaccurate, owing to the effects of harmonic generation, and is in error by almost 16.5% when the drive amplitude equals the bias. However, under such conditions harmonic distortion would be rather significant, so that drive levels are unlikely to be allowed to become this great in real transducer operation. Thus, the coupling coefficient based on the equivalent circuit is probably sufficiently accurate for practical applications. This means that the expression provided in Ref. 8 for the ''optimal'' bias level (i.e., the bias level that maximizes the coupling coefficient) can be considered to be accurate for applications in which the drive does not exceed 75% of the bias. The formula for the optimal bias is[8]

$$E_{\text{bias}}^{\text{optimal}} = \sqrt{\frac{2}{3a}}, \tag{57}$$

where a denotes the saturation parameter.[7] For the numerical value of the saturation parameter used in producing Fig. 1, Eq. (57) gives $E_{\text{bias}}^{\text{optimal}} \approx 0.95\,\text{MV/m}$, which is close to the 1 MV/m bias used in producing the results shown.

A careful examination of Fig. 1 also reveals another benefit of the manner in which the coupling coefficient has been defined here. In particular, it is clear from Fig. 1 that as the drive level is increased beyond the bias, the coupling coefficient declines significantly. (This is consistent with results reported elsewhere.[21,22]) Of course when the drive exceeds the bias, harmonic contamination of the output becomes increasingly important. Since the coupling coefficient declines sharply as harmonic contamination increases, it is seen that a natural metric for discriminating against such undesirable behavior has been realized. That is, if maximizing the coupling coefficient is a transducer design goal, achieving that maximum value based on the present definition automatically results in a design that discriminates against harmonic distortion. This is another distinction between the behavior of the definition given here and that given in Ref. 2. (The coupling coefficient of Ref. 2 approaches a nonzero limiting value as $E_{\text{drive}} \to \infty$, with no diminution owing to the effects of harmonic contamination. Cf. Ref. 2, Figs. 8 and 9.)

Next, we take up once again the issue of the dependence of k upon the parameter γ. Also, since the limit in Eq. (43) seems to require vanishingly small strains, as can be seen by referring back to Eq. (12), we will in addition consider the issue of the suitability of the present definition for large strains. Ignoring for the moment the external stress required to enforce Eq. (12), we can assume zero applied stress (including zero prestress, i.e., $T_0 = 0$) to estimate the strain produced by a large electric field. Under these circumstances $D_3 = \epsilon_0 \epsilon_{33}^T E_3 / \sqrt{1 + a E_3^2}$, which is the solution of Eq. (28) when $T_3 = P_0 = 0$. Assuming $E_3 = 2\,\text{MV/m}$, we get $D_3 \approx 0.26\,\text{C/m}^2$ for the numerical parameters introduced at the start of the present section. The strain that results from this D field can be computed from the product $Q_{33} D_3^2$ [cf. Eq. (29)], giving $S_3 \approx 0.001$. As noted previously, the value of k_{33} was found in the present numerical tests not to be significantly affected unless values in excess of $\gamma \approx 1.0 \times 10^{-2}\,\text{m}^2/\text{C}$ were considered. To see what strain such a value of γ implies we can use Eq. (12), but to do so also requires an estimate of $\hat{D}_3$. If we suppose that $E_{\text{bias}} = E_{\text{drive}} = 1\,\text{MV/m}$ produces the 2-MV/m electric field ($= E_{\text{bias}} + E_{\text{drive}}$ at the peak) currently under consideration, Eqs. (30), (32), (33), and (34) show that $\hat{D}_3 = D_3$ and $\hat{S}_3 = S_3$, again under the assumption that remanence and prestress are negligible. Although the D field declines somewhat with applied stress, typically even very large applied stresses reduce it by about 20% or less, so this effect is comparatively small. Thus, the estimate for D_3 of $0.26\,\text{C/m}^2$ just obtained may be considered a reasonable estimate for $\hat{D}_3$ and substituted into Eq. (12), together with $\gamma = 0.01\,\text{m}^2/\text{C}$, to get an estimated strain of $\hat{S}_3 = S_3 \approx 0.0026$. Note that this strain is more than $2\frac{1}{2}$ times the 0.001 strain value just shown to be produced by the electric field alone. [This significantly increased strain is a consequence of the external stress required to enforce Eq. (12).] Since this large strain was found in the present numerical tests to yield almost the same value of k as an infinitesimal strain, it is seen that the limit appearing in Eq. (43) imposes no significant restriction upon the value of strain for which the present definition is applicable. In fact there is clearly sufficient latitude available that instead of using the chosen limit one could use a numerical value of γ which, when substituted into Eq. (12), would produce *the actual maximum strain in a problem of interest*, without significantly affecting the value of the coupling coefficient. These calculations thus support the notion that the dependence of k on γ is exceedingly weak, and that the present definition of k is not limited to cases in which the strain is small.

It may seem surprising that the computed value of k is essentially the same whether the external stress used to effect the present definition is vanishingly small, or is large enough to cause a strain that is more than twice that produced by the electric field alone. However this may be understood on mathematical grounds by a brief reconsideration of Eq. (43), which the reader will recall constitutes the present definition of the coupling coefficient. The γ parameter affects the numerator and denominator of this expression in almost equal proportion. [Note in the linear case, it affects them in *exactly* equal proportion. Cf. Eqs. (19)–(22).] For the nonlinear case, the cancellation of γ between the numerator and denominator of Eq. (43) (for values of γ up to the specified amount) is not exact, but is nearly so.

V. DISCUSSION

The results obtained here differ substantially from those given in Ref. 2. The reader should not conclude from this that the earlier results are erroneous. The differences derive from the fact that there is a certain amount of latitude available in the definition of the coupling coefficient, and this is especially so when the material under consideration is non-

linear. Hence one can choose a definition that maximizes its relevance to the problem of interest, at least for the case of nonlinear materials. To be useful, of course, the chosen definition must be consistent with previous linear results. In the present article a generalized definition that recovers the coupling coefficient for linear systems, while still exhibiting the desired properties, has been given. These desired properties are (1) that the coupling coefficient vanishes identically for an unbiased electrostrictive ceramic and, (2) that the coupling coefficient declines in value as, owing to the material's nonlinear response, more of the input electrical energy is converted into undesirable output harmonics.

The approach used in Ref. 2 treats *any* conversion of electrical energy into mechanical energy on an equal footing, even if much of the mechanical output energy appears in harmonics of the drive frequency. That is, the area calculations used in that definition are insensitive to the process of harmonic generation. This explains why the approach of Ref. 2 produces large nonzero coupling coefficients for an unbiased electrostrictive ceramic at high drive-field levels, a situation in which harmonic distortion would generally be enormous.

Although the integrals used in the present approach obviously can also be interpreted as areas, the redefinition of variables used here results in a strong discrimination against energy injected into harmonics other than the first. These redefinitions tend to create compensating *negative* contributions when E_{drive} exceeds E_{bias}, a circumstance in which significant harmonic generation is expected. The identically zero coupling coefficient obtained here for the zero bias case results from a precise cancellation of negative and positive contributions. This cancellation can be seen by examining Eqs. (30), (32), and (55). A zero bias produces a negative E_{ref} for any nonzero E_{drive}. The negative value of E_{ref} produces a negative D_{ref}, at least when prestress is negligible. [Note the simplified expression for D_{ref} given in the last sentence of the paragraph containing Eq. (32).] The negative D_{ref} (actually 2 D_{ref}) precisely cancels a positive $\hat{D}_F$ via Eq. (55).

On the other hand when $E_{\text{drive}} > E_{\text{bias}}$ and E_{bias} is nonzero, the cancellation between positive and negative contributions is imperfect, owing to an asymmetry created by the nonzero E_{bias}. But in the limit as $E_{\text{drive}} \to \infty$ (at a fixed bias) this cancellation again becomes exact, which explains the decline in coupling coefficient seen at the right-hand side of Fig. 1. Of course as $E_{\text{drive}} \to \infty$, *any* fixed bias behaves effectively as if it were zero compared with the value of E_{drive}, which is increasing without bound. Thus, the diminution in the value of k seen at the right-hand side of Fig. 1 is not only consistent with the zero value obtained at zero bias, but is actually a consequence of it. And this behavior differs substantially from that seen in Figs. 8 and 9 of Ref. 2.

If any appearance of mechanical energy is considered to be useful in an application of interest regardless of the presence of harmonics, the definition of the coupling coefficient proposed in Ref. 2 is satisfactory, and indeed is probably preferable to the present definition. However in applications where first-harmonic fidelity is an important goal, the present definition should be favored.

VI. SUMMARY AND CONCLUSION

A new definition of the coupling coefficient for electrostrictive materials has been proposed. The new definition was shown to produce results consistent with standard piezoelectric theory when the equations produced by the electrostrictive theory are linearized appropriately. The definition also leads naturally to a generalized coupling coefficient suitable for biased piezoelectric ceramics. An enhancement of the coupling coefficient over that obtained for unbiased piezoelectric ceramics is predicted to occur, although the amount of the expected enhancement is not known owing to the absence of suitable experimental data. An experiment was suggested for determining the required parameters.

In analogy with the vanishing of the coupling coefficient for the piezoelectric case that occurs when the material becomes depoled, the new coupling coefficient becomes zero when no bias is applied to an electrostrictive ceramic (and remains zero under zero bias even when a significant prestress is present). Good agreement is seen between the generalized coupling coefficient and that deduced from an equivalent circuit based on a linearized version of the theory for drives as high as 75% of the bias. The present definition also discriminates against transducer designs that produce significant harmonic contamination.

ACKNOWLEDGMENT

The Office of Naval Research, Code 321, supported this work.

[1] D. Stansfield, *Underwater Electroacoustic Transducers* (Bath University Press, Melksham, Wiltshire, Great Britain, 1991), p. 49.
[2] C. L. Hom, S. M. Pilgrim, N. Shankar, K. Bridger, M. Massuda, and R. Winzer, ''Calculation of quasi-static electromechanical coupling coefficients for electrostrictive ceramic materials,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **41**, 542–551 (1994).
[3] A. H. Meitzler (chairman), *IEEE Standard on Piezoelectricity* (IEEE, New York, 1988), pp. 38–40.
[4] See Ref. 3, p. 38, Eq. (125).
[5] D. A. Berlincourt, D. R. Curren, and H. Jaffee, ''Piezoelectric and piezomagnetic materials and their function in transducers,'' in *Physical Acoustics*, edited by W. P. Mason, editor (Academic, New York, 1964), Vol. I A, Eq. (30), p. 190.
[6] W. P. Mason, *Piezoelectric Crystals and Their Application to Ultrasonics* (Van Nostrand, New York, 1950), p. 34, Table IV.
[7] J. C. Piquette and S. E. Forsythe, ''A nonlinear model of lead magnesium niobate (PMN),'' J. Acoust. Soc. Am. **101**, 289–296 (1997).
[8] J. C. Piquette and S. E. Forsythe, ''Generalized material model for lead magnesium niobate (PMN) and an associated electromechanical equivalent circuit,'' J. Acoust. Soc. Am. **104**, 2763–2772 (1998).
[9] M. B. Moffett, M. D. Jevenager, S. S. Gilardi, and J. M. Powers, ''Biased lead zirconate titanate as a high-power transduction material,'' J. Acoust. Soc. Am. **105**, 2248–2251 (1999).
[10] The end-electroded length expander bar for the case of linear piezoelectricity is treated in Ref. 5, pp. 227–228.
[11] The end-electroded length expander bar for the case of electrostriction is treated in Ref. 8, p. 2765.
[12] In Ref. 5, footnote 3a, p. 269 it is pointed out that Eq. (30) of Ref. 5 (for the coupling coefficient) does not apply to piezoelectric equations that are expressed in the form of Eqs. (8) and (9), as given here. However, it will be seen that the standard result for the piezoelectric coupling coefficient, expressed by the present Eq. (27), is nonetheless recovered. This is a consequence of the fact that the formula of Ref. 5 actually has been applied to the present Eqs. (13) and (14), and not to Eqs. (8) and (9). Equations (13) and (14) have the same structure as equation set (23) of Ref. 5, p. 188. That is, the variables T and E are expressed as functions of the variables S and D in both sets. Footnote 3a of Ref. 5 points out that the

coupling coefficient equation is indeed applicable to such equation sets.

[13] Equation (24) is Eq. (A3) of Ref. 8, and can be derived using Eq. (27) of Ref. 5.

[14] Equation (25) follows from a formula on p. 189 of Ref. 5, using the notation for the dielectric permittivity introduced in Ref. 7.

[15] Equation (26) can be derived by combining the final non-numbered equation given on p. 228 of Ref. 5, together with Eq. (35) of that reference, and Eq. (24) of the present work.

[16] See Ref. 5, Eq. (35), or Ref. 3, Eq. (125). The equation is expressed using the notation of Refs. 7 and 8.

[17] See Ref. 6, pp. 301–303.

[18] H. G. Baerwald and D. A. Berlincourt, ''Electromechanical response and dielectric loss of prepolarized barium titanate under maintained electric bias. I,'' J. Acoust. Soc. Am. **25**, 703–710 (1953).

[19] T. F. Hueter, D. P. Neuhaus, and J. Kolb, ''An experimental study of polarization effects in barium titanate ceramics,'' J. Acoust. Soc. Am. **26**, 696–703 (1954).

[20] It is important to note that the enhanced coupling coefficients reported in Ref. 18 were measured under *low-amplitude* drive conditions (cf. Ref. 18, p. 708, last paragraph), and *not* under the drive condition in which the amplitude is equal to the bias, a case that was also extensively considered there. The current Eq. (52) would predict no enhancement for this latter case, since $D_{\mathrm{ref}}=0$. However, Eq. (52) would not be applicable to such a high-amplitude drive condition; Eq. (44) would be used instead.

[21] H. C. Robinson, ''A comparison of nonlinear models for electrostrictive materials,'' Presentation to the 1999 IEEE Ultrasonics Symposium, 17–20 October 1999, Lake Tahoe, NV. Also, H. C. Robinson (private communication, 2000).

[22] This figure is also similar to the variation with bias of the coupling coefficient shown in C. L. Hom and N. Shankar, ''Modeling resonance tests for electrostrictive ceramics,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **46**, 1422–1430 (1999), Figs. 5 and 6. Note, however, those results are for *low*-amplitude drives. For *high*-amplitude drives, the coupling coefficient is shown to saturate to a high nonzero value in Ref. 2, Figures 3, 8, and 9, in contrast to the behavior seen in Fig. 1 here, where the high-amplitude coupling coefficient is seen to vanish in the limit of infinite amplitude.

Analysis of an asymmetrical piezoelectric annular bimorph using impedance and admittance matrices

Sung K. Ha[a] and Young H. Kim
Smart Structures and Materials Laboratory, Department of Mechanical Engineering, Hanyang University, 1271 Sal-Dong, Ansan, Kyounggi-do, Korea 425-791

(Received 13 November 2000; accepted for publication 3 April 2001)

A theoretical model is presented for the analysis of an asymmetrical piezoelectric annular bimorph (APAB) in dynamic harmonic motion. The APAB is assumed to be so thin that the thickness dependent stresses vanish throughout. The conjugate parameters of the admittance and impedance matrices are derived using the variational principle. Both the extensional and flexural motions are considered in deriving the motional equations and the boundary conditions. By using the symmetry characteristic of the admittance matrix, the derivation procedure for the impedance and admittance matrices is thus greatly simplified. The resonance and antiresonance frequencies and the effective electromechanical coupling factors are calculated using the matrices. These present methods are applied to three special cases: a single piezoelectric ring in extensional motion, a triple-layer APAB in series or parallel connection, and a triple-layer APAB with sensor and actuator. The results are compared with previous publication and finite element methods, and it is found that the present methods are very effective in analyzing the piezoelectric multilayer annular transducers. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1375139]

PACS numbers: 43.38.Ar, 43.38.Fx, 43.40.At, 43.40.Dx [SLE]

I. INTRODUCTION

Circular or annular types of multilayer piezoelectric transducers are often used in various fields, such as buzzers, telephone receivers, loud speakers,[1,2] ink jet printer heads, and so on. More recently, annular flextensional transducers that excite axisymmetric resonant modes are designed and applied to small particle ejectors for the deposition of liquids or solid particles.[3–5]

Impedance or admittance matrices are often used in the analysis of piezoelectric transducers since the lumped parameters conveniently describe the electromechanical behavior and show the equivalent circuit of the transducers. The resonance frequencies (RF) and the antiresonance frequencies (AF) and the effective electromechanical coupling factors (EECF) can be effectively calculated using the admittance matrix. Especially in the design of high power devices such as ultrasonic motors and transformers, it is important for the device to be operated at the antiresonance state because it can provide the same mechanical vibration level with less heat generation and with higher *Q*-factor than the resonance state.[6]

Numerous investigations have been thus made on the analysis and design of beam type piezoelectric transducers. Tanaka *et al.*[7,8] formulated basic equations for the multilayer piezoelectric beams using Hamilton's principle and Aoyagi *et al.*[9] presented a block equivalent circuit of multilayer piezoelectric benders. Wang *et al.*[10] discussed electromechanical coupling and output efficiency for cantilever bimorph and unimorph actuators. Smits *et al.*[11–13] derived the dynamic admittance matrix and calculated the electrical and mechanical AF together with the effective coupling factors of the cantilevered piezoelectric flexors. Cho *et al.*[14] presented a five-port equivalent electric circuit of piezoelectric bimorphs. Dobrucki *et al.*[15] presented finite element methods (FEM) for an axisymmetric multilayered piezoelectric shell, and Hui *et al.*[16] derived an exact solution for the axisymmetric response of piezoelectric circular plates. Lin *et al.*[17] present a model to analyze the coupling behavior between longitudinal and radial vibration of the disk resonator with the radius and thickness having the same dimensional magnitude. Iula *et al.*[18–20] presented a model for the theoretical characterization of thin piezoelectric rings in radial extensional motion and showed the variation of the effective coupling factors with the inner radius. Ha *et al.*[21] derived impedance and admittance matrices of a symmetrical piezoelectric annular bimorph (PAB) in series connection and analyzed PAB with the piezoelectric layers partially covering the shim layer using the matrices.

Nevertheless, an asymmetrical piezoelectric annular bimorph (APAB) has not been investigated until now. In this paper, the differential equations for the extensional and flexural motions of the triple-layer APAB are decoupled and the mechanical and electrical boundary conditions are derived using the variational principle. A systematic procedure of deriving an impedance matrix using the symmetry characteristic of the matrix is then presented and it is shown that the derivation procedure can be greatly simplified. Using the derived impedance matrix, the characteristic equations are then presented, from which the resonance and antiresonance frequencies and eventually the EECF can be calculated. A single piezoelectric ring in extensional motion, a triple-layer APAB in either series or parallel connections, and a triple-layer APAB with sensor and actuator are analyzed using the present methods, and the results are then compared with

[a]Author to whom correspondence should be addressed. Electronic mail: sungha@hanyang.ac.kr

 0001-4966/2001/110(1)/208/8/$18.00

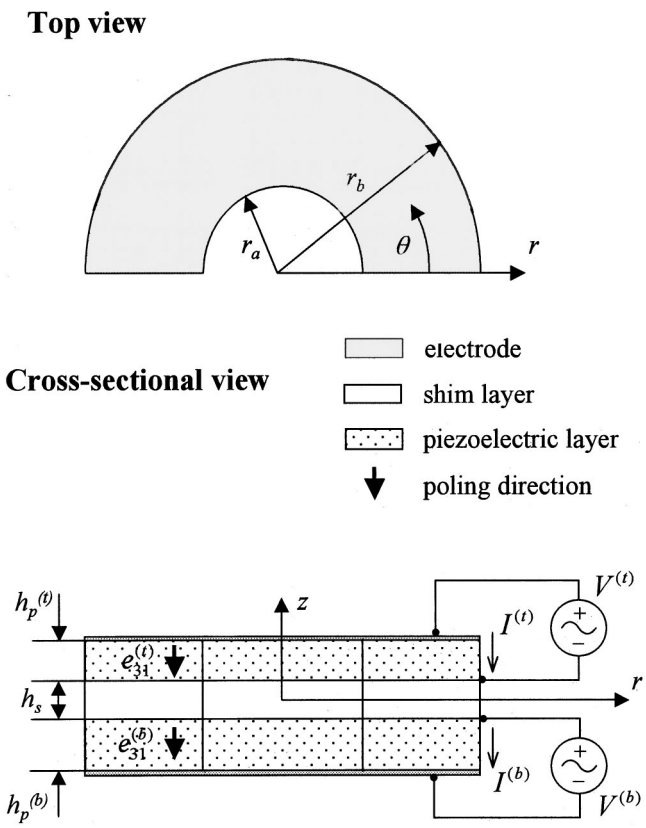


FIG. 1. Configuration of the triple-layer asymmetrical piezoelectric annular bimorph (APAB) with two electrical ports.

other publications and the finite element analysis. It is shown that the present methods are very effective in analyzing the piezoelectric annular type of transducers.

II. DIFFERENTIAL EQUATIONS AND BOUNDARY CONDITIONS

A triple-layer APAB for transducers is considered in this study, and the top and cross-sectional view in the cylindrical coordinate system r, θ, and z is shown in Fig. 1. The inner radius, the outer radius, the thickness of top and bottom piezoelectric layers, and the thickness of the middle shim are denoted by r_a, r_b, $h_p^{(t)}$, $h_p^{(b)}$, and h_s, respectively. The APAB is assumed to be thin and the top and bottom surfaces are supposed to be stress free. The z-direction-dependent stresses thus vanish throughout. Each piezoelectric layer is polarized in the direction perpendicular to the plate, i.e., parallel to the z-axis, and metallized or electroded on the top and bottom faces of the piezoelectric layer so that the r and θ directional electric fields E_r and E_θ also vanish. Under the assumption of axial symmetry, all the mechanical and electrical components are independent of θ. The two-dimensional constituent equations of the piezoelectric element are written as[6,22]

$$\begin{aligned} T_r &= c_{11}S_r + c_{12}S_\theta - e_{31}E_z, \\ T_\theta &= c_{12}S_r + c_{22}S_\theta - e_{31}E_2, \\ D_z &= e_{31}(S_r + S_\theta) + \varepsilon_{33}E_z. \end{aligned} \tag{1}$$

The material properties in Eq. (1) have the following relations:

$$\begin{aligned} c_{11} &= c_{22} = \frac{1}{s_{11}^E(1-\nu^2)}, \quad c_{12} = \nu c_{11}, \\ e_{31} &= \frac{d_{31}}{s_{11}^E(1-\nu)}, \quad \varepsilon_{33} = \varepsilon_{33}^T - \frac{2d_{31}^2}{s_{11}^E(1-\nu)}. \end{aligned} \tag{2}$$

T_r, T_θ, S_r, and S_θ in Eq. (1) are the radial and circumferential normal stresses and strains; D_z and E_z are the electric displacement and field, respectively; c_{ij} and s_{ij}^E are the elastic stiffness and compliance under the constant electric field; d_{31} and e_{31} are the piezoelectric strain and stress constant; ε_{33}^T and ε_{33} are the dielectric constant under the constant stress and constant strain, respectively; ν denotes a planar Poisson's ratio.

Based on the Kirchhoff assumption and the axial symmetric condition, the displacements are supposed to be

$$u_r(r,z) = u_R(r) - zu_{z,r}, \quad u_\theta(r,z) = 0, \quad u_z(r,z) = u_z(r). \tag{3}$$

u_R is the radial extensional displacement, i.e., the radial displacement u_r at the neutral axis ($z=0$) that will be determined later. The strain-displacement relationship can then be expressed as

$$S_r = u_{r,r} = u_{R,r} - zu_{z,rr}, \quad S_\theta = \frac{u_r}{r} = \frac{u_R}{r} - z\frac{u_{z,r}}{r}. \tag{4}$$

The electric field of each piezoelectric layer is approximated as

$$E_z^{(p)} = \frac{V^{(p)}}{h_p^{(p)}}, \tag{5}$$

where $V^{(p)}$ is the electric potential difference between the top and bottom surfaces of the piezoelectric layer; the superscript (p) denotes either a top layer (t) or a bottom layer (b).

The mechanical and electrical responses are supposed to be in harmonic motion with an angular frequency ω. The internal energy L for the motion of the APAB can then be defined as

$$L = \frac{1}{2}\int\int_v\int \{-\rho\omega^2(u_r^2+u_z^2) + T_rS_r + T_\theta S_\theta - D_zE_z\} \times r\,d\theta\,dz\,dr, \tag{6}$$

where v denotes the total volume of the APAB and ρ is the mass density. Since all the quantities are independent of the θ-direction, the variational principle yields

$$\delta L/2\pi = \int\int \{-\rho\omega^2 r(u_r\delta u_r + u_z\delta u_z) + r(T_r\delta S_r + T_\theta\delta S_\theta) - rD_z\delta E_z\}dz\,dr. \tag{7}$$

In the calculus of variations, u_R, u_z, and $V^{(q)}$ are considered as the independent variables. Using Eq. (3), performing the integration of the first term in Eq. (7) yields

$$\int\int \rho\omega^2 r(u_r\delta u_r+u_z\delta u_z)dz\,dr$$

$$=\int_{r_a}^{r_b}\rho_h\omega^2 r(u_R\delta u_R+u_z\delta u_z)dx, \tag{8}$$

where the rotational inertia is neglected, and ρ_h now indicates the mass density per unit area. Using Eq. (4), the integration by parts of the second term in Eq. (7) yields

$$\int\int r(T_r\delta S_r+T_\theta\delta S_\theta)dz\,dr$$

$$=\int_{r_a}^{r_b}\{-(rN_r)_{,r}+N_\theta\}\delta u_R\,dr$$

$$+\int_{r_a}^{r_b}\{(rM_r)_{,rr}-M_{\theta,r}\}\delta u_z\,dr+rN_r\delta u_R|_{r=r_a}^{r=r_b}$$

$$+rM_r\delta u_{z,r}|_{r=r_a}^{r=r_b}-(rM_r)_{,r}\delta u_z|_{r=r_a}^{r=r_b}+M_\theta\delta u_z|_{r=r_a}^{r=r_b}, \tag{9}$$

where the extensional loads and the flexural moments are defined as

$$N_r=\int T_r\,dz,\quad N_\theta=\int T_\theta\,dz, \tag{10a}$$

$$M_r=-\int T_r z\,dz,\quad M_\theta=-\int T_\theta z\,dz. \tag{10b}$$

Using Eq. (5), the integration of the third term in Eq. (7) yields

$$\int\int rD_z\delta E_z\,dz\,dr=\int\int_{(t)}rD_z^{(t)}\frac{\delta V^{(t)}}{h^{(t)}}dz\,dr$$

$$+\int\int_{(b)}rD_z^{(b)}\frac{\delta V^{(b)}}{h^{(b)}}dz\,dr$$

$$=Q^{(t)}\delta V^{(t)}+Q^{(b)}\delta V^{(b)}, \tag{11}$$

where the charge of the pth layer ($p=t$ or b) is thus defined as

$$Q^{(p)}=\frac{1}{h_p^{(p)}}\int\int_{(p)}rD_2^{(p)}\,dz\,dr. \tag{12}$$

Notice that, through the thickness direction, the charge is represented by an averaged electric displacement rather than by the value on the surface, which results from the common assumption in Eq. (5). Combining Eqs. (8), (9), and (11) finally yields

$$\delta L/2\pi=rN_r\delta u_R|_{r=r_a}^{r=r_b}+rM_r\delta u_{z,r}|_{r=r_a}^{r=r_b}+rR_r\delta u_z|_{r=r_a}^{r=r_b}$$

$$-Q^{(t)}\delta V^{(t)}-Q^{(b)}\delta V^{(b)} \tag{13}$$

and the extensional and flexural equilibrium equations:

$$(rN_r)_{,r}-N_\theta+\rho_h\omega^2 ru_R=0, \tag{14a}$$

$$\{(rM_r)_{,r}-M_\theta\}_{,r}-\rho_h\omega^2 ru_z=0, \tag{14b}$$

where the shear force R_r is defined as

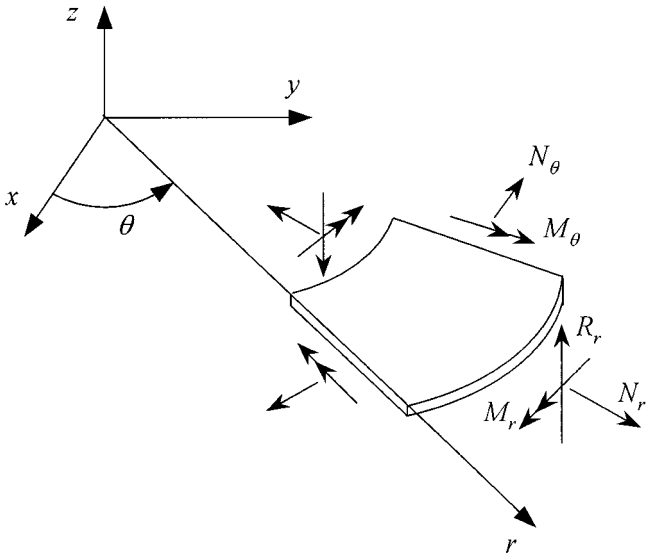


FIG. 2. The extensional loads N_r, N_θ, the flexural moments M_r, M_θ, and the shear force R_r, acting in the positive directions on the edges.

$$rR_r=-(rM_r)_{,r}+M_\theta. \tag{15}$$

The state of the extensional loads, the flexural moments, and the shear force which are defined in Eqs. (10) and (15) are shown in Fig. 2. Using Eqs. (1) and (4), performing the integration of Eq. (12) yields the expression for the electric charge of each piezoelectric layer:

$$Q^{(p)}=\frac{1}{h_p^{(p)}}\int\int_{(p)}e_{31}^{(p)}(S_r+S_\theta)r\,dz\,dr$$

$$+\frac{1}{h_p^{(p)}}\int\int_{(p)}\varepsilon_{33}^{(p)}\frac{V^{(p)}}{h_p^{(p)}}r\,dz\,dr$$

$$=e_{31}^{(p)}ru_R|_{r=r_a}^{r=r_b}-e_{31}^{(p)}z_e^{(p)}ru_{z,r}|_{r=r_a}^{r=r_b}+C^{(p)}V^{(p)}, \tag{16}$$

where the capacitance is defined as

$$C^{(p)}=\frac{(r_b^2-r_a^2)}{2}\frac{\varepsilon_{33}^{(p)}}{h_p^{(p)}}, \tag{17}$$

and $z_c^{(p)}$ denotes the z-directional center coordinate of the pth layer. Notice that all the mechanical forces and the electric charges are divided by 2π, as shown in Eq. (13).

Substituting Eqs. (1), (4), and (5) into Eq. (10) yields the following expressions for the extensional loads and the flexural moments:

$$\begin{pmatrix}N_r\\N_\theta\\M_r\\M_\theta\end{pmatrix}=\begin{bmatrix}A_{11}&A_{12}&B_{11}&B_{12}\\A_{21}&A_{22}&B_{21}&B_{22}\\B_{11}&B_{12}&D_{11}&D_{12}\\B_{21}&B_{22}&D_{21}&D_{22}\end{bmatrix}\begin{pmatrix}u_{R,r}\\u_R/r\\u_{z,rr}\\u_{z,r}/r\end{pmatrix}+\begin{pmatrix}N_r^*\\N_\theta^*\\M_r^*\\M_\theta^*\end{pmatrix}, \tag{18}$$

where the equivalent extensional loads N_r^*, N_θ^* and the equivalent flexural moments M_r^*, M_θ^* that result from the applied voltage are defined as

$$N_r^*=N_\theta^*=-\int e_{31}E_z\,dz=-e_{31}^{(t)}V^{(t)}-e_{31}^{(b)}V^{(b)}, \tag{19a}$$

$$M_r^* = M_\theta^* = \int e_{31} E_z z \, dz = z_c^{(t)} e_{31}^{(t)} V^{(t)} + z_c^{(b)} e_{31}^{(b)} V^{(b)}. \tag{19b}$$

The stiffness matrices A_{ij}, B_{ij}, and D_{ij} in Eq. (18) are defined as

$$(A_{ij}, B_{ij}, D_{ij}) = \int c_{ij}(1, -z, z^2) dz. \tag{20}$$

As can be seen in Eqs. (14) and (18), the differential equations of the extensional and flexural motions can be decoupled by vanishing the matrix B_{ij}. This is made possible by properly choosing the position of $z=0$, i.e., the neutral axis. Let z_0 be the distance from a reference axis, $z'=0$, e.g., the bottom surface of the laminate, to the neutral axis, $z=0$. The distance z_0 is determined so that $B_{11}(=B_{22})$ vanish:

$$B_{11} = B_{22} = -\int c_{11}^E (z' - z_0) dz' = 0, \quad z_0 = \frac{\int c_{11} z' \, dz'}{\int c_{11} \, dz'}. \tag{21}$$

If the Poisson's ratios of all the layers are the same, B_{12} is also zero with z_0 in Eq. (21) and the distance z_0 is uniquely determined. However, in other cases, B_{12} is neglected and the decoupling is performed approximately.

Then substituting Eq. (18) with the decoupled condition into Eq. (14) yields the decoupled extensional and flexural differential equations:

$$\frac{\partial^2 u_R}{\partial r^2} + \frac{1}{r}\frac{\partial u_R}{\partial r} - \frac{u_R}{r^2} + \lambda_N^2 u_R = 0, \tag{22a}$$

$$\frac{\partial^4 u_z}{\partial r^4} + \frac{2}{r}\frac{\partial^3 u_z}{\partial r^3} - \frac{1}{r^2}\frac{\partial^2 u_z}{\partial r^2} + \frac{1}{r^3}\frac{\partial u_z}{\partial r} - \lambda_M^4 u_z = 0, \tag{22b}$$

where the parameters λ_N and λ_M are functions of the angular frequency ω, each defined as

$$\lambda_N^2 = \frac{\rho_h \omega^2}{A_{11}}, \quad \lambda_M^4 = \frac{\rho_h \omega^2}{D_{11}}. \tag{23}$$

Notice that the electric voltage disappeared in Eq. (22) since the derivative of $V^{(p)}$ in Eq. (19) with respect to r vanished; the electric charge or voltage coupled with the displacements appears in the boundary conditions as can be seen in Eqs. (16) and (19).

III. IMPEDANCE AND ADMITTANCE MATRICES

In this section, the impedance and admittance matrices are derived using the results obtained in the previous section. Equation (13) can be now written as

$$\delta L / 2\pi = \begin{pmatrix} \mathbf{F}_N \\ \mathbf{F}_M \\ -\mathbf{Q} \end{pmatrix}^T \delta \begin{pmatrix} \mathbf{u}_N \\ \mathbf{u}_M \\ \mathbf{V} \end{pmatrix}, \tag{24}$$

where each vector is defines as

$$\mathbf{F}_N = \begin{pmatrix} -r_a N_r(r_a) \\ r_b N_r(r_b) \end{pmatrix}, \quad \mathbf{u}_N = \begin{pmatrix} u_R(r_a) \\ u_R(r_b) \end{pmatrix}, \tag{25a}$$

$$\mathbf{F}_M = \begin{pmatrix} -r_a M_r(r_a) \\ -r_a R_r(r_a) \\ r_b M_r(r_b) \\ r_b R_r(r_b) \end{pmatrix}, \quad \mathbf{u}_M = \begin{pmatrix} u_{z,r}(r_a) \\ u_z(r_a) \\ u_{z,r}(r_b) \\ u_z(r_b) \end{pmatrix}, \tag{25b}$$

$$\mathbf{Q} = \begin{pmatrix} Q^{(t)} \\ Q^{(b)} \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} V^{(t)} \\ V^{(b)} \end{pmatrix}. \tag{25c}$$

As often noticed in the electromechanical system, Eq. (24) manifests the reciprocity conditions of the APAB with the multi-electrical ports.

The general solution of Eq. (22) can be written as

$$u_R(r) = \Phi_N^T \mathbf{A}_N, \tag{26a}$$

$$u_z(r) = \Phi_M^T \mathbf{A}_M, \tag{26b}$$

where

$$\Phi_N = (J_1(\lambda_N r) \;\; Y_1(\lambda_N r))^T, \tag{27a}$$

$$\Phi_M = (J_0(\lambda_M r) \;\; Y_0(\lambda_M r) \;\; I_0(\lambda_M r) \;\; K_0(\lambda_M r))^T, \tag{27b}$$

$$\mathbf{A}_N = (A_1 \;\; A_2)^T, \quad \mathbf{A}_M = (A_3 \;\; A_4 \;\; A_5 \;\; A_6)^T, \tag{27c}$$

where J_i and Y_i are the Bessel functions of the first and second kind of order i, respectively. I_i and K_i are modified Bessel functions of the first and second kind of order i, respectively. The coefficient vectors $\mathbf{A}_N, \mathbf{A}_M$ are to be determined using the boundary conditions.

Using Eqs. (18), (19a), and (26a), the extensional force boundary condition of Eq. (25a) can be represented in a matrix form which is the function of displacements at the boundary and the voltages:

$$\mathbf{F}_N = \mathbf{B}_N^F \mathbf{A}_N + \mathbf{C}_N^E \mathbf{V}, \tag{28}$$

where the matrices $\mathbf{B}_N^F$ and $\mathbf{C}_N^E$ are defined as

$$\mathbf{B}_N^F = A_{11} \begin{bmatrix} -r_a \lambda_N J_N^{0a} + J_N^{1a} - \alpha I_N^{1a} & -r_a \lambda_N Y_N^{0a} + Y_N^{1a} - \alpha Y_N^{1a} \\ r_b \lambda_N J_N^{0b} - J_N^{1b} + \alpha J_N^{1b} & r_b \lambda_N Y_N^{0b} - Y_N^{1b} + \alpha Y_N^{1b} \end{bmatrix}, \tag{29a}$$

$$\mathbf{C}_N^E = \begin{bmatrix} r_a e_{31}^{(t)} & r_a e_{31}^{(b)} \\ -r_b e_{31}^{(t)} & -r_b e_{31}^{(b)} \end{bmatrix}, \tag{29b}$$

where J_N^{0a} denotes $J_0(\lambda_N r_a)$, etc., and α is defined as A_{12}/A_{11}. Using Eqs. (18), (19b), and (26b), the flexural moment boundary condition of Eq. (25b) is similarly expressed in a matrix form:

$$\mathbf{F}_M = \mathbf{B}_M^F \mathbf{A}_M + \mathbf{C}_M^E \mathbf{V}, \tag{30}$$

where the matrices $\mathbf{B}_M^F$ and $\mathbf{C}_M^E$ are defined as

$$\mathbf{B}_M^F = D_{11}\lambda_M^2 \begin{bmatrix} r_a J_M^{0a} - \beta J_M^{1a} & r_a Y_M^{0a} - \beta Y_M^{1a} & -r_a I_M^{0a} + \beta I_M^{1a} & -r_a K_M^{0a} - \beta K_M^{1a} \\ \lambda_M r_a J_M^{1a} & \lambda_M r_a Y_M^{1a} & \lambda_M r_a I_M^{1a} & -\lambda_M r_a K_M^{1a} \\ -r_b J_M^{0b} + \beta J_M^{1b} & -r_b Y_M^{0b} + \beta Y_M^{1b} & r_b I_M^{0b} - \beta I_M^{1b} & r_b K_M^{0b} + \beta K_M^{1b} \\ -\lambda_M r_b J_M^{1b} & -\lambda_M r_b Y_M^{1b} & -\lambda_M r_b I_M^{1b} & \lambda_M r_b K_M^{1b} \end{bmatrix}, \tag{31a}$$

$$\mathbf{C}_M^E = \begin{bmatrix} -r_a e_{31}^{(t)} z_c^{(t)} & -r_a e_{31}^{(b)} z_c^{(b)} \\ 0 & 0 \\ r_b e_{31}^{(t)} z_c^{(t)} & r_b e_{31}^{(b)} z_c^{(b)} \\ 0 & 0 \end{bmatrix}, \tag{31b}$$

where β is defined as $(1-D_{12}/D_{11})/\lambda_M$

Similar to the force boundary conditions, the displacement boundary conditions of Eqs. (25a) and (25b) can be also expressed in a matrix form by using Eq. (26):

$$\mathbf{u}_N = \mathbf{B}_N^u \mathbf{A}_N, \tag{32a}$$

$$\mathbf{u}_M = \mathbf{B}_M^u \mathbf{A}_M, \tag{32b}$$

where $\mathbf{B}_N^u$ and $\mathbf{B}_M^u$ are defined as

$$\mathbf{B}_N^u = \begin{bmatrix} J_N^{1a} & Y_N^{1a} \\ J_N^{1b} & Y_N^{1b} \end{bmatrix}, \tag{33a}$$

$$\mathbf{B}_M^u = \begin{bmatrix} -\lambda_M J_M^{1a} & -\lambda_M Y_M^{1a} & \lambda_M I_M^{1a} & -\lambda_M K_M^{1a} \\ J_M^{0a} & Y_M^{0a} & I_M^{0a} & K_M^{0a} \\ -\lambda_M J_M^{1b} & -\lambda_M Y_M^{1b} & \lambda_M I_M^{1b} & -\lambda_M K_M^{1b} \\ J_M^{0b} & Y_M^{0b} & I_M^{0b} & K_M^{0b} \end{bmatrix}. \tag{33b}$$

Notice that the procedure of Eqs. (24)–(33) can be also applied to other types of piezoelectric transducer with the proper redefinitions of $\mathbf{B}^F$, $\mathbf{B}^u$, and $\mathbf{C}^E$ according to the applications.

As can be seen in Eq. (16), the electric charge of each piezoelectric layer has relationships with both the mechanical displacements and the electric voltages. Considering the definitions of $\mathbf{C}_N^E$ in Eq. (29b) and $\mathbf{C}_M^E$ in Eq. (31b), the charge vector of Eq. (25c) can be now expressed as

$$\mathbf{Q} = -(\mathbf{C}_N^E)^T \mathbf{u}_N - (\mathbf{C}_M^E)^T \mathbf{u}_M + \mathbf{C}\mathbf{V}, \tag{34}$$

where the matrix $\mathbf{C}$ is a diagonal matrix with each diagonal term represented by Eq. (17). Notice that Eqs. (28), (30), and (34) show the electromechanical reciprocity conditions.

For the derivation of impedance matrix, velocity vectors $\mathbf{U}_N$ and $\mathbf{U}_M$ and current vector $\mathbf{I}$ are used instead of the displacement vectors $\mathbf{u}_N$ and $\mathbf{u}_M$ and the charge vector $\mathbf{Q}$, respectively. Since we are considering all the physical quantities in the harmonic response, the following relations are used:

$$\mathbf{U}_N = j\omega \mathbf{u}_N, \quad \mathbf{U}_M = j\omega \mathbf{u}_M, \tag{35a}$$

$$\mathbf{I} = j\omega \mathbf{Q}, \tag{35b}$$

where the component $I^{(p)}$ of the current vector $\mathbf{I}$ is the current per unit angle which flows the p-th layer in the thickness direction. Eliminating the coefficient vectors $\mathbf{A}_N$ and $\mathbf{A}_M$ from Eqs. (28), (30), and (32), then using Eqs. (34) and (35), we obtain

$$\begin{pmatrix} \mathbf{F}_N \\ \mathbf{F}_M \\ -\mathbf{I} \end{pmatrix} = \begin{bmatrix} \mathbf{Z}_N & \mathbf{0} & \mathbf{C}_N^E \\ & \mathbf{Z}_M & \mathbf{C}_M^E \\ \text{sym.} & & -\mathbf{Z}_C^{-1} \end{bmatrix} \begin{pmatrix} \mathbf{U}_N \\ \mathbf{U}_M \\ \mathbf{V} \end{pmatrix}, \tag{36}$$

where

$$\mathbf{Z}_N = \frac{1}{j\omega} \mathbf{B}_N^F (\mathbf{B}_N^u)^{-1}, \tag{37a}$$

$$\mathbf{Z}_M = \frac{1}{j\omega} \mathbf{B}_M^F (\mathbf{B}_M^u)^{-1}, \tag{37b}$$

$$\mathbf{Z}_C = \frac{1}{j\omega} \mathbf{C}^{-1}. \tag{37c}$$

Notice that the negative sign in the currents and the symmetry of the matrix are manifested by the results of the variational analysis, i.e., Eq. (24). The exchange of the current vector for the voltage vector in Eq. (36) easily yields the impedance matrix:

$$\begin{pmatrix} \mathbf{F}_N \\ \mathbf{F}_M \\ \mathbf{V} \end{pmatrix} = \begin{bmatrix} \mathbf{Z}_N + \mathbf{C}_N^E \mathbf{Z}_C (\mathbf{C}_N^E)^T & \mathbf{C}_N^E \mathbf{Z}_C (\mathbf{C}_M^E)^T & \mathbf{C}_N^E \mathbf{Z}_C \\ & \mathbf{Z}_M + \mathbf{C}_M^E \mathbf{Z}_C (\mathbf{C}_M^E)^T & \mathbf{C}_M^E \mathbf{Z}_C \\ \text{sym.} & & \mathbf{Z}_C \end{bmatrix} \times \begin{pmatrix} \mathbf{U}_N \\ \mathbf{U}_M \\ \mathbf{I} \end{pmatrix}. \tag{38}$$

The exchange of the force vectors for the velocity vectors in Eq. (36) yields the admittance matrix:

$$\begin{pmatrix} \mathbf{U}_N \\ \mathbf{U}_M \\ \mathbf{I} \end{pmatrix} = \begin{bmatrix} \mathbf{Y}_N & \mathbf{0} & -\mathbf{Y}_N \mathbf{C}_N^E \\ & \mathbf{Y}_M & -\mathbf{Y}_M \mathbf{C}_M^E \\ \text{sym.} & & \mathbf{Y}_E \end{bmatrix} \begin{pmatrix} \mathbf{F}_N \\ \mathbf{F}_M \\ \mathbf{V} \end{pmatrix}, \tag{39}$$

which enables us to calculate the mechanical and electrical responses of the piezoelectric transducers due to harmonic excitation by either forces at the boundary or voltages in each piezoelectric layer. Notice that extensional and flexural motions can be generated by the exciting voltage due to the nonzero matrix $\mathbf{C}_N^E$ and $\mathbf{C}_M^E$. The mechanical and electrical admittance matrices in Eq. (39) is defined as

$$\mathbf{Y}_N = (\mathbf{Z}_N)^{-1}, \quad \mathbf{Y}_M = (\mathbf{Z}_M)^{-1}, \tag{40a}$$

$$\mathbf{Y}_E = (\mathbf{C}_N^E)^T \mathbf{Y}_N \mathbf{C}_N^E + (\mathbf{C}_M^E)^T \mathbf{Y}_M \mathbf{C}_M^E + j\omega \mathbf{C}. \tag{40b}$$

In case of no external load, $\mathbf{Y}_E$ represents the two ports electrical admittance matrix with components of $\mathbf{Y}_E(p,q)$ relat-

ing the p-th current with the q-th voltage: The poles and zeros of $\mathbf{Y}_E$ give the characteristic equations that yield the RF ω_r and the AF ω_a of electric current. The EECF k_{eff} is then calculated using the RF and the AF by applying the known expression[23]

$$k_{\text{eff}}^2 = \frac{\omega_a^2 - \omega_r^2}{\omega_a^2}. \tag{41}$$

The EECF is commonly used as an index of the capability of the piezoelectric transducer to convert mechanical into electrical energy, or vice versa, at each resonance frequency.

It is also important to determine the mechanical antiresonance frequencies for various applications.[13] Using the upper-right term in Eq. (39), the mechanical antiresonance frequencies can be determined from the condition:

$$\mathbf{Y}_N \mathbf{C}_N^E = 0, \quad \mathbf{Y}_M \mathbf{C}_M^E = \mathbf{0}. \tag{42}$$

In the cases of other boundary conditions than all mechanically free conditions considered above, the dimensions of the impedance matrix in Eq. (38) and the admittance matrix in Eq. (39) are reduced according to the specified displacement conditions. The RF, the AF, and the EECF can then be calculated as above.

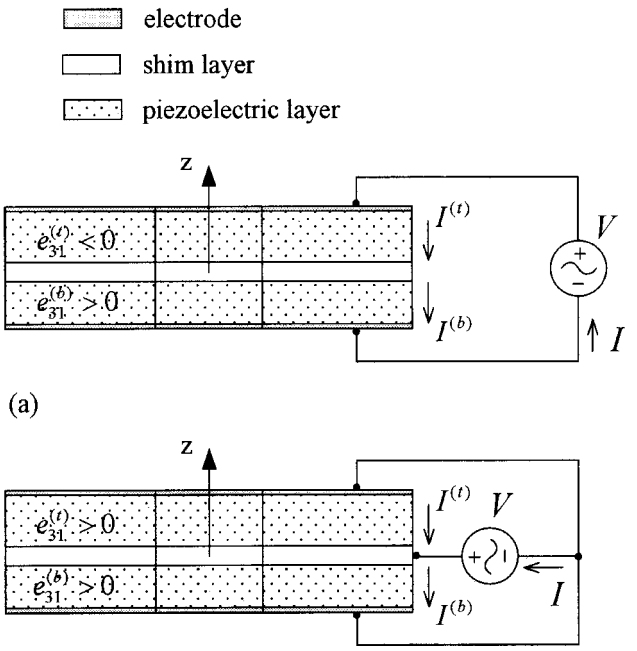


FIG. 3. Two types of the APAB of which two piezoelectric layers are used for actuating. (a) APAB in series connection; (b) APAB in parallel connection.

IV. RESULTS AND DISCUSSION

In this section, several types of annular piezoelectric transducers are investigated using the impedance and admittance matrices derived in the previous section. In the first special case, a single piezoelectric annular ring in radial extensional motion is analyzed.[18–20] In this case, the capacitance C, the extensional force-voltage matrix $\mathbf{C}_N^E$, and the flexural moment-voltage matrix $\mathbf{C}_M^E$ in Eqs. (17), (29b), and (31b) simply reduce to

$$C = \frac{(r_b^2 - r_a^2)}{2} \frac{\varepsilon_{33}}{h_p}, \quad \mathbf{C}_N^E = e_{31} \begin{bmatrix} r_a \\ -r_b \end{bmatrix}, \quad \mathbf{C}_M^E = \mathbf{0}. \tag{43}$$

The electrical admittance of Eq. (40b) also reduces to

$$Y_E = (\mathbf{C}_N^E)^T \mathbf{Y}_N \mathbf{C}_N^E + j\omega C. \tag{44}$$

Then the current I is expressed as follows:

$$I = -(\mathbf{C}_N^E)^T \mathbf{U}_N + j\omega C V. \tag{45}$$

The results are in agreement with the expression presented by Iula *et al.*[18] but are more simplified.

In the second special case, the effects of the thickness ratio of the triple-layer APAB in either parallel or series connection on the RF and the AF are examined. For the series connection, as shown in Fig. 3(a), the two electrical ports are reduced to one electrical port and its voltage and current are expressed as

$$V = V^{(t)} = V^{(b)}, \quad I = I^{(t)} = I^{(b)}. \tag{46}$$

The two electrical ports are also reduced to one electrical port for the parallel connection, as shown in Fig. 3(b):

$$V = -V^{(t)} = V^{(b)}, \quad I = -I^{(t)} + I^{(b)}. \tag{47}$$

In the numerical calculations, the top and bottom layers are PZT G1195N (point group 6 mm) piezoceramics with the following parameters: $1/s_{11}^E = 61.0\,\text{GPa}$, $\rho = 7600.0\,\text{kg/m}^3$, $\nu = 0.3$, $d_{31} = 254.0\,\text{pm/V}$, and $\varepsilon_{33}^T = 15.0\,\text{nF/m}$. The shim layer is made of stainless steel with $1/s_{11}^E = 200.0\,\text{GPa}$, $\rho = 7830.0\,\text{kg/m}^3$, and $\nu = 0.3$. Two cases of boundary conditions are considered, (1) free-free and (2) free-clamped, each at the inner surface $r_a = 20.0\,\text{mm}$ and the outer surface $r_b = 50.0\,\text{mm}$. The total thickness of the APAB and the thickness of the middle shim are set to 1.0 mm and 0.6 mm, respectively. The RF and the AF of the fundamental mode are calculated for the thickness ratios of the top and bottom piezoelectric layers $h_p^{(t)}/h_p^{(b)}$ ranging from 0 to 1, shown in Fig. 4(a) for series connection and 4(b) for parallel connection. Since no investigation on the APAB has been performed before, three-dimensional FEM[24] are used to verify the results by the present methods. The element has four degrees of freedom at a node, i.e., three displacements and one electric potential. Forty elements are used for the radial direction, and nine elements for the thickness direction. In the finite element analysis, the RF and the AF are calculated by the modal analysis under the condition of short circuit and open circuit, respectively.[25] The RF and the AF calculated by the impedance matrix are lower than those by the FEM with the relative error of 1% when the ratio $h_p^{(t)}/h_p^{(b)}$ is equal to 1 and 2% as the ratio $h_p^{(t)}/h_p^{(b)}$ approaches to 0, which is caused by the assumption of constant electric field in each piezoelectric layer. The RF and the AF of the APAB under the free-clamped conditions are higher than those of the APAB under the free-free conditions but they are in a similar trend with respect to the ratio $h_p^{(t)}/h_p^{(b)}$. As the ratio $h_p^{(t)}/h_p^{(b)}$ approaches 0, the RF of the APAB increase exponentially because of the increase of bending rigidity. When the APAB is connected in parallel, the difference between the RF and the AF decrease more drastically than the case of series connection as the ratio $h_p^{(t)}/h_p^{(b)}$ approaches 0, which results in the

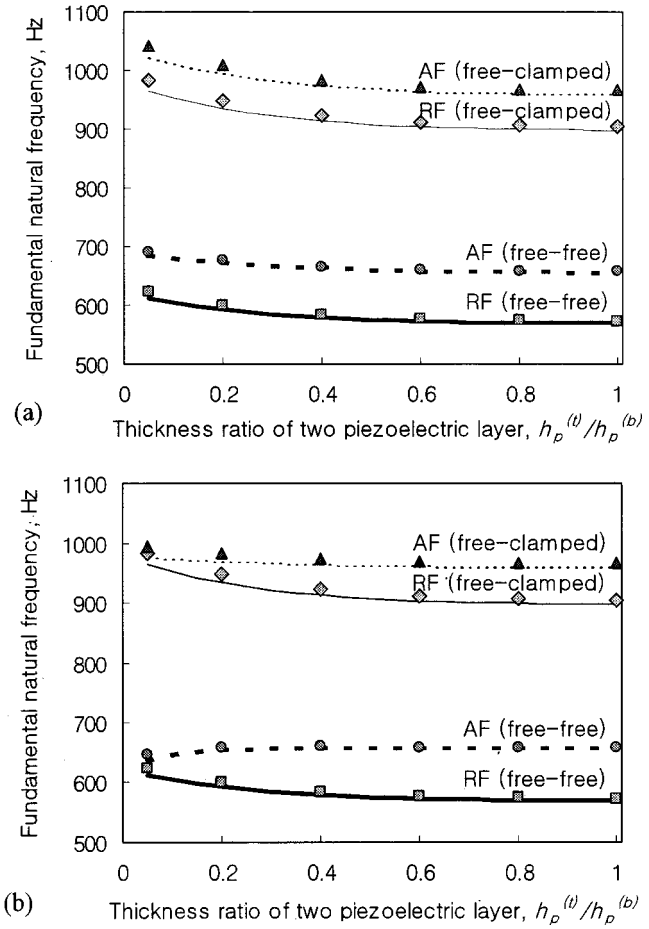


FIG. 4. The variations of the resonance frequencies (RF) and the antiresonance frequencies (AF) of the APAB (a) in series connection and (b) in parallel connection vs the thickness ratio $h_p^{(t)}/h_p^{(b)}$; $h_s=0.6$ mm and $h_p^{(t)}+h_p^{(b)}=0.4$ mm. Lines and symbols represent the results by the impedance matrix and the FEM, respectively.

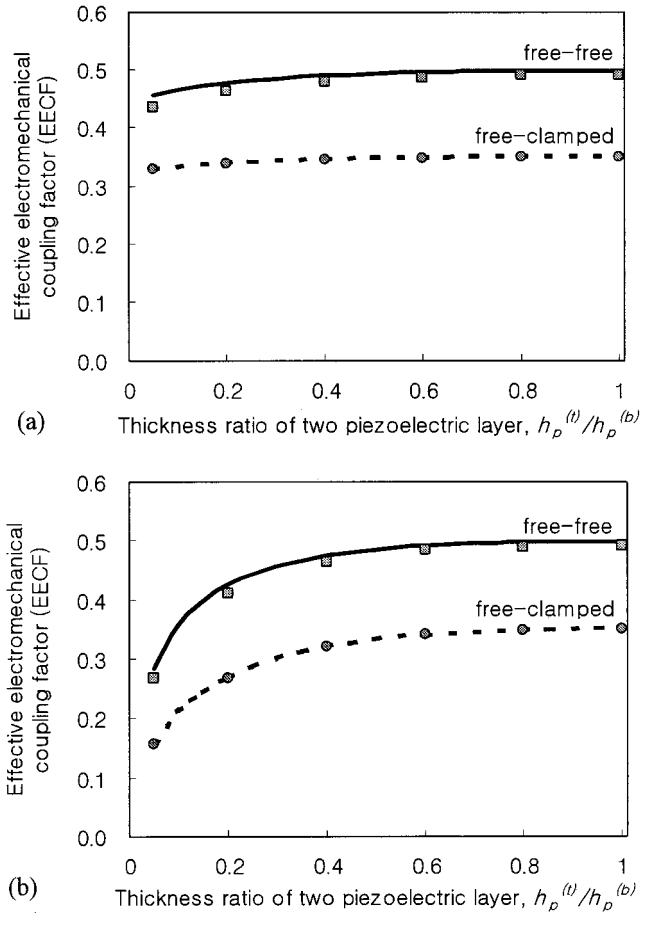


FIG. 5. The variations of the effective electromechanical coupling factors (EECF) of the APAB (a) in series connection and (b) in parallel connection vs the thickness ratio $h_p^{(t)}/h_p^{(b)}$; $h_s=0.6$ mm and $h_p^{(t)}+h_p^{(b)}=0.4$ mm. Lines and symbols represent the results by the impedance matrix and the FEM, respectively.

higher EECF of the APAB in series connection than those of the APAB in parallel connection. The variations of EECF versus the ratio $h_p^{(t)}/h_p^{(b)}$ are calculated using Eq. (41) and shown in Fig. 5 for the two electric connections. The results by the present method agree well with those by the FEM.

In the third special case, a triple-layer APAB is considered where one piezoelectric layer used for actuator and the other used for sensor. This type of transducer can be applied to the resonant sensor for static force measurement.[26,27] The voltage at the sensor, e.g., $V^{(b)}$, is caused by the deformation due to the actuator voltage, e.g., $V^{(t)}$. The same geometry and material properties as the second case are used in the numerical calculation. The ratio of the voltage at the sensor to the exciting voltage $V^{(b)}/V^{(t)}$ versus the ratio $h_p^{(t)}/h_p^{(b)}$ are calculated using the impedance matrix when the exciting frequency is the half of the RF and shown in Fig. 6. For the given geometry and material properties, the voltage ratio $V^{(b)}/V^{(t)}$ has maximum value when the ratio $h_p^{(t)}/h_p^{(b)}=0.1$ and 0.3 for the free-clamped and the free-free conditions, respectively. Good agreements are shown between the results by the impedance matrix and the FEM.

V. CONCLUSION

In this paper, the conjugate parameters for the admittance matrix of the triple-layer APAB with separate electrical ports are derived using the variational principle; the electric charge is represented by the averaged electric displacement

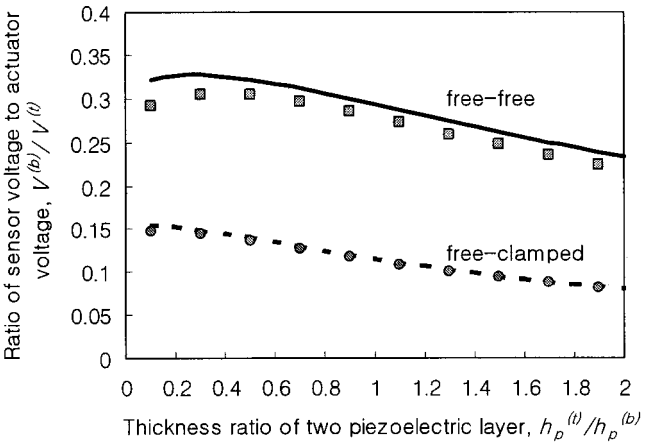


FIG. 6. The ratio of the voltage at the sensor to the exciting voltage $V^{(b)}/V^{(t)}$ vs the thickness ratio $h_p^{(t)}/h_p^{(b)}$ when the exciting frequency is the half of the RF; $h_s=0.6$ mm and $h_p^{(t)}+h_p^{(b)}=0.4$ mm. Lines and symbols represent the results by the impedance matrix and the FEM, respectively.

over the thickness rather than by that on the surface, which results from the common assumption of the constant electric field in each piezoelectric layer. It is also shown that the evaluation or derivation of the admittance and impedance matrices can be simplified using the reciprocity conditions that are often noticed in the electromechanical system. As a result, the total admittance matrix is explicitly expressed in terms of the mechanical admittance, the electromechanical coupling relationships, and the electrical admittance. Using the derived matrices, the effects of the thickness ratio of the two piezoelectric layers on the performance of the APAB are clearly revealed. It is demonstrated that the derived matrices can be applied to various types of annular bimorph with separate electrical ports as well as a single piezoelectric ring. It is expected that the simplified matrix manipulation of the impedance matrix presented in this paper can be easily extended to other types of piezoelectric transducers for future applications.

ACKNOWLEDGMENTS

This work was supported by the center of IDOT and the Brain Korea 21 Project.

[1] J. Bost, ''A new type of tweeter horn employing a piezoelectric driver,'' J. Audio Eng. Soc. **23**, 796–801 (1978).
[2] J. Bost, ''A new piezoelectric driver enhances horn performance,'' J. Audio Eng. Soc. **28**, 244–249 (1980).
[3] G. Perçin, L. Levin, and B. T. Khuri-Yakub, ''Piezoelectrically actuated droplet ejector,'' Rev. Sci. Instrum. **68**, 4561–4563 (1997).
[4] G. Perçin, A. Atalar, F. L. Degertekin, and B. T. Khuri-Yakub, ''Micromachined two-dimensional array piezoelectrically actuated transducers,'' Appl. Phys. Lett. **72**, 1397–1399 (1998).
[5] G. Perçin, T. S. Lundgren, and B. T. Khuri-Yakub, ''Controlled ink-jet printing and deposition of organic polymers and solid particles,'' Appl. Phys. Lett. **73**, 2375–2377 (1998).
[6] K. Uchino, *Piezoelectric Actuator and Ultrasonic Motors* (Kluwer Academic, Boston, 1996).
[7] H. Tanaka, ''Generalized basic equations for bending motions of piezoelectric bars from Hamilton's principle,'' J. Acoust. Soc. Am. **95**, 1768–1772 (1994).
[8] H. Tanaka and R. Aoyagi, ''Derivation of equivalent circuit of multilayered bending vibration,'' in *Proceedings of Acoustic Engineering*, Tohoku Univ., Japan, pp. 262–263, Feb., 1992 (in Japanese).
[9] R. Aoyagi and H. Tanaka, ''Equivalent circuit analysis of piezoelectric bending vibrators,'' Jpn. J. Appl. Phys. **33**, 3010–3014 (1995).
[10] Q. M. Wang, X. H. Du, B. Xu, and L. Eric Cross, ''Electromechanical coupling and output efficiency of piezoelectric bending actuator,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **46**, 638–646 (1999).
[11] J. G. Smits, S. I. Dalke, and T. K. Cooney, ''The constituent equations of piezoelectric bimorphs,'' Sens. Actuators A **28**, 41–61 (1991).
[12] J. G. Smits and A. Ballato, ''Dynamic admittance matrix of piezoelectric cantilever bimorphs,'' J. Microelectromech. Syst. **3**, 105–112 (1994).
[13] J. G. Smits, W. Choi, and A. Ballato, ''Resonance and antiresonance of symmetric and asymmetric cantilevered piezoelectric flexors,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **44**, 250–258 (1997).
[14] Y. S. Cho, Y. E. Park, C. S. Han, and S. K. Ha, ''Five-port equivalent electric circuit of piezoelectric bimorph beam,'' Sens. Actuators A **84**, 140–148 (2000).
[15] A. B. Dobrucki and P. Pruchnicki, ''Theory of piezoelectric axisymmetric bimorph,'' Sens. Actuators A **58**, 203–212 (1997).
[16] H. L. Hui, ''Axisymmetric response of circular plates with piezoelectric layers: An exact solution,'' Int. J. Mech. Sci. **40**, 1265–1279 (1998).
[17] S. Y. Lin, ''Coupled vibration analysis of piezoelectric ceramic disk resonators,'' J. Sound Vib. **218**, 205–217 (1998).
[18] A. Iula, N. Lamberti, and M. Pappalardo, ''A model for the theoretical characterization of thin piezoceramic rings,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 370–375 (1996).
[19] A. Iula, N. Lamberti, and M. Pappalardo, ''Analysis and experimental evaluation of a new planar piezoelectric accelerometer,'' IEEE/ASME Trans. Mechatron. **4**, 207–212 (1999).
[20] A. Iula, N. Lamberti, R. Carotenuto, and M. Pappalardo, ''Analysis of the radial symmetrical modes of thin piezoceramic rings,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **46**, 1047–1049 (1999).
[21] S. K. Ha and Y. H. Kim, ''Impedance and admittance matrices of symmetric piezoelectric annular bimorphs and their applications,'' J. Acoust. Soc. Am. **108**, 2125–2133 (2000).
[22] B. A. Auld, *Acoustic Fields and Waves in Solids*, 2nd ed. (Krieger, Melbourne, 1990).
[23] IEEE Standard on Piezoelectricity, ANSI/IEEE Std. 176-1987, p. 43 (1988).
[24] ANSYS, ''ANSYS User's Manual,'' Swanson Analysis Systems, Inc. Houston, PA, 1995.
[25] D. Boucher, M. Lagier, and C. Maerfeld, ''Computation of the vibrational modes for piezoelectric array transducer using a mixed finite element-perturbation method,'' IEEE Trans. Sonics Ultrason. **SU-28**, 318–330 (1981).
[26] C. Gehin, Y. Teisseyre, and C. Barthod, ''Piezoelectric resonant sensor for static force measurement,'' Colloque Interdisciplinaire en Instrumentation, Cachan, France, pp. 119–125, 1998.
[27] C. Gehin, C. Barthod, and Y. Teisseyre, ''Design and characterization of a new force resonant sensor,'' Sens. Actuators A **84**, 65–69 (2000).

Boundary effect of a viscous fluid on a longitudinally vibrating bar: Theory and application

Corinne M. Darvennes and Sally J. Pardue
Department of Mechanical Engineering, Tennessee Technological University, Box 5014, Cookeville, Tennessee 38505

(Received 10 January 2000; revised 2 April 2001; accepted 9 April 2001)

A theoretical model has been developed to represent the interaction of a bar in longitudinal vibration with its surrounding viscous fluid. This model was generated to better predict the length of timber piles used in bridge foundation. Experiments were performed with sand and with several water–glycerin mixtures surrounding an aluminum bar, showing a significant effect of the bounding material on the vibration characteristics of the bar. The viscous fluid was observed to lower the measured frequencies, indicating system mass loading, while the presence of the sand caused an upward shift in frequencies, indicating increased system stiffness. Hamilton's principle was used to model the longitudinal bar motion with the fluid loading due to the viscous boundary, following the procedure of a known soil–bar interaction model. It was found that the fluid model gives an excellent prediction of the natural frequencies of the bar. The modal mass and damping were not predicted quite as accurately, but the fluid model offered a significant improvement over a simple bar model. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1377293]

PACS numbers: 43.40.At, 43.40.Le, 43.40.Cw, 43.20.Tb, 43.20.Ks [CBB]

I. INTRODUCTION

The model presented in this paper was developed as part of the development phase of a project on determining the length of bridge piles using random vibration. The project goal was to develop and test a field measurement system for detecting, on site, the length of timber piles used in bridge foundations. These data are necessary to verify the integrity of the bridges and also to determine their usability under changing traffic patterns. In some cases, no records of the pile lengths were kept at the time of construction. In other cases, the original length is known but the timber piles might have broken or be rotted at the base. Currently, the most commonly used technique to evaluate piles is a visual inspection of their dry portion, close to the water surface; this provides some indication on the integrity of the wood but no definite data on pile length. A more sophisticated method is available,[1] but it is too cumbersome for on-site measurements and analysis by crews with limited technical training.

In the first phase of this project, a simple continuous model of the pile under longitudinal vibration was proposed which led to a relationship between length, wave speed, and natural frequencies ($L=c/2f$, for a free–free bar). To verify the concept, a simple lab experiment was performed using an aluminum rod. The use of aluminum avoided the complications of analyzing vibration data from wood, associated with the inhomogeneity of the wood grain structure. The preliminary experiments were also performed to check for the effects of the surrounding material on the bar's vibration response; did significant changes in the vibration signature produce an observable error in the length prediction using the simple equation? Since the pile of a bridge is exposed to various media such as air, water, mud, and different soils, depending on the bridge location, it was important to assess the influence these materials might have on the measured longitudinal vibration. In the laboratory, a decision was made to focus on easily obtained materials that could be readily characterized for later comparison and expansion of the studies. Dry sand was chosen for the soil type; glycerin–water mixtures, in different proportions, were used to assess the effects of a viscous watery mix such as a mud slurry.

The effects of these two materials on the vibration characteristics were indeed measurable, and consequently an effort was made to theoretically model these effects. The viscous fluid was observed to lower the measured frequencies, indicating system mass loading, while the presence of the sand caused an upward shift in frequencies, indicating increased system stiffness. In both cases, the damping was also observed to increase, although more significantly with the sand. The experiments performed and the resulting measurements are summarized in Sec. III.

A literature review provided several models for the soil–bar interaction (see Refs. 2 and 3, for example) but no fluid–bar model applicable to our situation was found. The approach of Ref. 2 was followed and modified to develop an appropriate viscous fluid–bar interaction model. This model, presented in Sec. II, was derived for an infinite surrounding medium, representative of a bridge pile. Then, we were able to find how much of the medium influences the bar vibration and define a fluid loading layer. For practical reasons, we were not able to experiment with an infinite surrounding medium. However, the cylinders used to contain the experiments were large enough to include the fluid loading layer and, thus, approximate an infinite medium. A comparison of analytical and experimental data, in Sec. IV, shows that including the effect of the material surrounding the bar significantly improves the length prediction.

II. THEORY

The formulation of the governing equation of longitudinal motion is derived using Hamilton's principle, which re-

 0001-4966/2001/110(1)/216/9/$18.00

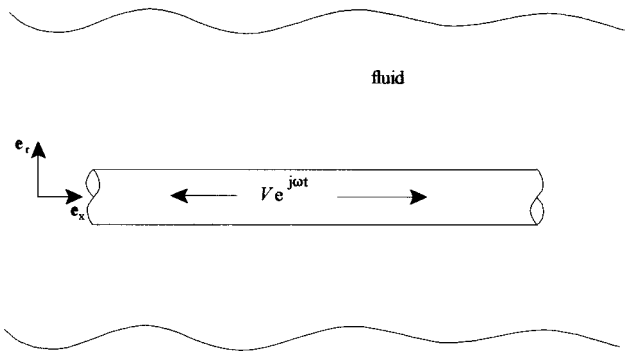


FIG. 1. General geometry of the cylinder in fluid.

lies on an analytical approach using variational principles on the energy of the system as a whole. For a continuous system, it is possible to obtain not only the differential equation of motion but also the associated boundary conditions needed to solve the differential equation.

A. Soil–bar interaction model

References 2 and 3 provide the basis of the soil–bar interaction model. The motion of a soil layer is treated as a cylinder with inner radius R_0, the bar's outer radius, and an outer radius extending to infinity. The equation for shear wave propagation in the soil is solved as a continuous model[3] which is used to define a complex vertical stiffness per unit length of the soil acting on the bar, k_{soil}. The equation for the bar displacement, u, with soil effects included is

$$EA\frac{\partial^2 u}{\partial x^2}-C_{\text{bar}}\frac{\partial u}{\partial t}-k_{\text{soil}}u=\rho_{\text{bar}}A\frac{\partial^2 u}{\partial t^2}, \tag{1}$$

where E, A, C_{bar}, and ρ_{bar} are the modulus of elasticity, cross-sectional area, damping coefficient, and density of the bar material, respectively, and k_{soil} is the complex stiffness of the soil. This equation can be simplified by assuming negligible damping, $C_{\text{bar}}=0$. This assumption is reasonable for the aluminum bar used in our experiments, but may not be correct for wooden timbers.

B. Fluid–bar interaction model

The motivation for the model is to theoretically explain the apparent mass loading as well as damping of the vibration of an aluminum bar surrounded by glycerin that was observed experimentally (as reported in Sec. III). The premise is to use the solution of the fluid motion about a uniform cylinder vibrating longitudinally with rigid body motion to approximate the fluid excited about a cylinder undergoing standing wave motion, following a procedure similar to that described for the soil–bar interaction. The geometry is shown in Fig. 1. The fluid velocity vector, $\mathbf{V}(r,\theta,x,t)$, has components v_r, v_θ, and v_x.

Several assumptions are made to yield a simple model with an analytical solution.

(1) The bar is considered long enough that changes in the fluid motion at and beyond the ends of the bar can be disregarded.
(2) The bar is surrounded by a fluid that extends to infinity in the radial direction. Indeed, only a small region of fluid around the bar affects the systems motion.
(3) The bar oscillates harmonically in the longitudinal direction with a velocity given as $v_x^{\text{bar}}=e^{j\omega t}\partial u(x,t)/\partial t$, where $u(x,t)$ is the longitudinal displacement of the bar at its periphery.
(4) The bar is vibrating in an incompressible fluid (density, ρ, and kinematic viscosity, ν, are constant).
(5) The fluid motion is axisymmetric, since we are considering the longitudinal motion of an axisymmetric bar. The fluid velocity can be expressed as $\mathbf{V}=\mathbf{V}(r,x,t),v_\theta=0$.
(6) The flow is fully developed with respect to x, $\mathbf{V}=\mathbf{V}(r,t)$. This is the least robust of the assumptions because the fluid motion is caused by standing waves in the bar, which are dependent on x. However, the amplitude of the longitudinal waves is small, and the radial vibration is small for a thin bar;[4] therefore, we can assume that the changes in fluid motion due to longitudinal variations in bar velocity are significantly smaller than the radial variations in fluid velocity. The x dependence of the bar velocity will be reintroduced later to calculate the kinetic energy of the fluid with more accuracy.
(7) Since the fluid motion is due only to the longitudinal vibration of the rod, we do not expect fluid motion in the radial direction, $v_r=0$. This can also be derived from mass conservation combined with assumptions (5) and (6).
(8) There is no outside pressure source.
(9) The fluid is initially at rest, $\mathbf{V}(t=0)=0$.

The boundary conditions are as follows: because of viscosity, the lateral bar motion is directly transmitted to the fluid, i.e., at $r=R_0$ (where R_0 is the bar radius), $v_x(R_0,t)=e^{j\omega t}\partial u(x,t)/\partial t$ and far enough away from the bar, the fluid motion becomes null, i.e., $v_x(\infty,t)=0$.

Under these assumptions, the problem is reduced to determining the longitudinal fluid velocity $\mathbf{V}=v_x(r,t)\mathbf{e}_x$ and the x-momentum equation simplifies to[5]

$$\frac{\partial v_x}{\partial t}=\frac{\nu_x}{r}\frac{\partial v_x}{\partial r}+\nu\frac{\partial^2 v_x}{\partial r^2}, \tag{2}$$

where ν is the kinematic viscosity of the fluid. Assuming a solution of the form

$$v_x(r,t)=v_1(r)e^{j\omega t}\partial u/\partial t, \tag{3}$$

where $v_1(R_0)=1$ and $v_1(\infty)=0$, and substituting it into Eq. (2) yields the following ordinary differential equation for the function $v_1(r)$:

$$\frac{d^2 v_1}{dr^2}+\frac{1}{r}\frac{dv_1}{dr}-j\frac{\omega}{\nu}v_1=0. \tag{4}$$

Letting $\alpha^2=\omega/\nu$, the equation can be identified as a Bessel equation. The general solutions are modified Bessel functions of order 0:[6] $v_1(r)=AI_0(\alpha r j^{3/2})+BK_0(\alpha r j^{1/2})$. In order to satisfy the second boundary condition of $v_1(r\to\infty)=0$, the constant A must equal 0 since $I_0(\alpha r j^{3/2})\to\infty$ as r

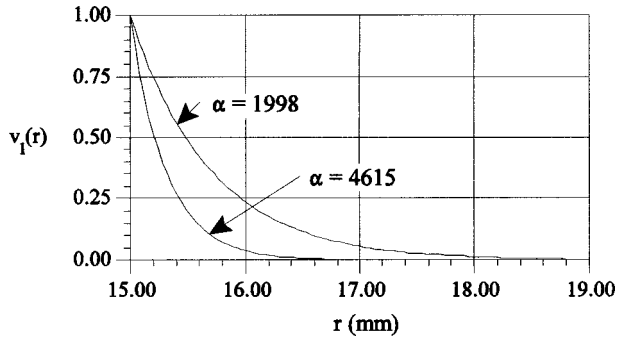


FIG. 2. Fluid velocity as a function of radius for glycerin surrounding an aluminum bar 15 cm in radius excited at 750 and 4000 Hz.

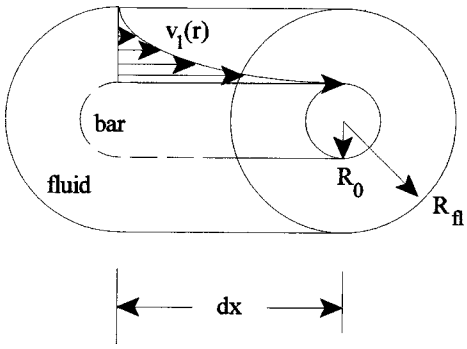


FIG. 3. Graphical representation of the fluid velocity and the fluid layer thickness.

$\rightarrow\infty$. Applying the first boundary condition, $v_1(R_0)=1$, the solution becomes

$$v_1(r)=\frac{K_0(\alpha r j^{1/2})}{K_0(\alpha R_0 j^{1/2})}, \quad r>R_0. \tag{5}$$

To establish the participation of the fluid surrounding the bar, it is proposed to determine the value of r at which the magnitude of the fluid velocity is approximately one percent of the bar velocity, $v_1(r)=0.01$. This value of r is called R_{fl}, thickness of the fluid loading layer, and is determined numerically. In our experiments, two frequencies were considered, 750 and 4000 Hz, representing the upper and lower frequency limits of the random excitation; the surrounding fluid was glycerin, with a kinematic viscosity of $\nu=1.18\times10^{-3}\,\mathrm{m^2/s}$, at 20 °C, which yields $1998\,\mathrm{m^{-1}}\leqslant\alpha=\sqrt{(\omega/\nu)}\leqslant4615\,\mathrm{m^{-1}}$. The function $v_1(r)$ is plotted in Fig. 2 for the upper and lower limits on the frequency parameter, α. The values of R_{fl} can be read from the figure as 18 mm (3 mm out from the rod surface) at 750 Hz and 16.5 mm (1.5 mm out from the rod surface) at 4000 Hz. It can be seen from the figure that more fluid is activated at the lower frequencies. This makes sense, since at the lower frequency the fluid mass has more time to move and excite the neighboring particles.

C. Kinetic energy of fluid

The next step is to establish the kinetic energy of the fluid for inclusion in the system kinetic energy term in the Hamiltonian approach

$$T_{\mathrm{fl}}=\frac{1}{2}mv^2=\frac{1}{2}\int_0^L\int_{R_0}^{R_{\mathrm{fl}}}2\pi\rho_{\mathrm{fl}}rv_1(r)^2\left(\frac{\partial u(x,t)}{\partial t}\right)^2 dr\,dx, \tag{6}$$

where

$$m=\int_{R_0}^{R_{\mathrm{fl}}}2\pi\rho_{\mathrm{fl}}r\,dr$$

represents the mass per unit length of the fluid in the annular cylinder defined between R_0, the bar radius, and R_{fl}, where the fluid velocity has dropped to one percent of the bar velocity. The fluid velocity is given by Eq. (3), $v=v_1(r)\partial u(x,t)/\partial t$. Originally, we assumed that the velocity of the bar was constant with respect to x along the boundary; recall assumption (6). As mentioned then, we now partially remove this assumption by considering that the fluid velocity takes on the magnitude of the local bar velocity. This concept is depicted graphically in Fig. 3.

The kinetic energy of the fluid in motion about the aluminum bar can be rewritten in terms of the bar longitudinal velocity, $v_x^{\mathrm{bar}}(x,t)=\partial u(x,t)/\partial t$, and the mass per unit length of the activated fluid, m_{actfl}, as

$$T_{\mathrm{fl}}=\frac{1}{2}\int_0^L m_{\mathrm{actfl}}v_x^{\mathrm{bar}}(x,t)^2dx,$$

$$\text{where}\quad m_{\mathrm{actfl}}=2\pi\rho_{\mathrm{fl}}\int_{R_0}^{R_{\mathrm{fl}}}rv_1(r)^2dr. \tag{7}$$

The kinetic energy of the total system can be written as

$$T_{\mathrm{sys}}=T_{\mathrm{bar}}+T_{\mathrm{fl}}=\frac{1}{2}\int_0^L(m_{\mathrm{bar}}+m_{\mathrm{actfl}})v_x^{\mathrm{bar}}(x,t)^2dx, \tag{8}$$

where T_{bar} and m_{bar} are the kinetic energy and the mass per unit length of the aluminum bar, respectively. The effect of the fluid mass can now be incorporated in the motion of the bar by simply adjusting the mass of the bar with the activated fluid mass. A similar procedure is used in one-degree-of-freedom vibratory systems to account for the mass of a spring by adding 1/3 of the spring mass to the system mass.[7]

D. Energy loss

An energy loss term modeling the fluid damping on the motion of the bar is required to complete the model of the interaction of the fluid and the bar. The shear stress in the fluid is given by $\tau(r)=\mu_{\mathrm{fl}}\partial v_{\mathrm{fl}}/\partial r$, where μ_{fl} is the absolute viscosity of the fluid ($\mu=\nu\rho$). Multiplying the shear stress at the wall by the contact area yields the force on the bar due to the fluid motion as

$$F_{nc}=2\pi R_0\mu_{\mathrm{fl}}\left.\frac{dv_1(r)}{dr}\right|_{r=R_0}\int_0^L v_x^{\mathrm{bar}}dx=C_{\mathrm{fl}}\int_0^L v_x^{\mathrm{bar}}dx, \tag{9}$$

where C_{fl} is a damping constant per unit length defined by

$$C_{\mathrm{fl}}=2\pi R_0\mu_{\mathrm{fl}}\left.\frac{dv_1(r)}{dr}\right|_{r=R_0}.$$

The differential equation for the system can now be stated as

TABLE I. Mass and damping of the activated fluid, for glycerin and a 90/10 glycerin–water mix.

	Glycerin $\nu=1.11\times10^{-3}$ m^2/s, $\rho=1264$ kg/m^3		90% Glycerin, 10% Water $\nu=2.08\times10^{-4}$ m^2/s, $\rho=1148$ kg/m^3	
ω(rad/s)	m_{actfl} (kg/m)	C_{fl} (N s/m^2)	m_{actfl} (kg/m)	C_{fl} (N s/m^2)
950(2π)	0.0385	203.7	0.0153	71.5
1900(2π)	0.0275	277.2	0.0108	93.8
2850(2π)	0.0215	330.2	0.0088	108.7
3800(2π)	0.0194	372.9	0.0076	119.8

$$EA\frac{\partial^2 u}{\partial x^2}-(m_{\text{bar}}+m_{\text{actfl}})\frac{\partial^2 u}{\partial t^2}-C_{\text{fl}}\frac{\partial u}{\partial t}=0, \tag{10}$$

where E and A are the constant modulus of elasticity and cross-sectional area of the aluminum bar, respectively, m_{bar} is the mass per unit length of the bar, m_{actfl} is the mass per unit length of the activated fluid, and C_{fl} is the damping constant of the fluid. The parameters m_{actfl} and C_{fl} are functions of the fluid distribution curve, $v_1(r)$. The calculations of m_{actfl} and C_{fl} for various fluid viscosities, ν, and frequencies of excitation, ω, are tabulated in Table I. As expected, the water–glycerin mix being less dense and less viscous than pure glycerin will contribute less mass and damping to the vibrating bar.

III. EXPERIMENTS

A. Experimental setup

The experimental program examined the longitudinal vibration response of a free–free cylindrical bar 2.74 m in length, 30 mm in diameter, made of aluminum. Material constants for the specific aluminum alloy used in the experiment were carefully determined. The rod was weighed and the aluminum density, ρ, was calculated as 2775 kg/m^3. The modulus of elasticity, E, of the aluminum was calculated to be 7.52×10^{10} N/m^2 from the measured natural frequencies of the free–free suspended bar.

Three configurations were considered: the aluminum rod in air, immersed in a viscous fluid, and surrounded by dry masonry sand. The excitation used in all the experiments was a random force input. Random excitation was selected over impulse excitation to allow for the control of the input force with respect to frequency content and amplitude.

1. Bar suspended in air

The bar was suspended horizontally from overhead supports by two thin steel wires, as shown in Fig. 4. The theoretically predicted natural frequencies were expected to match the experimental values closely. A force transducer and piezoelectric actuator were attached into one end of the bar via a specially machined bolt. The accelerometer was screwed to the opposite end of the bar.

2. Bar surrounded by fluid

The same aluminum bar was suspended horizontally inside a clear acrylic tube, as seen in Fig. 5. The acrylic tube was much larger than the fluid loading layer, R_{fl}, and thus approximated an infinite surrounding medium. Two different seals were tried. The first approach was to stretch and secure a thin latex membrane across the face of the flanges which were flush with the end of the aluminum bar. The transducers were attached to the bar through the membrane. The second approach was to use a flat ring of rubber sheeting as the sealing device, which left the ends of the aluminum bar exposed for attachment of transducers. The rubber seal was chosen for the majority of the experiments because the latex was prone to leaks. The advantage of the latex, though, was that it caused virtually no change to the bar's response, whereas the rubber seals added stiffness to the system.

Three different fluids were used: pure water, pure glycerin, and a 90% glycerin–10% water mix. The kinematic viscosities of the fluids were measured with a Canon–Fenske viscometer at room temperature, 20 °C. The viscosities were measured as 1 centistoke (mm^2/s) for tap water, 1110 centistokes for glycerin, and 208 centistokes for the 90/10 mix. The effects of three orders of magnitude of fluid viscosity were thus examined.

3. Bar surrounded by sand

A second aluminum bar with the same dimensions as the first bar was placed in a vertical position and surrounded by sand; see Fig. 6. The bar was supported and centered by two wooden end plates. The sand was contained by two sections of 15.24-cm-diameter metal ducting (large enough to ap-

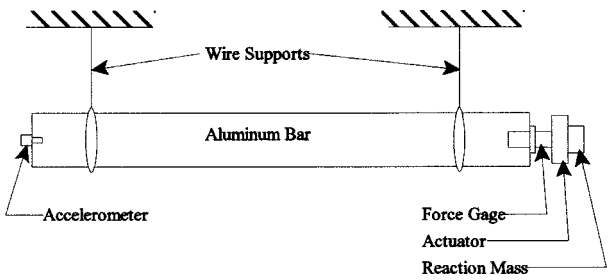


FIG. 4. Aluminum bar suspended in air.

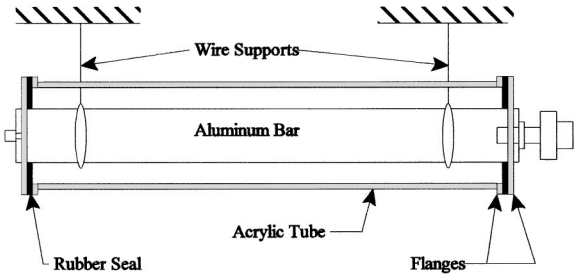


FIG. 5. Aluminum bar suspended inside the sealed acrylic tube.

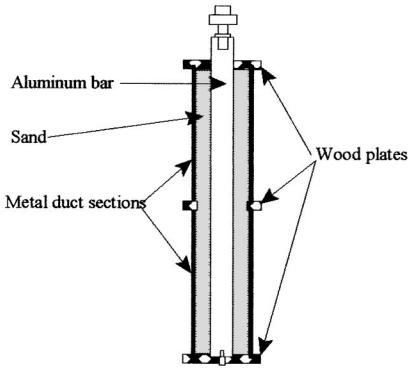


FIG. 6. Aluminum bar in the vertical sand bed.

proximate an infinite surrounding). The total length of surrounding sand was 2.44 m. The upper end of the aluminum bar extended 30.5 cm beyond the top plate. The force transducer and actuator were located on the top end of the aluminum bar and the accelerometer was located at the bottom end of the bar.

Measurements were made on the system in four different configurations. First, the system was assembled completely without sand to get a set of baseline measurements, since the end conditions for the aluminum bar were no longer free–free and it was desirable to see the effects of the new boundary conditions. Next, only the first section of ducting was filled with sand (called half-sand configuration). Then, the whole assembly was filled with sand (called full sand). Finally, the bar was pulled 30.5 cm upward through the sand. The fourth set of measurements was taken because unexpected results were observed with the second and third sets of data, due to the influence of the base plate.

4. Measurement equipment

A schematic of the measurement setup is shown in Fig. 7. It consists of two computers, a forcing device, and transducers for measuring force and acceleration. One computer supplied the random excitation signal to the forcing device. The other computer contained a data acquisition card to record the signals from the force and acceleration transducers. More specific detail on the equipment can be found in Ref. 8.

The reference frequency spectrum was specified as constant from 750 to 4000 Hz with an amplitude such that the output signal had a value of 0.5 V rms. The range of frequency excitation was sufficient to excite the first four resonances of the 2.74-m aluminum bar in all three systems. However, the third resonance of the bar occurred near an anomaly in the frequency spectrum of the shaker characterized by a ripple effect in the otherwise flat response of the shaker between 2500 and 3000 Hz.

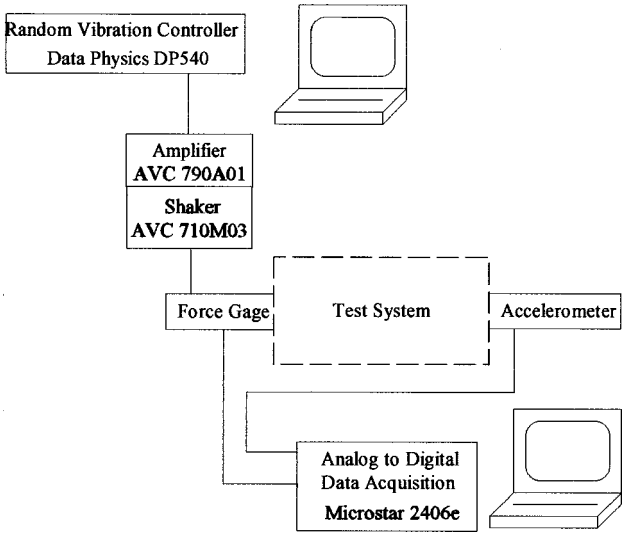


FIG. 7. Measurement equipment.

The known input (force) and output (acceleration) were used to compute the system transfer function or frequency response function (FRF). The procedures used to compute the system FRF and subsequently analyze it for parameter identification are standard procedures, described in Refs. 9 and 10. The FRF of a multiple-degree-of-freedom (dof) system exhibits peaks at the natural frequencies (ω_i) of the system. Since the peaks are well separated, single dof (sdof) systems are used to model each peak. A simplex algorithm[11] was used to optimize each of these sdof systems to identify a gain, $\text{Gain}_i = 1/M_i$, a natural frequency, $\omega_i = \sqrt{K_i/M_i}$, and a damping ratio, $\zeta_i = C_i/2M_i\omega_i$, where M_i, K_i, and C_i are the modal parameters or mass, stiffness, and damping coefficient, respectively, of the sdof system used to represent peak *i*.

B. Experimental results

The effects of various numerical processing and signal digitization factors on the spectral density estimates were evaluated and are reported in Ref. 12. The signal-analysis parameters used in the data reported here are as follows (unless otherwise noted): number of records, $n_d = 10$; bandwidth, $B_S = |H_{\max}|/10$, where $|H_{\max}|$ is the maximum magnitude of the FRF; sampling rate, $\Delta t = 40\,\mu$s, and resolution,

TABLE II. Aluminum bar in air.

	Bar in air			Empty fluid assembly rubber seals			Empty sand assembly $B_S = \lvert H_{\max}\rvert/\sqrt{2}$		
Mode	Gain	f (Hz)	ζ	Gain	f (Hz)	ζ	Gain	f (Hz)	ζ
1	26.7	949.1	0.0007	24.6	956.1	0.0030	30.5	958.3	0.0009
2	25.2	1898.2	0.0005	24.7	1903.3	0.0007	⋯	⋯	⋯
3	25.8	2850.7	0.0014	27.9	2854.2	0.0018	31.6	2841.4	0.0021
4	25.3	3795.7	0.0005	23.7	3801.0	0.0008	22.4	3785.8	0.0023

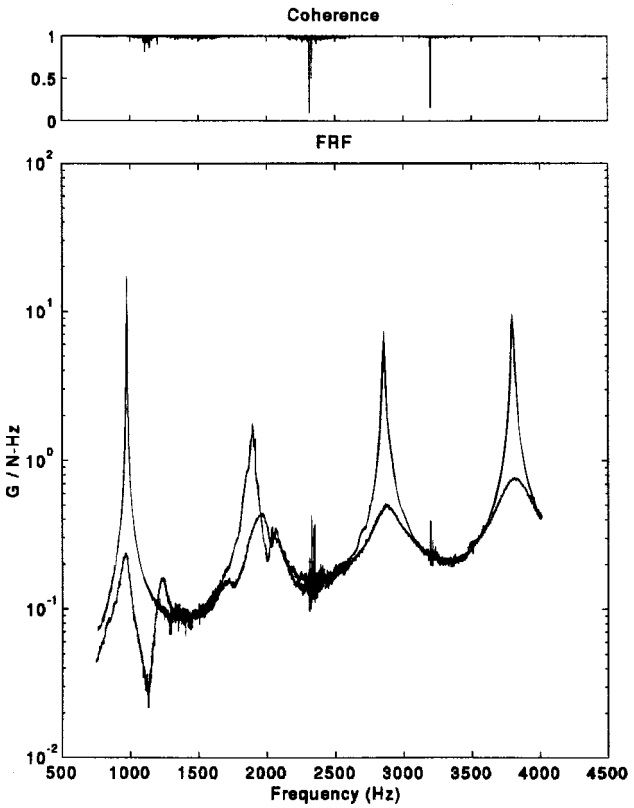


FIG. 8. FRF and coherence for aluminum bar in the sand assembly; (i) no sand (sharp peaks) and (ii) half sand; the FRF is an acceleration over force per time (G/N Hz).

$\Delta f = 0.76$ Hz. Each test was performed several times and showed excellent repeatability, with no visible changes in the FRFs.

1. Aluminum bar in air

Table II provides gain, frequency, and damping ratio for the first four modes of the bar suspended in air, in the three configurations described in Sec. III A.

The bar suspended in air provides the baseline for comparison for all the experimentation. Comparing the resonant frequencies in air and in the empty acrylic tube with rubber seals reveals upward frequency shifts on all the modes. This can be attributed to the increased stiffness of the system caused by the rubber seals used to secure the fluid inside the acrylic tube. Considering the resonances of the bar in the empty sand assembly, we observe a shift upward on the fundamental mode and downward shifts on the third and fourth modes. The damping on each mode is also increased.

Note that no data are available for mode 2 in the empty sand assembly. Because of large damping, the bandwidth of data, B_S, was changed to $|H_{\max}|/\sqrt{2}$ and the second peak was not able to be fitted with this width of data selected. This can be seen in Fig. 8, which shows the FRF and coherence in sand (empty assembly and half-sand configurations). Also visible in this figure is an unexpected peak following the second mode resonance. The coherence is essentially 1 in this frequency range, so the phenomenon cannot be discounted as a data-processing flaw. This peak can be explained as a by-product of the assembly conditions. The base of the bar is resting on a wooden plate whose modes of vibration may have influenced the response of the bar. The boundary condition at the base end of the bar is no longer free, but could be considered a flexible connection. The condition at the top end of the bar is still free, although the wooden plate slipped over the bar and resting 0.3 m down from the top may contribute a slight stiffness addition to the system.

2. Aluminum bar surrounded by fluid

The results of adding water, a glycerin–water mix, and pure glycerin to the system of the aluminum bar in the acrylic tube are shown in Table III. Water and mix data were taken with the rubber seals, while glycerin data were recorded with latex seals.

Adding water has little effect on the resonant frequencies: comparing Tables II and III shows that the first, second, and fourth frequencies are the same within one Δf resolution; discrepancies on the third frequency are attributed to the shaker anomaly in this frequency region (described in Sec. III A). Comparing the damping estimates for air and water in the sealed tube shows some interesting results: the value of ζ increases on modes 1, 3, and 4 and decreases on mode 2, and in mode 2, ζ is now an order of magnitude lower than the value for the other three modes. We have no definite explanation for this observation.

The glycerin–water mix shows a downward shift in the frequencies and an increase in damping. The largest frequency shift is in the third mode, as indicated by a 0.4% difference from the pure water case.

Glycerin was tested in the tube sealed with latex. Before addressing the effects of the glycerin on the system, the empty tube was examined with latex seals: the latex seals were very flexible and did not exhibit the stiffness increase noted with the rubber seals. Therefore, the pure glycerin data were compared with the tabular results for the freely suspended bar. When glycerin was added to the tube, the fre-

TABLE III. Aluminum bar in liquid.

	Water–rubber seals			Mix–rubber seals			Glycerin–latex seals		
Mode	Gain	f (Hz)	ζ	Gain	f (Hz)	ζ	Gain	f (Hz)	ζ
1	25.3	955.1	0.0041	26.0	953.0	0.0046	26.5	942.5	0.0077
2	25.3	1901.7	0.0003	26.0	1897.7	0.0027	25.6	1888.5	0.0056
3	25.0	2851.2	0.0024	24.1	2840.5	0.0033	25.0	2835.5	0.0056
4	24.5	3801.9	0.0012	25.2	3793.2	0.0021	23.2	3781.9	0.0043

TABLE IV. Aluminum bar in sand.

	Half sand			Full sand			Full sand—raised bar		
Mode	Gain	f (Hz)	ζ	Gain	f (Hz)	ζ	Gain	f (Hz)	ζ
1	64.4	963.5	0.0336	73.6	961.2	0.0424	26.6	998.4	0.1334
2	30.3	1951.8	0.0384	21.4	1910.3	0.1413	37.3	1893.7	0.0534
3	27.7	2876.6	0.0369	34.2	2876.3	0.0586	22.6	2892.2	0.0621
4	21.7	3813.5	0.0304	24.9	3849.0	0.0552	25.2	3832.3	0.0311

quencies of the system were observed to shift downward significantly, as would be expected from an added mass.

3. Aluminum bar surrounded by sand

The results of placing the bar in the three sand configurations described in Sec. III A are shown in Table IV. The damping ratios in the half-sand are essentially frequency independent. It is seen that the various sand amounts strongly influence the frequencies of the first two modes, but have little effect on the higher modes. This will be discussed in Sec. IV D.

IV. COMPARISON OF EXPERIMENTAL AND ANALYTICAL DATA

The models of fluid and soil interactions are considered. The pure glycerin case provided the clearest indication of significant effects of a surrounding fluid on the vibration of the aluminum bar, as seen in Table II. Therefore, this case is presented for comparison between theory and experiment. For the soil interaction comparison, the case when the bar has been pulled up 0.3 m through the sand is presented, because it helps eliminate the boundary condition effects of the base plate on the measurements.

A. Natural frequency

Table V summarizes the measured and analytical values for the natural frequencies of the bar in air, in glycerin, and in sand. To evaluate the natural frequencies in glycerin, it was assumed that the stiffness of the system remained unaltered by the addition of the fluid, since the fluid is not capable of elastic resistance to motion. Table V shows an excellent agreement between measured and predicted frequency values both in air and in glycerin, with the largest difference, on mode 1 in glycerin, only 0.2%.

The predicted resonances in the sand were evaluated assuming free–free boundary conditions, neglecting the damping per unit length of the bar, C_{bar}, and using the parameters presented in Ref. 13 for homogeneous layers of soil, while the model stated in Ref. 14 was used to estimate the shear modulus for the sand, $G=6.0\times10^6\ \text{N/m}^2$, even though it was empirically developed for undisturbed earth and not small columns of soil above ground.

The predicted third and fourth frequencies are in the neighborhood of the experimentally determined values. The agreement on the first two modes is not as good. The assumption of free–free boundary conditions used to obtain the theoretical data is probably not a good one for the bar embedded in sand, since the mass of the sand is acting like an end mass on the bar. This end mass was neglected in our theoretical model because we were interested in the effect of the material surrounding the bar and the free–free boundary simplified the equation for the bar length. Other boundary conditions still need to be considered, to further improve the present model.

TABLE V. Comparison of the bar's natural frequencies in air, glycerin, and sand.

	Measured			Calculated		
Mode	f_{air} (Hz)	f_{glycerin} (Hz)	f_{sand} (Hz)	f_{air} (Hz)	f_{glycerin} (Hz)	f_{sand} (Hz)
1	949	942	998	949	940	1063
2	1898	1888	1893	1898	1885	1954
3	2851	2836	2892	2847	2832	2884
4	3796	3782	3832	3796	3778	3824

B. Modal mass

Table VI summarizes the measured and analytical values for the modal mass of the bar in air and in glycerin. The measured values of the bar in air, M_{air}, match the predicted modal value of $M=2.72$ only on mode 1. This is probably due to the error introduced by using sdof fitting of the data in the parameter identification stage. However, the deviation from the true value of $M_{\text{air}}=2.72$ was no more than 5.5%, an acceptable level. Comparing the values of the bar in glycerin, M_{glycerin}, with the value for the aluminum bar in air appears to indicate only a small change in modal mass; however, the associated shift in frequencies this additional mass causes is seen to be significant in Table V.

TABLE VI. Comparison of the bar's mass in air, and in glycerin.

	Measured			Calculated	
Mode	f_{air} (Hz)	M_{air} (kg)	M_{glycerin} (kg)	M_{air} (kg)	M_{glycerin} (kg)
1	949	2.72	2.70	2.72	2.770
2	1898	2.57	2.61	2.72	2.758
3	2851	2.63	2.55	2.72	2.749
4	3796	2.58	2.36	2.72	2.746

C. Damping

Table VII summarizes the measured and analytical values for the modal damping of the bar in air and in glycerin. For the aluminum, the damping per unit length is taken as constant and assumed to be $C_{\text{bar}}=20\ \text{N s/m}^2$.

The measured values of C in Table VII indicate a dependence on frequency, implying that an assumption of

TABLE VII. Comparison of the bar's damping in air and in glycerin.

	Measured				Calculated	
Mode	f_{air} Hz	f_{glycerin} Hz	C_{air} N s/m	C_{glycerin} N s/m	C_{air} N s/m	C_{glycerin} N s/m
1	949	942	27.43	246.2	19.22	279.4
2	1898	1888	27.43	346.9	28.75	380.2
3	2851	2836	27.43	508.8	42.60	452.9
4	3796	3782	27.43	482.3	81.69	511.5

TABLE VIII. Errors in predicted length.

	Simple model			Advanced model			
System	f (Hz)	L_{pred} (m)	% Diff	f (Hz)	L_{pred} (m)	% Diff	Model
Bar	$f_1=949$	2.74	0.0				$L=ic_0/2\,f_i$
	$f_4=3796$	2.74	0.0				
Glycerin	$f_1=942$	2.76	0.7	942	2.74	0.0	$L^2=\frac{(i\pi)^2 EA}{(m_{\text{bar}}+m_{\text{actfl}})\omega_i^2+j\omega_i C_{\text{fl}}}$
	$f_4=3782$	2.75	0.4	3782	2.74	0.0	
Sand	$f_1=998$	2.61	4.7	998	2.81	2.5	$L^2=\frac{(i\pi c_0)^2}{\omega_i^2-k_{\text{soil}}/m_{\text{bar}}}$
	$f_4=3832$	2.71	1.1	3832	2.73	0.4	

C_{bar}=constant made for the analytical model was incorrect. The model used during parameter identification of the experimental data was viscous damping (proportional to velocity); however, the damping in the aluminum bar by itself is hysteretic; that is, energy is dissipated by internal friction between various planes of material during deformation. The energy lost per vibration cycle is proportional to the square of the amplitude of vibration and to the hysteretic damping constant, h. For the aluminum, h can be calculated from the modal C and is found to be approximately $0.001\,K$, where K is the modal stiffness. The measured data also yield the same relation for h.

Comparing the modal estimates of C obtained from the experimental data for glycerin with the predicted values shows the largest percent difference on mode 1 with 13.5% difference. The damping ratios of the experimental data for glycerin are not constant with respect to frequency, which was expected from our fluid interaction model; however, the model does not precisely emulate the actual damping. The experimental data for the glycerin system were therefore analyzed to assess whether the damping might be treated as proportional damping, $[C]=\boldsymbol{\alpha}[M]+\boldsymbol{\beta}[K]$, but the values obtained for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ using different sets of measured $\boldsymbol{\omega}$ and ζ were not constant, indicating that the damping cannot be considered proportional.

D. Error in length prediction

The overall goal of this research is to establish a means of identifying the unknown length of a timber pile. The resonance frequencies of the structure are used to determine the length. The experimental results for the aluminum bar are used to demonstrate the method of predicting length from measured resonances. First, a simple model length prediction, assuming a free–free bar, is applied to the three sets of measurements: the bar in air, in glycerin, and in sand. Then, the models for liquid and soil interactions are used to improve the length estimate. The results for the measured frequencies, the predicted length, and the percent difference from the known length (2.74 m) are tabulated in Table VIII for modes 1 and 4. The wave speed in the aluminum material is taken to be $c_0=5200$ m/s.

The solutions for the fluid and sand models are evaluated assuming free–free boundary conditions. Modeling the fluid loading on the bar shows a significant improvement on the length prediction. The fluid model yields an exact length in glycerin, an excellent refinement over the 0.4% and 0.7% differences obtained with the simple model.

Predicted lengths in sand are also greatly improved by the use of the soil interaction model. The length prediction from the first resonance retains the highest error, 2.5% (substantially down from 4.7%, though). The use of a mass loaded boundary condition for the bar would shift the fundamental frequency more than the upper resonances. The assumption of free–free boundary conditions is therefore more likely to hold true for the higher resonances. This tendency is demonstrated by observing the improvement in predicted length using the fourth measured resonance. The error in the length prediction has improved, from 1.1% to 0.4%, using mode 4.

V. CONCLUSIONS

The natural frequencies of standing waves in a bar can be used to predict the bar length. A model was derived to model the effect of a fluid on the vibration of a bar. A similar procedure was found in the literature to model soil interaction. The validity of these two models was tested experimentally using a bar in glycerin and in sand, respectively. It was found that the fluid model gives an excellent prediction of the natural frequencies of the bar. The modal mass and damping were not predicted quite as accurately, but the fluid model offered a significant improvement over the simple bar model (which completely neglects the effect of the surrounding fluid). Likewise, the soil model significantly improved the frequency and length predictions.

ACKNOWLEDGMENTS

The help and useful discussions provided by Dr. R. J. Houghton are gratefully acknowledged. This work was supported by the Tennessee Department of Transportation.

[1] R. A. Douglas and J. D. Holt, ''Determining the Length of Installed Timber Pilings by Dispersive Wave Propagation Methods,'' FHWA/NC/94-002 (Center for Transportation Engineering Studies, North Carolina State University, 1993).
[2] K. Sun and J. A. Pires, ''Simplified approach for pile and foundation interaction analysis,'' J. Geotech. Eng. **119**, 1462–1479 (1993).
[3] M. Novak and F. Aboul-Ella, ''Impedance functions of piles in layered media,'' J. Eng. Mech. Div., Am. Soc. Civ. Eng. **104** (EM6), 643–661 (1978).

[4] K. F. Graff, *Wave Motion in Elastic Solids* (Dover, New York, 1975).
[5] R. B. Bird, W. E. Stewart, and E. N. Lightfoot, *Transport Phenomena* (Wiley, New York, 1960).
[6] N. W. McLachlan, *Bessel Functions for Engineers*, 2nd ed. (Oxford University Press, New York, 1961).
[7] A. Dimaragonas, *Vibration for Engineers*, 2nd ed. (Prentice-Hall, Englewood Cliffs, NJ, 1995), p. 764.
[8] S. J. Pardue, and J. R. Houghton, ''Embedded Pile Length Determination by Controlled Random Excitation,'' Proceedings of the 13th International Modal Analysis Conference, SEM, 1437–1443 (1995).
[9] J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures*, 2nd ed. (Wiley, New York, 1986).
[10] D. J. Ewins, *Modal Testing: Theory and Practice* (Wiley, New York, 1984).
[11] J. Nelder and R. Mead, ''A simplex method for function minimization,'' Comput. J. (UK) **7**, 308–313 (1965).
[12] S. J. Pardue, ''Longitudinal Random Vibration Analysis Applied to a Bar Surrounded by Fluid and Sand Media,'' Ph.D. dissertation, Tennessee Technological University, Cookeville, TN, 1995.
[13] A. S. Veletsos and A. K. Dotson, ''Vertical and torsional vibration of foundations in inhomogeneous media,'' J. Geotech. Eng. **114**, 1002–1021 (1988).
[14] S. Prakash, *Soil Dynamics* (McGraw-Hill, New York, 1981).

Finite-element vibration analysis of a cantilever plate with a central circular hole subject to an in-plane moving (rotating) load

Fawzi M. A. El-Saeidy
Damro-Salman, Desouk, Kafr-El-Sheikh, Egypt

(Received 29 September 1999; accepted for publication 22 January 2001)

In rotating radial ball bearings supported on elastic casings with the bearing outer ring lightly fitted into the housing, the force due to the ball elastic contact is indeed a rotating load vector rolling over the housing. For accurate estimation of the casing annulus (circular hole) dynamic deformations, which in turn affect bearing tolerances and the magnitude of the generated forces, the effect of the load rotation (motion) on the housing vibration should be considered. Considering the integral casing and outer ring to be a plate, an isoparametric plane stress finite-element (FE) based analytical procedure is presented for dynamic analysis of housing as affected by load vector rotation. The equations of motion are obtained using Lagrange's equations and decoupled using the normal coordinates representation and solved using a special numerical integration scheme. The computations are carried out using the FE program DAMRO 1. Results in both time and frequency domains are discussed and it is found that higher load rotational speeds decrease deformations and increase the smoothness of the annulus surface. With lower speeds, the deformations are always positive (contraction), and at higher speeds they exhibit both contraction and expansion around the annulus circumference. The vibrations measured at the casing outer surface show that the spectrum in the direction of casing rigid support attracts more system natural frequencies compared to the orthogonal direction spectrum. This underlines the importance of the vibration measuring probe(s) orientation with respect to the casing rigid support direction to capture all the important vibrations.
 [DOI: 10.1121/1.1360237]

PACS numbers: 43.40.Dx [PJR]

I. INTRODUCTION

Flat plates are widely used in practice, ranging from, for example, a single plate supporting a shaft bearing to a collection of plates constituting a gear box casing, etc. The applied loads on the elastic member can be pure in-plane, or lateral, or a combination, and they can be stationary or moving. As a result, the plate will vibrate in its plane, or laterally, or both. The study of flat plates under lateral moving loads has received much work using integral transformations methods[1] and approximate techniques such as FEM.[2] To the author's knowledge, a flat plate subject to in-plane moving loads has not been studied before, and the published studies are limited to in-plane stationary concentrated/distributed forces, to primarily determine the plate buckling loads and/or the related vibrations; see, for example, Refs. 3–5. Taha and Crookall[6] used an existing FE program (using triangular plate elements) to compute in-plane deformations of a cantilver thin plate with a central circular hole subject to in-plane stationary (nonrotating) radial forces due to contact between roller bearing rollers and rings. The bearing outer ring was treated as an integral part of the plate (casing). Also, plate finite elements were used, using existing FE programs, to model machine casings. See, for example, Ref. 7 where the linear springs and dampers representing the supporting bearing were connected to the housing wall at fixed locations.

In mechanical systems incorporating radial ball bearings, the ball force due to contact with the bearing rings is indeed a rotating vector rolling over the support; see, for example, Refs. 8, 9, and 10 where the housing with the outer ring fitted in it was treated rigid (except the Hertzian local elastic contacts). For elastic housings, the effect of the ball force rotation (motion) on the casing annulus (hole) surface dynamic deformations under a moving ball which, in turn, affect bearing tolearances and the magnitude of the generated (reaction) forces should be considered. The contribution of this paper is to present an FE formulation for the in-plane vibration analysis of a flat plate with a central circular hole subject to an in-plane rotating (moving) load vector. Although applications to rotating machinery incorporating radial ball bearings supported on elastic housings have been a guiding consideration, the analysis may have other similar applications.

II. ANALYTICAL MODEL

Figure 1 shows the FE discretization of a cantilever plate rigidly fixed along the X axis with a central circular hole subject to an in-plane rotating load vector $\mathbf{F}$. Here, the internal nodes (number 2, 4, 6, and 8; see Fig. 2) of each element are not shown. XYZ is a global coordinate system where the Z axis (not shown) is pointing out of the paper. In Fig. 2, which describes the details of a typical plate finite element under the action of the external moving (rotating) load vector $\mathbf{F}$, the axes $\overline{XY}$ are fixed in space and parallel to the XY axes with the origin at the geometrical center of the annulus in the undeformed state (zero speed position). In Fig. 2, $\mathbf{F}$ is shown within the spatial domain of the kth finite element. The radial position of the load vector measured from the horizontal direction (point O_j is on the $\overline{X}$ axis) is given by θ_j

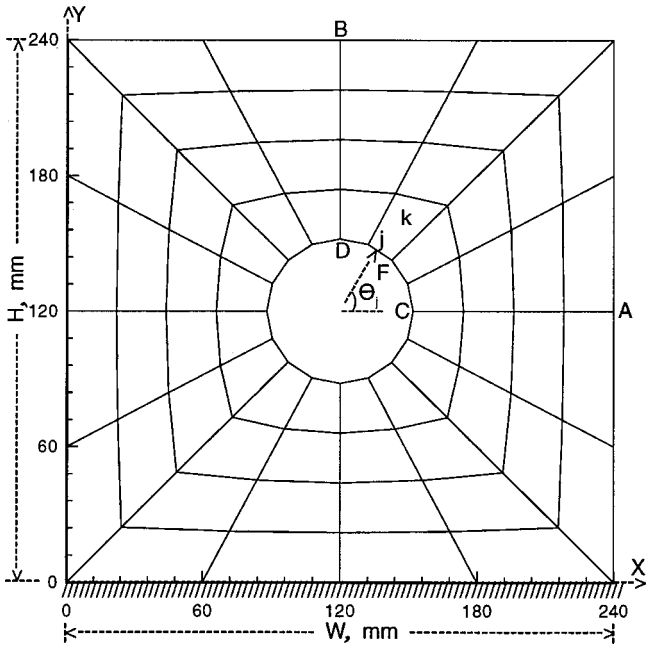


FIG. 1. FE discretization of the cantilever plate.

$=\Omega_c t$ (rad), where Ω_c is the angular speed (rad/s) of the rotating load vector and t is the global time in seconds measured from the point O_j on the global boundary (the annulus circumference). The plate element has eight nodes with two degrees of freedom each. Recall that the desired accuracy of the solution and idealization of the curved boundary (the circular hole surface in our case) can be achieved by using, for example, a fewer number of eight-node isoparametric elements compared to an increased number of six-node triangle elements with curved sides or vice versa. This is provided the number of degrees of freedom per node in each choice is the same. Thus the selection of the eight-node C° element is justified.

Let $\mathbf{d}_k^e$ and $\mathbf{h}_k^e$ denote, respectively, the global vector of the undeformed state position and the global vector of the elastic deformations of a generic point within the spatial domain of the kth element, such that

$$\mathbf{d}_k^e=[x \quad y]^T, \quad \mathbf{h}_k^e=[u \quad v]^T, \tag{1}$$

where the superscript e denotes the element whereas the subscript k stands for its number. x and y are the global positions

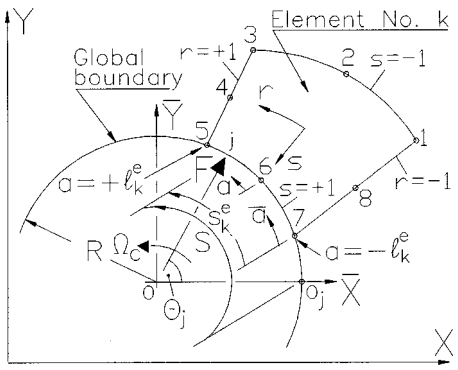


FIG. 2. A global boundary plate finite element subject to an in-plane moving (rotating) load.

along the X and Y directions, respectively. u and v are the global elastic displacements along the X and Y, respectively. Let $\mathbf{z}_k^e$ stands for the element 16×1 nodal points global coordinates vector and $\mathbf{q}_k^e$ refers to the kth element global displacement field vector, such that

$$\mathbf{z}_k^e=[x_j \quad y_j]^T, \quad \mathbf{q}_k^e=[u_j \quad v_j]^T, \quad j=1,2,..,8. \tag{2}$$

The element global coordinate vector, $\mathbf{z}_k^e$, and the elastic deformations vector, $\mathbf{h}_k^e$, are related, respectively, to the vector of the element nodal points coordinates and the vector of the element nodal points variables, such that

$$\mathbf{d}_k^e=\mathbf{N}\mathbf{z}_k^e, \quad \mathbf{h}_k^e=\mathbf{N}\mathbf{q}_k^e, \quad \mathbf{N}=\begin{bmatrix} N_j & 0 \\ 0 & N_j \end{bmatrix}, \quad j=1,2,..,8. \tag{3}$$

$\mathbf{N}$ is the 2×16 overall matrix of the element C° shape functions with its enteries (needed for the analysis of the next section) are

$$\begin{bmatrix} N_1 \\ N_2 \\ N_3 \\ N_4 \\ N_5 \\ N_6 \\ N_7 \\ N_8 \end{bmatrix} = \begin{bmatrix} \frac{-1}{4}(1-r)(1-s)(1+r+s) \\ \frac{1}{2}(1-r^2)(1-s) \\ \frac{-1}{4}(1+r)(1-s)(1-r+s) \\ \frac{1}{2}(1+r)(1-s^2) \\ \frac{-1}{4}(1+r)(1+s)(1-r-s) \\ \frac{1}{2}(1-r^2)(1+s) \\ \frac{-1}{4}(1-r)(1+s)(1+r-s) \\ \frac{1}{2}(1-r)(1-s^2) \end{bmatrix}. \tag{4}$$

r and s are the element serendipity coordinates.

Expressed in the XYZ system, the plate element kinetic energy and the potential energy are $T_k^e=\frac{1}{2}(\dot{\mathbf{q}}_k^e)^T\mathbf{M}_k^e\dot{\mathbf{q}}_k^e$, $P_k^e=\frac{1}{2}(\mathbf{q}_k^e)^T\mathbf{K}_k^e\mathbf{q}_k^e$, respectively, where $\mathbf{M}_k^e$ and $\mathbf{K}_k^e$ are the element consistent mass and stiffness matrices, respectively, for linearly elastic isotropic material for plane stress.

A. Plate finite-element load vector due to in-plane moving load

Figure 2 depicts a typical plate finite element (number k) sharing the global boundary (the hole circumference) under the action of an in-plane moving load vector, where $\overline{a}$ is an element physical coordinate with its origin at node number 7 and is measured along the global boundary. Also, let us introduce another physical coordinate a that runs along the annulus circumference with its origin at the element midspan point along global boundary (node number 6). The element physical length along the global boundary is $2l_k^e$ such that $a=+l_k^e$ at node number 5 and $a=-l_k^e$ at node number 7. The global dynamic load vector, $\mathbf{F}$, has two components along the X and Y axes, F_x and F_y, respectively,

$$\mathbf{F}=[F_x \quad F_y]^T=F_0[\cos\Omega t \quad \sin\Omega t]^T. \tag{5}$$

$F_0=\overline{F}_0 t_k^e$ is a constant in which $\overline{F}_0$ is a force per unit plate thickness and t_k^e is the plate thickness. The instantenous global position of the moving load along the global boundary is determined by a global distance, $S=Vt$, measured from the instant $\mathbf{F}$ starts rotation at the point O_j. $V=\Omega_c R$ is the circumferential speed of the moving load where R is the radius of the annulus and t is the global time coordinate with its

origin at O_j. While the vector $\mathbf{F}$ is within the kth element global boundary, its position relative to node number 7 is measured by a local distance, $s_k^e = V\overline{t}$, such that $\overline{t}$ is a local (an element) time coordinate measured from the instant the moving (rotating) load vector reaches node number 7. This local measure and the global measure, S, are related, such that

$$s_k^e = S - 2(k-1)l_k^e. \tag{6}$$

k is the element number. It should be noted, however, that both S and s_k^e are measured along the same global boundary. Expressed in the XYZ frame, the element global external load vector due to the moving load vector is given by $\mathbf{F}_k^e$, such that

$$\mathbf{F}_k^e = \begin{cases} \mathbf{F}\delta(\overline{a} - V\overline{t}) & \text{for } 2(k-1)l_k^e < S < 2kl_k^e \\ \underline{0}_2 & \text{otherwise} \end{cases}. \tag{7}$$

Here, $\underline{0}_2$ is a 2×1 null vector. $\delta(\overline{a} - V\overline{t})$ is the dimensional Dirac delta function with both $\overline{a}$ and $V\overline{t}$ as defined earlier. Since the units of this delta function are those of $1/\overline{a}$, vector $\mathbf{F}_k^e$ may be interpreted as a load per unit length [Eq. (10)]. To rewrite $\delta(\overline{a} - V\overline{t})$ in terms of the coordinates (r,s), let us introduce a new nondimensional coordinate, $\overline{g}$ (called hereafter the element global boundary natural coordinate):

$$\overline{g} = \frac{a}{l_k^e}. \tag{8}$$

$\overline{g}$, whose origin (not shown) is at node number 6, changes its value in a linear fashion from -1 at node number 7 to $+1$ at node number 5 along the global boundary.

From Fig. 2, the relation between the kth element two physical coordinates $\overline{a}$ and a is $\overline{a} = a + l_k^e$. Thus $\delta(\overline{a} - V\overline{t}) = \overline{\delta}(a + l_k^e - V\overline{t})$ which upon using Eq. (8), and utilizing the generalized delta function properties, we get

$$\delta(\overline{a} - V\overline{t}) = \frac{1}{l_k^e}\overline{\delta}\left(\overline{g} - \left(-1 + \frac{V\overline{t}}{l_k^e}\right)\right). \tag{9}$$

$\overline{\delta}(.)$ is the nondimensional delta function. The term $(-1 + V\overline{t}/l_k^e)$ assumes a value of -1 when the load arrives at node 7 and a value of $+1$ when the load departs at node number 5. Thus this time-dependent, and linearly varying, term measures the change in the coordinate, $\overline{g}$, as load progresses (rotates). Substitute $V\overline{t} = s_k^e$ into the right hand side of Eq. (9) and then into Eq. (7), and we get

$$\mathbf{F}_k^e = \begin{cases} \mathbf{F}\dfrac{1}{l_k^e}\overline{\delta}\left(\overline{g} - \left(\dfrac{s_k^e}{l_k^e} - 1\right)\right) & \text{for } 2(k-1)l_k^e < S < 2kl_k^e \\ \underline{0}_2 & \text{otherwise} \end{cases}. \tag{10}$$

Recall that $\mathbf{h}_k^e = \mathbf{N}\mathbf{q}_k^e$, the virtual change in the displacement field $\mathbf{h}_k^e$, $\delta\mathbf{h}_k^e$, is

$$\delta\mathbf{h}_k^e = [\,\delta u \quad \delta v\,]^T = \mathbf{N}\delta\mathbf{q}_k^e. \tag{11}$$

$\delta\mathbf{q}_k^e$ is the virtual change in $\mathbf{q}_k^e$. The virtual work of vector $\mathbf{F}_k^e$ is given by δW_k^e:

$$\delta W_k^e = \int_{-l_k^e}^{l_k^e} (\delta\mathbf{h}_k^e)^T \mathbf{F}_k^e \, da. \tag{12}$$

In this global boundary line integral, da (not shown) is the differential (infinitesimal) arc length on the global boundary. da may be computed using the global components of the infinitesimal chord length approximating it, dx and dy, in the X and Y directions, respectively, and the Pathagorean theorem. To this end

$$da = \sqrt{(dx)^2 + (dy)^2}. \tag{13}$$

x and y are the global coordinates of a general point on the element global boundary. From Eqs. (1)–(3), $[x\,y]^T = \mathbf{N}[x_j\,y_j]^T$. Recall that $\mathbf{N} = \mathbf{N}(r,s)$ and use the chain rule of differentiations, the total differentials dx and dy are

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = \sum_{j=1}^{Nen} \begin{bmatrix} \dfrac{\partial N_j}{\partial r}x_j & \dfrac{\partial N_j}{\partial s}x_j \\ \dfrac{\partial N_j}{\partial r}y_j & \dfrac{\partial N_j}{\partial s}y_j \end{bmatrix} \begin{bmatrix} dr \\ ds \end{bmatrix}. \tag{14}$$

Nen is the element number of nodes. Since on the global boundary $s = +1$ and r is a spatial variable, $ds = 0$. Now, from Eq. (14) into (13), then

$$da = R_k^e(r,1)\,dr, \tag{15}$$

$$R_k^e(r,1) = \sqrt{\left(\sum_{j=1}^{Nen} \frac{\partial N_j(r,1)}{\partial r}x_j\right)^2 + \left(\sum_{j=1}^{Nen} \frac{\partial N_j(r,1)}{\partial r}y_j\right)^2}.$$

In Eqs. (15), the partial derivatives of the element shape matrix enteries are computed along the global boundary where $s = +1$ and r changes its value in a linear fashion as explained above. The elements, N_j, in Eqs. (4), show that only the components $N_5 - N_7$ will contribute to the right hand side of R_k^e's equation. And this not only saves CPU time but also serves as an essential pivot in deriving Eq. (21) as follows.

Use Eq. (11), then $(\delta\mathbf{h}_k^e)^T = (\delta\mathbf{q}_k^e)^T(\mathbf{N})^T$ and recall that $\mathbf{N} = \mathbf{N}(r,s)$ should be replaced by $\mathbf{N}(r,1)$ and then substitute from Eqs. (15) into (12), and since $\delta\mathbf{q}_k^e$ constitutes a set of only time dependent variables, we get

$$\delta W_k^e = (\delta\mathbf{q}_k^e)^T \int_{-1}^{1} (\mathbf{N}(r,1))^T \mathbf{F}_k^e R_k^e(r,1)\,dr. \tag{16}$$

Substitute from Eq. (10) into (16) and adjust the vector dimensions, then in a compact form we have

$$\delta W_k^e = (\delta\mathbf{q}_k^e)^T \mathbf{Q}_k^e, \tag{17}$$

such that

$$\mathbf{Q}_k^e = \begin{cases} \displaystyle\int_{-1}^{1} \mathbf{Z}_1(r)\,dr & \text{for } 2(k-1)l_k^e < S < 2kl_k^e \\ \underline{0}_{16} & \text{otherwise} \end{cases}, \tag{18}$$

$$\mathbf{Z}_1(r) = R_k^e(r,1)(\mathbf{N}(r,1))^T \frac{\mathbf{F}}{l_k^e}\overline{\delta}\left(\overline{g} - \left(\frac{s_k^e}{l_k^e} - 1\right)\right).$$

$\mathbf{Q}_k^e$ is the element global generalized load vector associated with the global generalized deformation vector $\mathbf{q}_k^e$, where $\underline{0}_{16}$ is a 16×1 null vector. The integral in Eq. (18) in its current form is not amenable to evaluation because the delta function is expressed in terms of the coordinate, $\overline{g}$, whereas the integration is with respect to r. However, this hurdle can be removed as follows. As it has been pointed out previously, the element nodal points numbers 1–4 and 8 are free from any direct influence of the external applied load vector and this makes the coordinate r only affect the external forces applied to nodal points 5–7. That is, the applied forces at the element global boundary nodes are not explicit functions of any external forces that may exist at any of the other five nodal points and the effect of r on the element external forces is limited to the linear variation in its value between nodes number 7 and 5 on the global boundary. The same effect of the linear variation in value, along the global boundary, is a requirement for the coordinate $\overline{g}$, as well, as can be viewed from the above presentation. In other words, both the r and $\overline{g}$ coordinates are equivalent in their effect on the computation of the element load vector. Therefore, in the above integral, $\overline{g}$ may be replaced by r and the reverse is not allowed. To this end, Eq. (18) can be rewritten as

$$\mathbf{Q}_k^e=\begin{cases}\displaystyle\int_{-1}^{1}\mathbf{Z}_2(r)\,dr & \text{for } 2(k-1)l_k^e<S<2kl_k^e\\ \underline{0}_{16} & \text{otherwise}\end{cases}, \tag{19}$$

where the time dependent function $\mathbf{Z}_2(r)$ is the same as $\mathbf{Z}_1(r)$ [Eq. (18)] except that $\overline{g}$ is replaced by r. Now, utilizing the properties of the Dirac delta function, Eq. (19) becomes

$$\mathbf{Q}_k^e=\begin{cases}\mathbf{H}_k^e(s_k^e)\dfrac{F}{l_k^e}, & \text{for } 2(k-1)l_k^e<S<2kl_k^e\\ \underline{0}_{16}, & \text{otherwise}\end{cases}, \tag{20}$$

$$\mathbf{H}_k^e(s_k^e)=R_k^e\left(\frac{s_k^e}{l_k^e}-1,1\right)\left(\mathbf{N}\left(\frac{s_k^e}{l_k^e}-1,1\right)\right)^T.$$

Thus the element global generalized load vector, $\mathbf{Q}_k^e$, associated with the generalized displacement vector, $\mathbf{q}_k^e$, can be evaluated at any instantaneous position along the element global boundary by evaluating the matrix function $\mathbf{H}_k^e(s_k^e)$ at the corresponding instantaneous and nondimensional position $(-1+s_k^e/l_k^e)$, and then multiplying it by the global load vector $\mathbf{F}$, divided by the element half-length along global boundary, l_k^e. However, if nodal point number 6 is located at the midspan point of the element global boundary, which is the assumption from the start of the analysis, the function $R_k^e(-1+s_k^e/l_k^e,1)$ will always be equal to l_k^e. That is, $\mathbf{Q}_k^e$ is given by

$$\mathbf{Q}_k^e=\begin{cases}\left(\mathbf{N}\left(\dfrac{s_k^e}{l_k^e}-1,1\right)\right)^T\mathbf{F}, & \text{for } 2(k-1)l_k^e<S<2kl_k^e\\ \underline{0}_{16}, & \text{otherwise}\end{cases}. \tag{21}$$

However, in this case, Eq. (20) can be used to check the success of the formulation computer implementation. That is, a successful digital implementation should always yield the instantenous value of the function $R_k^e[-1+(s_k^e/l_k^e),1]$ equal to l_k^e while the load vector is traveling along the kth element global boundary.

Let $[u_d\, v_d]^T$ denotes the global displacements vector of the point under the moving load. This deformation field vector can be obtained by specializing $\mathbf{h}_k^e$ of Eq. (3) to the rolling point on the element global boundary domain, such that

$$\begin{bmatrix}u_d\\ v_d\end{bmatrix}=\mathbf{N}\left(\frac{s_k^e}{l_k^e}-1,1\right)\mathbf{q}_k^e, \tag{22}$$

where u_d and v_d are the global motions along the X and Y axes, respectively. Similarily, if $[x_d\, y_d]^T$ denotes the global coordinates vector of the point under the moving load, then $\mathbf{d}_k^e$ of Eq. (3) gives

$$\begin{bmatrix}x_d\\ y_d\end{bmatrix}=\mathbf{N}\left(\frac{s_k^e}{l_k^e}-1,1\right)\mathbf{z}_k^e. \tag{23}$$

If the element nodal points and its serendipity coordinates (r,s) are arranged in such a manner that on the global boundary the coordinate r has a constant value whereas s changes its value in a linear fashion, the right hand side of Eq. (21) has to be changed accordingly and the ensuing necessary changes can easily be identified. Although presenting the final equations necessary for producing the reported results, namely Eqs. (21)–(23), may be sufficient for some readers. However, for others with an interest to extend the work further to, for example, plates under a 3-D load vector, etc., analyses leading to these equations may be of interest.

B. Equations of motion and solution scheme

Use the element energy expressions and apply Lagrange's equations, the element equations of motion are $\mathbf{M}_k^e\ddot{\mathbf{q}}_k^e+\mathbf{K}_k^e\mathbf{q}_k^e=\mathbf{Q}_k^e$. Let the system overall global mass and stiffness matrices be $\mathbf{M}$ and $\mathbf{K}$, respectively, and $\mathbf{Q}$ be the time-dependent overall global load vector, each obtained by assembling the contribution from the individual elements using the 1-D array scheme for banded/sparse matrices.[11] Let $\mathbf{q}$ denote the system overall global displacement field, and the global system of equations of motion is $\mathbf{M}\ddot{\mathbf{q}}+\mathbf{K}\mathbf{q}=\mathbf{Q}$. Premultiply by $(\mathbf{\Phi}_N)^T$ and use the transformations $[\mathbf{q}\,\dot{\mathbf{q}}\,\ddot{\mathbf{q}}]^T=\mathbf{\Phi}_N[\eta\,\dot{\eta}\,\ddot{\eta}]^T$, where η is the modal displacement vector and $\mathbf{\Phi}_N$ is the normalized (with respect to $\mathbf{M}$) modal matrix obtained by solving the eigenvalue problem, we get $\ddot{\eta}+\underline{\omega}^2\eta=\overline{\mathbf{F}}$, where $\overline{\mathbf{F}}=(\mathbf{\Phi}_N)^T\mathbf{Q}$ is the load vector expressed in the normal coordinates system and $\underline{\omega}^2$ is the spectral matrix. The scheme used to solve the above uncoupled system is

$$\eta_s(t_j+\Delta t)=\left[\eta_s(t_j)-\frac{\overline{F}_s(t_j)}{(\omega_s)^2}\right]\cos(\omega_s\Delta t)+\frac{\dot{\eta}_s(t_j)}{\omega_s}\sin(\omega_s\Delta t)+\frac{\overline{F}_s(t_j)}{(\omega_s)^2},$$

TABLE I. Natural frequencies used in the modal analysis, Hz.

$(\overline{f}_1, \overline{f}_2) = (71.242, 182.858)$
$(\overline{f}_3, \overline{f}_4) = (210.097, 290.553)$
$(\overline{f}_5, \overline{f}_6) = (301.367, 361.885)$
$(\overline{f}_7, \overline{f}_8) = (451.762, 472.995)$
$(\overline{f}_9, \overline{f}_{10}) = (555.931, 572.175)$
$(\overline{f}_{11}, \overline{f}_{12}) = (594.796, 601.397)$
$(\overline{f}_{13}, \overline{f}_{14}) = (670.564, 685.403)$
$(\overline{f}_{15}, \overline{f}_{16}) = (751.222, 775.623)$

$$\dot{\eta}_s(t_j+\Delta t) = \left[\frac{\overline{F}_s(t_j)}{(\omega_s)} - \omega_s \eta_s(t_j)\right]\sin(\omega_s \Delta t) + \dot{\eta}_s(t_j)\cos(\omega_s \Delta t),$$
$$\ddot{\eta}_s(t_j+\Delta t) = \overline{F}_s(t_j) - (\omega_s)^2 \eta_s(t_j+\Delta t), \qquad (24)$$
$$s = 1,2,3,.....,Nf.$$

ω_s and Nf are the circular natural frequency and number of retained modes, respectively. The $\eta_s(.)$ and $\dot{\eta}_s(.)$ expressions have been presented in Ref. 12 without derivation. In their derivation,[8] the force $\overline{F}_s(t_j)$ is assumed to be constant during time interval $t_j \leqslant t \leqslant t_{j+1}$ (i.e., time step has to be small). In this work, $\Delta t = 2\times 10^{-7}$ s.

III. RESULTS AND DISCUSSION

The input data are: $F_0 = 500$ N, $H = W = 240$ mm, $E^e = 200$ GPa, $\rho^e = 7500$ kg/m^3, $\nu^e = 0.3$, $\Omega_c = 25$, 100, 400, and 1000 rad/s. In the FE discretization (Fig. 1), the element midspan nodal points (numbers 2, 4, 6, and 8) are masked. The eigenvalue problem is solved using the subspace iteration method[11] for different numbers of elements and samples of the results are $(\overline{f}_1^{32}, \overline{f}_1^{48}, \overline{f}_1^{64}) = (71.595, 71.525, 71.242)$, $(\overline{f}_2^{32}, \overline{f}_2^{48}, \overline{f}_2^{64}) = (183.954, 183.803, 182.858)$, $(\overline{f}_3^{32}, \overline{f}_3^{48}, \overline{f}_3^{64}) = (210.620, 210.469, 210.097)$, $(\overline{f}_{12}^{32}, \overline{f}_{12}^{48}, \overline{f}_{12}^{64}) = (603.498, 602.532, 601.397)$ Hz, where $\overline{f}$ is the natural frequency and superscript denotes mesh number of elements used. For accurate mapping of the casing curved global boundary (annulus surface), we used the 64 elements discretization. Table I tabulates the first 16 natural tones used in the modal analysis.

Figure 3 depicts the global deformations of the plate global boundary circular surface, in the form of a polar plot, under the moving (rotating) load for different values of the rotational speed w_c (Ω_c in the formulation). Here, x_d and y_d are the X and Y global positions of the moving point around the circumference, computed using Eq. (23), and u_d and v_d are their counterparts global deformations, calculated using Eq. (22). The surface $[x_d, y_d]$, which corresponds to the global boundary configuration when $\Omega_c = 0$ is a circle of radius $R = 32$ mm, and this demonstrates success of the formulation computer implementation. Both u_d and v_d are multiplied by a factor $\alpha = 10^4$ and then superimposed on the corresponding instantenous positions x_d and y_d, respectively, and plotted as $(x_d + \alpha u_d)$ vs $(y_d + \alpha v_d)$. Recall that the load starts and terminates rotation at point C, we remark that the deformable

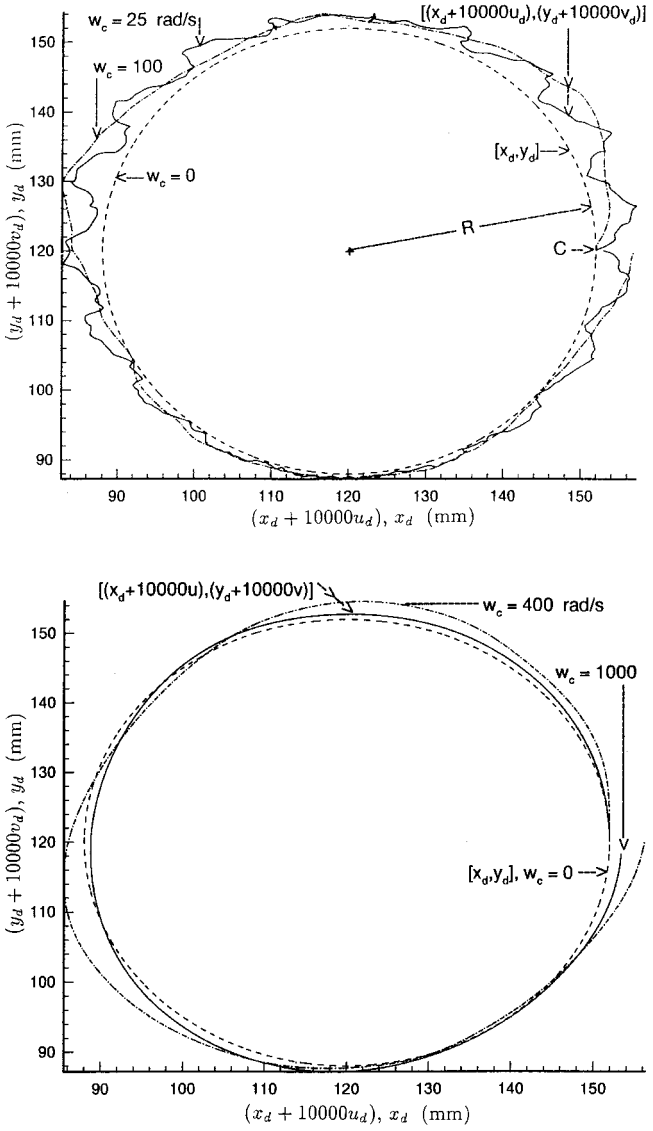


FIG. 3. Plots $(x_d + \alpha u_d)$ vs $(y_d + \alpha v_d)$, x_d vs y_d ($\alpha = 10\,000$), ($\Omega_c = 0{,}25$, 100, 400, 1000 rad/s).

surface, $[(x_d + \alpha u_d), (y_d + \alpha v_d)]$, does not close onto itself with the minimum deviation at the closing point (point C) corresponds to the plot for $\Omega_c = 25$ rad/s. Also, this speed produces the overall largest elastic deformations, under the moving load vector $\mathbf{F}$, with a maximum value of about 0.6 μm. As the speed of rotation increases, the deformation amplitude decreases and the surface becomes more smoother. This is because at lower speeds the elastic motions have enough time to build up. With $\Omega_c = 25{,}100$ rad/s, the deformation amplitudes are always positive (contraction) with the region of the overall minimum amplitudes stacked in the neighborhood of $\theta_j = 0{,}180°$. And with $\Omega_c = 400{,}1000$ rad/s, the elastic deformable line (surface) expands and crosses the line of zero deformation [surface $[x_d, y_d]$] with the region of minimum amplitudes slightly shifted above $\theta_j = 180°$. As we see, higher rotational speeds reduce deformation and increase smoothness of the annulus surface.

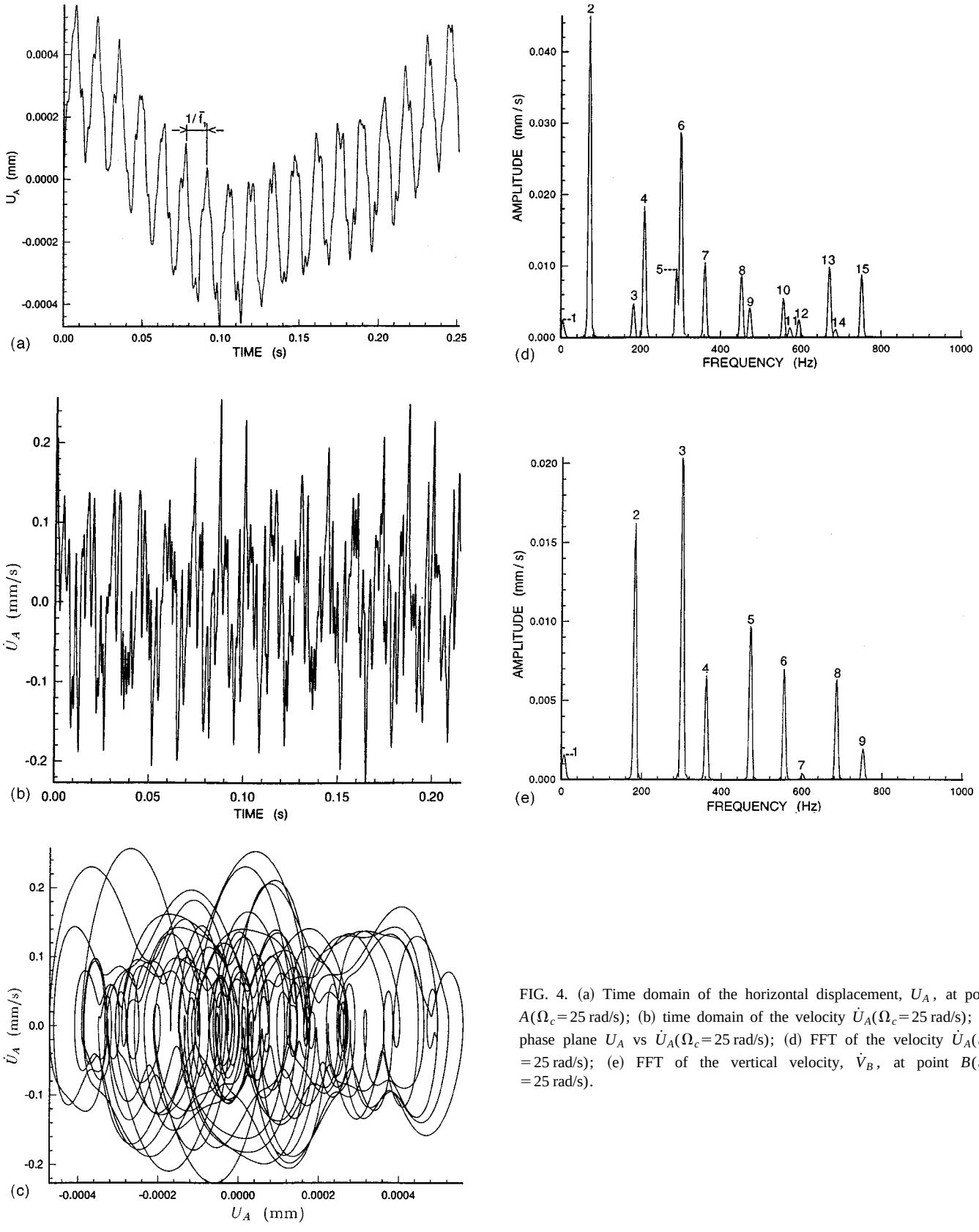


FIG. 4. (a) Time domain of the horizontal displacement, U_A, at point $A(\Omega_c = 25\ \text{rad/s})$; (b) time domain of the velocity $\dot{U}_A(\Omega_c = 25\ \text{rad/s})$; (c) phase plane U_A vs $\dot{U}_A(\Omega_c = 25\ \text{rad/s})$; (d) FFT of the velocity $\dot{U}_A(\Omega_c = 25\ \text{rad/s})$; (e) FFT of the vertical velocity, $\dot{V}_B$, at point $B(\Omega_c = 25\ \text{rad/s})$.

Prior to FFT computations, the time record is zero meaned and windowed using a Hanning window (information about the mathematics of frequency domain analysis using FFT may be found in Ref. 13). The frequency resolution used in Figs. 4(d)–(e) is 2.384 186 Hz. For $\Omega_c = 25\ \text{rad/s}$ ($f_c = 3.9773\ \text{Hz}$), Figs. 4(a) and (b) depict time domains of the displacement, U_A, and the velocity, $\dot{U}_A$, of point A on the casing outer surface, respectively, measured in the XYZ system. Figure 4(c) illustrates the corresponding phase plane where the response chaos can be seen. In Fig.

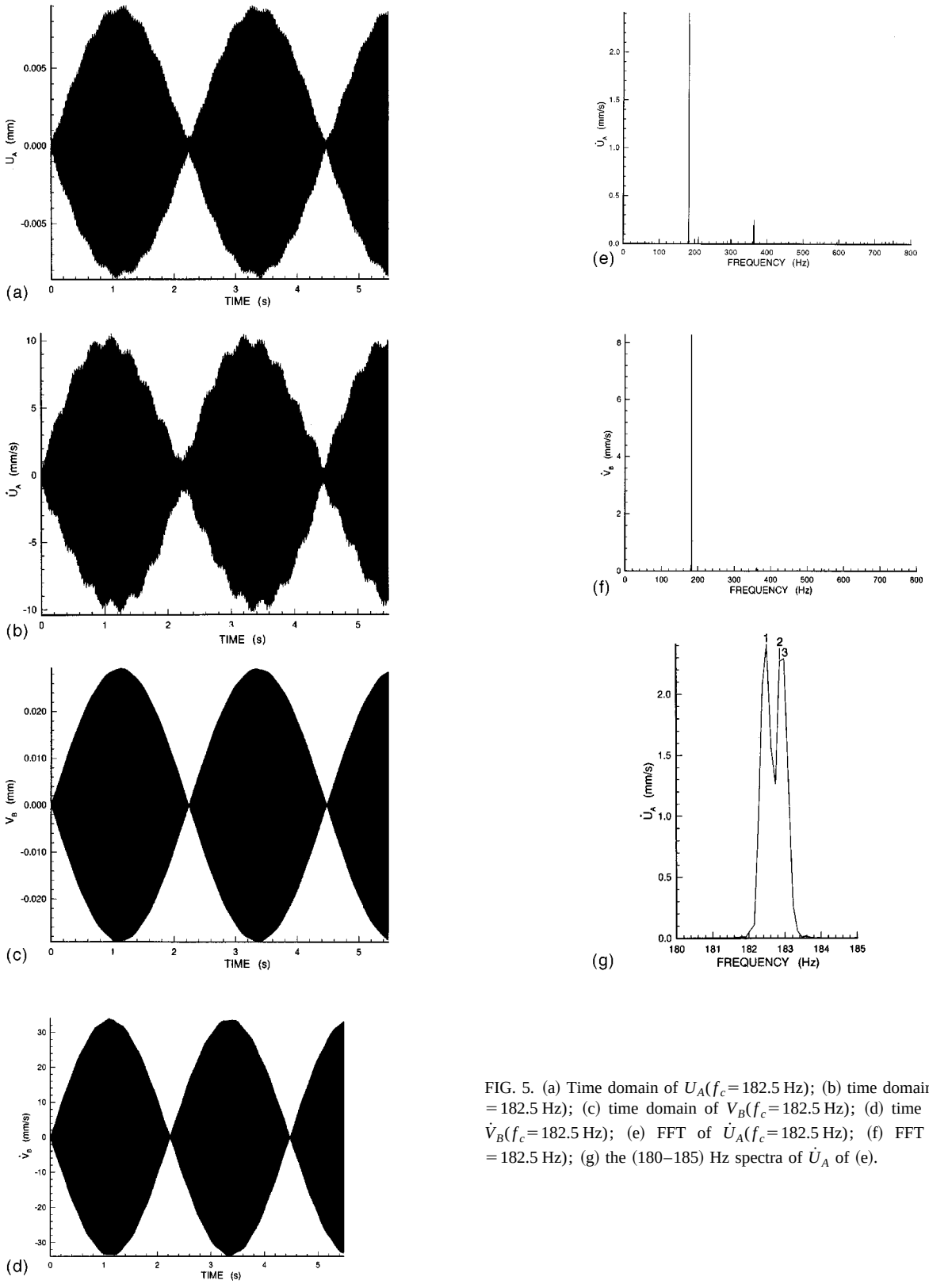


FIG. 5. (a) Time domain of $U_A(f_c=182.5\,\text{Hz})$; (b) time domain of $\dot{U}_A(f_c=182.5\,\text{Hz})$; (c) time domain of $V_B(f_c=182.5\,\text{Hz})$; (d) time domain of $\dot{V}_B(f_c=182.5\,\text{Hz})$; (e) FFT of $\dot{U}_A(f_c=182.5\,\text{Hz})$; (f) FFT of $\dot{V}_B(f_c=182.5\,\text{Hz})$; (g) the (180–185) Hz spectra of $\dot{U}_A$ of (e).

4(a), the external excitation periodic time, $\tau_c=1/f_c=0.2514\,\text{s}$, comprises 18 oscillations of period $1/\overline{f_1}\,\text{s}$ each, where $\overline{f_1}=71.242\,\text{Hz}$ is the casing first natural frequency (Table I). Each of these oscillations is modulated apparently by higher natural frequencies. For the sake of brevity no other time domain plots are included. The spectra of $\dot{U}_A$ and $\dot{V}_B$ are given in Figs. 4(d) and (e), respectively, where the horizontal response is of a higher energy content compared to that measured at point *B*. This is because the cantilever rigid support is along the *X* axis and f_c is closer to $\overline{f_1}$, the first natural tone in the *X* direction. This is supplemented by the fact that the force $F_x=F_0\cos\Omega_c t$ has its maximum

value instantly at the start of rotation, whereas F_y commences with a zero value (i.e., the horizontal motion has an ample time to build up). In each of Figs. 4(d) and (e), f_1 is attributed to f_c. The dominant tone in the X the direction is $f_2 = 71.526\,\text{Hz} = \overline{f_1}$ and along the Y direction, $f_3 = 300.41$ Hz $= \overline{f_5}$, which in turn coincides with f_6 in Fig. 4(d). In Fig. 4(d), we also have $(f_3, f_4) = (183.58, 209.81) = (\overline{f_2}, \overline{f_3})$, $(f_5, f_7) = (290.87, 362.37) = (\overline{f_4}, \overline{f_6})$, $(f_8, f_9) = (450.61, 472.07) = (\overline{f_7}, \overline{f_8})$, $(f_{10}, f_{11}) = (555.51, 572.20) = (\overline{f_9}, \overline{f_{10}})$, $(f_{12}, f_{13}) = (593.66, 669.96) = (\overline{f_{11}}, \overline{f_{13}})$, $(f_{14}, f_{15}) = (686.65, 751.02) = (\overline{f_{14}}, \overline{f_{15}})$ Hz. In Fig. 4(e), $f_2 = \overline{f_2}$ and $f_4 - f_9$ are due to $\overline{f_6}$, $\overline{f_8}$, $\overline{f_9}$, $\overline{f_{12}}$, $\overline{f_{14}}$, and $\overline{f_{15}}$, respectively.

In the field, a sudden (unexpected) change in the rotor rotational speed can relocate the supporting ball bearing outer ring ball passing frequency (number of balls $\times$bearing cage frequency) close to/at one of the casing natural tones. In the following example we study housing response under such possible adverse operating conditions. We select $f_c = 182.5$ Hz close, from below, to $\overline{f_2} = 182.858$ Hz (Table I). Although further disturbance in rotor speed can lead to catastrophic failure, the information provided by this example is useful in supporting the conclusions provided from the spectra discussed above; the casing rigid support along the X axis entices more natural tones to participate (i.e., to be visible) in the X direction spectrum compared to the spectrum along the Y direction. In other words, the revelations of Fig. 4 underline the importance of the vibration measuring probe(s) orientation(s) with respect to the X and the Y directions, to avoid missing threatening vibration peaks. Since f_c is high, to reduce CPU time we solved for 10^3 cycles and used a frequency resolution of 0.198 682 15 Hz in the frequency domain analysis. The results in time domain are shown in Figs. 5(a)–(d) for U_A, $\dot{U}_A$, V_B, and $\dot{V}_B$, respectively, where we see more riggles in the response of point A compared to the one at point B. This is due to effect of the cantilever rigid support along the X axis, as explained above. Also both signals remind us of a vibration signal with beats. Because the forcing frequency is close to the first natural tone in the vertical direction $(\overline{f_2})$, the vibrations at point A are of higher amplitudes. The spectra of $\dot{U}_A$ and $\dot{V}_B$ are shown in Figs. 5(e) and (f), respectively. In each spectrum we have a main spectral line around forcing frequency. When the 5 Hz including and around this spike are shown on an expanded scale [see Fig. 5(g) for $\dot{U}_A$] we find two peaks $f_1 = f_c$ and $(f_2 + f_3)/2 = \overline{f_2}$. It is expected that a further increase in the frequency resolution will separate them further. The second harmonic of each of these two peaks is visible in Fig. 5(e) and not in Fig. 5(f).

IV. CONCLUSIONS

Analytical procedures, using an isoparametric FE formulation, are presented for the vibration analysis of a cantilever plate with a central circular hole (annulus) subject to an in-plane moving (rotating) load vector. The analyses are implemented in the FE program DAMRO 1 and used to highlight the effect of the load rotational speed on the annulus surface elastic deformations under moving load vector. And it is concluded that: (1) Higher load rotational speeds produce less deformations and increase the smoothness of the annulus surface; (2) The annulus dynamic deformations associated with lower speeds ($\Omega_c = 25{,}100$ rad/s) are always positive (contraction). And at higher speeds (400, 1000 rad/s), these deformations exhibit both contraction and expansion around the annulus circumference; (3) The vibration measuring probe(s) orientation with respect to the casing rigid support direction should be selected carefully to avoid missing catastrophic failure enducing vibration peaks.

[1] L. Fryba, *Vibration of Solids and Structures Under Moving Load* (Noordhoff, Leiden, 1972).
[2] M. R. Taheri and E. C. Ting, ''Dynamic response of plate to moving loads: Finite element method,'' Comput. Struct. **34**, 509–521 (1988).
[3] A. W. Leissa and E. F. Ayoup, ''Vibration and buckling of a simply supported rectangular plate subjected to a pair of in-plane concentrated forces,'' J. Sound Vib. **127**, 155–171 (1988).
[4] P. A. Laura and R. H. Gutherrez, ''Transverse vibrations of a thin rectangular plate subjected to a non-uniform stress distribution field,'' J. Sound Vib. **210**, 559–565 (1998).
[5] T. Y. Yang, *Finite Element Structural Analysis* (Prentice-Hall, New York, 1988).
[6] M. M. A. Taha and J. R. Crookall, ''Rolling bearings for machine tools—Comparative evaluation by a new experimental technique and by finite element analysis,'' Int. J. Mach. Tools Des. Res. **17**, 179–190 (1977).
[7] F. K. Choy, Y. F. Ruan, Y. K. Tu, J. J. Zakrajsek, and D. P. Townsend, ''Modal analysis of multiple gear systems coupled with gearbox vibrations: Actuators incorporating flexible casing effects,'' ASME J. Mech. Des. **114**, 486–497 (1992).
[8] F. M. A. El-Saeidy, ''Finite element modeling of rotor-shaft-rolling bearing systems with consideration of bearing nonlinearities,'' J. Vib. Control **4**, 541–602 (1998).
[9] F. M. A. El-Saeidy, ''Finite element dynamic analysis of a rotating shaft with or without nonlinear boundary conditions subject to a moving load,'' Nonlinear Dyn. **21**, 377–408 (2000).
[10] A. Sankaravelu, S. T. Noah, and C. P. Burger, ''Bifurcation and chaos in ball bearings,'' *ASME Winter Annual Meeting, Symposium on Stochastic and Nonlinear Dynamics, AMD-Vol. 192, DE-Vol. 78*, Chicago, pp. 313–325 (1994).
[11] F. M. A. El-Saeidy, DAMRO 1: *A General Purpose Finite Element Program*, 1993.
[12] M. L. Adams, ''Nonlinear dynamics of flexible multi-bearing rotors,'' J. Sound Vib. **71**, 129–144 (1980).
[13] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1975).

Vibration of annular sector plates from three-dimensional analysis

K. M. Liew[a)]
Centre for Advanced Numerical Engineering Simulations, School of Mechanical and Production Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798

T. Y. Ng
Institute of High Performance Computing, 89C Science Park Drive, #02-11/12, The Rutherford, Singapore Science Park 1, Singapore 118261

B. P. Wang
Department of Mechanical Engineering, University of Texas, Arlington, Texas 76019

(Received 7 September 2000; revised 7 April 2001; accepted 14 April 2001)

Accurate three-dimensional elasticity solutions are presented for the free-vibration analysis of annular sector plates with arbitrary (but continuous) boundary conditions. The Ritz procedure is used to minimize the integral energy functional, which is formulated in a cylindrical polar coordinate system, in the derivation of the governing eigenvalue equation. The vibratory characteristics of annular sector plates with variations in boundary condition, sector angle, and thickness are examined. The accuracy of the method is validated through appropriate convergence and comparison studies. The first known results presented here provide a basis for understanding the physics of vibrations of annular sector plates and can be used to assess numerical results obtained from various refined two-dimensional theories. © *2001 Acoustical Society of America.*
[DOI: 10.1121/1.1377868]

PACS numbers: 43.40.Dx [CBB]

I. INTRODUCTION

Plates form important design components in many engineering applications such as parts found in aerospace, mechanical, and marine structures. These applications are often exposed to severe vibration conditions. The accurate determination of the vibratory characteristics of these plate elements thus becomes an important task for designers and engineers. Vast amounts of literature exist for the free vibration of plates; see Liew *et al.* (1995, 1998). Compared to work done on rectangular plates, much less research emphasis has been given to the free vibration of sector and annular sector plates. Approximate solutions for sector plates based on classical plate theory (CPT) with various edge conditions on the circular and radial boundaries have been previously obtained; see Ben-Amoz (1959); Westmann (1962); Rubin (1975); and Bhattacharya and Bhowmic (1975). Maruyama and Ichinomiya (1981) reported the experimental work on sector plates with fully clamped boundary condition. Narita (1985) presented frequency results based on thin-plate approximations for completely free circular plates. Approximate results were also obtained by Kim and Dickinson (1989a, 1989b) for thin isotropic/composite annular and circular plates. Liew and Lam (1993) presented vibration solutions for thin annular sector plates with arbitrary combinations of boundary conditions, via the Rayleigh–Ritz method. The nature of stationarity of the natural frequencies at the natural modes in the Rayleigh–Ritz method was discussed by Bhat (1997). Also, the recurrence scheme for the generation of two-dimensional boundary characteristic orthogonal polynomials to be used in the Rayleigh–Ritz minimization for the study of plate vibration was elucidated by Bhat *et al.* (1998). Notable recent works on annular plate vibration include Laura *et al.* (1999, 2000).

For the treatment of moderately thick sector plates, refined plate theories that incorporate the transverse shear deformation should be employed. Bapu Rao *et al.* (1977) and Guruswamy and Yang (1979) proposed various Reissner shear deformation sector plate finite elements for approximating the vibration solutions. The finite strip method was applied for the analysis of annular sector Mindlin plates by Mizusawa (1991). Srinivasan and Thiruvenkatachari (1985) formulated solutions for the free vibration of transverse isotropic annular sector Mindlin plates. Using the Ritz method and a unique set of two-dimensional polynomial functions, Xiang *et al.* (1993) studied the free vibrations of moderately thick Mindlin annular sector plates. Natural frequencies for a wide range of sector plates with different boundary conditions were reported. Qatu (1994) examined the validity of various nonlinear shear deformation theories for laminated plates. The superposition-Galerkin method was developed by Gorman (1996, 1997) for obtaining accurate frequency characteristics of Mindlin plates with torsional elastic edge supports.

In contrast to the above studies, which used two-dimensional theories for their analysis, the present paper adopts the three-dimensional elasticity theory to compute the vibration frequencies and mode shapes of annular sector plates. The first known three-dimensional elasticity results for the vibration of thick circular plates can be attributed to

a)Author to whom correspondence should be addressed. Electronic mail: mkmliew@ntu.edu.sg

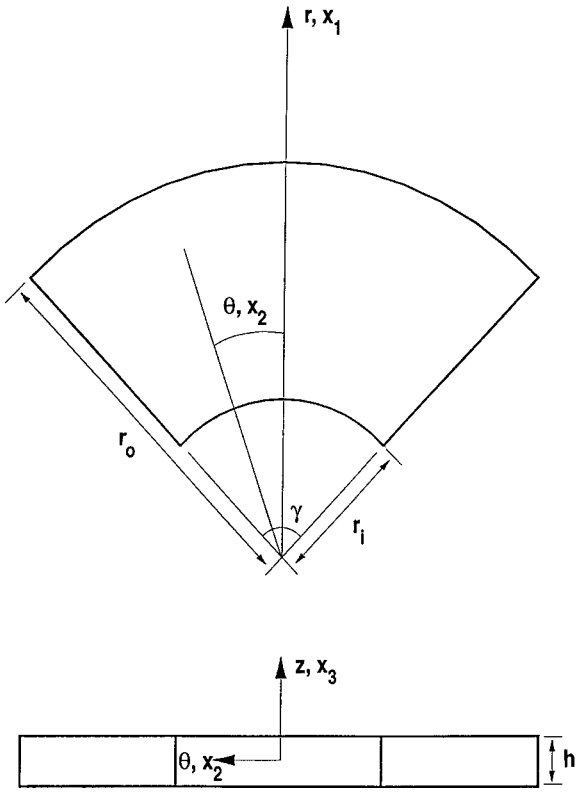


FIG. 1. Geometry and dimensions of an annular sector plate.

Hutchinson (1979, 1984), who employed the Mathieu series in the solution procedure. However, three-dimensional vibration solutions for annular sector plates of arbitrary (but continuous) combination of boundary conditions are not available in the literature, and will be presented here. Analytical solutions for this problem are difficult to obtain and approximate techniques are usually employed for such analysis. In this study, the solutions to the problem are made possible by employing the Ritz method with sets of one- and two-dimensional polynomial functions as trial functions approximating the surface and thickness displacements. Following the standard Ritz minimization procedure, the governing eigenvalue equation is derived. First known vibration solutions in terms of nondimensional frequency parameters and mode shapes are presented for several annular sector plates subject to various combinations of boundary conditions. Needless to say, these three-dimensional elasticity solutions are notably important because they form a real basis for assessing the solutions determined from the two-dimensional plate theories.

II. PROBLEM DEFINITION

The geometric configuration of the homogeneous, isotropic annular sector plate under consideration is depicted in Fig. 1. The plate is of constant thickness h, inner radius r_i, outer radius r_o, and sector angle γ. The plate geometry and dimensions are defined in a cylindrical coordinate system (r,θ,z). The corresponding displacement components at a generic point are u_1, u_2, and u_3 in the radial, circumferential, and thickness directions, respectively. Various combinations of boundary conditions are considered. For example, in this study, SCSF denotes an annular sector plate with hard, simply supported edges at the inner and outer radii, clamped and free boundary conditions at the radial edges at $\theta = \pm \gamma/2$. The vibration frequencies and mode shapes of this plate are to be determined from a three-dimensional displacement-based polynomials-Ritz method.

III. METHOD OF ANALYSIS

The linear elastic strain energy component $\hat{V}$ for a plate in cylindrical coordinates can be written in integral form as

$$\hat{V}=\hat{E}\int_{r_i}^{r_o}\int_{-\gamma/2}^{\gamma/2}\int_{-h/2}^{h/2}[\hat{A}_1^2+(1-2\nu)\times(\hat{A}_2+\tfrac{1}{2}\hat{A}_3)]r\,dr\,d\theta\,dz, \tag{1}$$

where

$$\hat{E}=\frac{E}{2(1+\nu)(1-2\nu)}, \tag{2}$$

$$\hat{A}_1=\varepsilon_{rr}+\varepsilon_{\theta\theta}+\varepsilon_{zz}, \tag{3}$$

$$\hat{A}_2=\varepsilon_{rr}^2+\varepsilon_{\theta\theta}^2+\varepsilon_{zz}^2, \tag{4}$$

$$\hat{A}_3=\varepsilon_{r\theta}^2+\varepsilon_{rz}^2+\varepsilon_{\theta z}^2, \tag{5}$$

and E is the Young's modulus, ν the Poisson's ratio, and the strain components in the cylindrical polar coordinate for small deformation are given as

$$\varepsilon_{rr}=\frac{\partial u_1}{\partial r};\quad \varepsilon_{\theta\theta}=\frac{u_1}{r}+\frac{\partial u_2}{r\partial\theta};\quad \varepsilon_{zz}=\frac{\partial u_3}{\partial z}, \tag{6}$$

$$\varepsilon_{r\theta}=\frac{\partial u_1}{r\partial\theta}+\frac{\partial u_2}{\partial r}-\frac{u_2}{r};\quad \varepsilon_{rz}=\frac{\partial u_1}{\partial z}+\frac{\partial u_3}{\partial r}, \tag{7}$$

$$\varepsilon_{\theta z}=\frac{\partial u_2}{\partial z}+\frac{\partial u_3}{r\partial\theta}. \tag{8}$$

For free-vibration analysis, the kinetic energy $\hat{T}$ can be expressed as

$$\hat{T}=\frac{\rho}{2}\int_{r_i}^{r_o}\int_{-\gamma/2}^{\gamma/2}\int_{-h/2}^{h/2}[\dot{u}_1^2+\dot{u}_2^2+\dot{u}_3^2]r\,dr\,d\theta\,dz \tag{9}$$

$$\dot{u}_1=\frac{\partial u_1}{\partial t};\quad \dot{u}_2=\frac{\partial u_2}{\partial t};\quad \dot{u}_3=\frac{\partial u_3}{\partial t}, \tag{10}$$

where ρ is the mass density per unit volume, and t is time. For linear, small-strain, simple harmonic motion, the displacement components assume the following forms:

$$u_\alpha(r,\theta,z,t)=U_\alpha(r,\theta,z)e^{i\omega t};\quad \alpha=1,2,3, \tag{11}$$

where ω denotes the frequency of vibration.

For simplicity and convenience in mathematical formulation, the cylindrical coordinates (r,θ,z) are transformed into a set of nondimensional parameters $(\overline{x_1},\overline{x_2},\overline{x_3})$ by the following relations:

$$\overline{x_1}=\frac{2r-(r_o+r_i)}{r_o-r_i};\quad \overline{x_2}=\frac{2\theta}{\gamma};\quad \overline{x_3}=\frac{z}{h}. \tag{12}$$

The nondimensional displacement amplitude functions in the radial $U_1(\overline{x_1},\overline{x_2},\overline{x_3})$, circumferential $U_2(\overline{x_1},\overline{x_2},\overline{x_3})$, and thickness $U_3(\overline{x_1},\overline{x_2},\overline{x_3})$ directions are approximated by a combination of one- and two-dimensional orthogonal polynomials in the following forms:

$$U_\alpha(\overline{x_1},\overline{x_2},\overline{x_3})=\sum_{m=1}^{M}\sum_{n=1}^{N} C_{mn}^{\alpha}\phi_m^{\alpha}(\overline{x_1},\overline{x_2})\psi_n^{\alpha}(\overline{x_3});$$
$$\alpha=1,2,3, \tag{13}$$

in which C_{mn}^{α} are the unknown coefficients, and $\phi_m^{\alpha}(\overline{x_1},\overline{x_2})$ and $\psi_n^{\alpha}(\overline{x_3})$ are the one- and two-dimensional orthogonal polynomial functions. It should be noted that the functions $\phi_m^{\alpha}(\overline{x_1},\overline{x_2})$ and $\psi_n^{\alpha}(\overline{x_3})$ for different boundary conditions can be obtained from the earlier references, Xiang *et al.* (1993), Liew and Lam (1993), and Liew *et al.* (1993). Explicitly, these functions are

$$\phi_m^{\alpha}(\overline{x_1},\overline{x_2})=(\overline{x_2}-1)^{\Omega_1}(\overline{x_1}+1)^{\Omega_2}(\overline{x_2}+1)^{\Omega_3}(\overline{x_1}-1)^{\Omega_4}, \tag{14}$$

where $(\overline{x_2}-1)$ is the function for the left straight edge; $(\overline{x_1}+1)$ the function for the inner radius; $(\overline{x_2}+1)$ the function for the right straight edge; $(\overline{x_1}-1)$ the function for the outer radius; and Ω_j, $j=1,2,3,4$, depending on the support edge conditions, is determined by the following:

For the basic function ϕ_m^3

$\Omega_j=0$ if the jth edge is free (F), (15a)

$\Omega_j=1$ if the jth edge is clamped (C) or simply supported (S). (15b)

For the basic function ϕ_m^1

$\Omega_j=0$ if the jth edge is free (F) or simply supported (S) in the θ direction, (16a)

$\Omega_j=1$ if the jth edge is clamped (C) or simply supported (S) in the r direction. (16b)

For the basic function ϕ_m^2

$\Omega_j=0$ if the jth edge is free (F) or simply supported (S) in the r direction, (17a)

$\Omega_j=1$ if the jth edge is clamped (C) or simply supported (S) in the θ direction. (17b)

For antisymmetric thickness modes, the basic functions $\psi_n^{\alpha}(\overline{x_3})$ are

$$\psi_n^1(\overline{x_3})=\psi_n^2(\overline{x_3})=\overline{x_3};\quad \psi_n^3(\overline{x_3})=1. \tag{18a}$$

For symmetric thickness modes, the basic functions $\psi_n^{\alpha}(\overline{x_3})$ are

$$\psi_n^1(\overline{x_3})=\psi_n^2(\overline{x_3})=1;\quad \psi_n^3(\overline{x_3})=\overline{x_3}. \tag{18b}$$

Let $\hat{\Pi}$ be the energy functional given by

$$\hat{\Pi}=\tilde{V}-\tilde{T}, \tag{19}$$

where $\tilde{V}$ and $\tilde{T}$ are the maximum strain and kinetic energies respectively, of the plate which are derived by substituting Eq. (9) into the respective energy expressions in Eqs. (1) and (9) with the periodic component eliminated.

The minimization of the functional in Eq. (19) with respect to the coefficients

$$\frac{\partial\hat{\Pi}}{\partial C_{mn}^{\alpha}}=0;\quad \alpha=1,2,3 \tag{20}$$

leads to the governing eigenvalue equation of the form

$$(\hat{\mathbf{K}}-\hat{\lambda}^2\hat{\mathbf{M}})\hat{\mathbf{C}}=\mathbf{0}, \tag{21}$$

where

$$\hat{\mathbf{K}}=\begin{bmatrix}\hat{\mathbf{k}}^{11} & \hat{\mathbf{k}}^{12} & \hat{\mathbf{k}}^{12}\\ & \hat{\mathbf{k}}^{22} & \hat{\mathbf{k}}^{23}\\ \text{Sym} & & \hat{\mathbf{k}}^{33}\end{bmatrix}, \tag{22}$$

$$\hat{\mathbf{M}}=\begin{bmatrix}\hat{\mathbf{m}}^{\mathbf{11}} & 0 & 0\\ & \hat{\mathbf{m}}^{22} & 0\\ \text{Sym} & & \hat{\mathbf{m}}^{33}\end{bmatrix}, \tag{23}$$

and

$$\hat{\mathbf{C}}=\{\mathbf{C}^1\quad\mathbf{C}^2\quad\mathbf{C}^3\}^T. \tag{24}$$

The explicit form of the respective elements in the stiffness submatrices $\hat{\mathbf{k}}^{\alpha\beta}$ are given by

$$\hat{k}_{mjnk}^{1}=\frac{(1-\nu)r_o}{\tilde{\Lambda}_1}\left[\frac{1}{\hat{\vartheta}_1^2}(\hat{\mathbf{I}}_{mj}^{1010;1})_{11}+(\hat{\mathbf{I}}_{mj}^{0000;-1})_{11}\right](\hat{\mathbf{J}}_{nk}^{00})_{11}$$
$$+\frac{\nu}{\tilde{\Lambda}_1}\left(\frac{r_o}{\hat{\vartheta}_1}\right)[(\hat{\mathbf{I}}_{mj}^{0010;0})_{11}+(\hat{\mathbf{I}}_{mj}^{1000;0})_{11}](\hat{\mathbf{J}}_{nk}^{00})_{11}$$
$$+\frac{r_o}{\tilde{\Lambda}_2}\left[\frac{1}{h^2}(\hat{\mathbf{I}}_{mj}^{0000;1})_{11}(\hat{\mathbf{J}}_{nk}^{11})_{11}\right.$$
$$\left.+\frac{4}{\gamma^2}(\hat{\mathbf{I}}_{mj}^{0101;-1})_{11}(\hat{\mathbf{J}}_{nk}^{00})_{11}\right], \tag{25}$$

$$\hat{k}_{mjnk}^{12}=\frac{(1-\nu)r_o}{\tilde{\Lambda}_1}\left(\frac{2}{\gamma}\right)(\hat{\mathbf{I}}_{mj}^{0001;-1})_{12}(\hat{\mathbf{J}}_{nk}^{00})_{12}+\frac{\nu}{\tilde{\Lambda}_1}\left(\frac{2r_o}{\gamma\hat{\vartheta}_1}\right)$$
$$\times(\hat{\mathbf{I}}_{mj}^{1001;0})_{12}(\hat{\mathbf{J}}_{nk}^{00})_{12}+\frac{1}{\tilde{\Lambda}_2}\left[\frac{2r_o}{\gamma\hat{\vartheta}_1}(\hat{\mathbf{I}}_{mj}^{0110;0})_{12}\right.$$
$$\left.-\frac{2r_o}{\gamma}(\hat{\mathbf{I}}_{mj}^{0100;-1})_{12}\right](\hat{\mathbf{J}}_{nk}^{00})_{12}, \tag{26}$$

$$\hat{k}_{mjnk}^{13}=\frac{\nu}{\tilde{\Lambda}_1}\left[\left(\frac{r_o}{h\hat{\vartheta}_1}\right)(\hat{\mathbf{I}}_{mj}^{1000;1})_{13}+\left(\frac{r_o}{h}\right)(\hat{\mathbf{I}}_{mj}^{0000;0})_{13}\right](\hat{\mathbf{J}}_{nk}^{01})_{13}$$
$$+\frac{1}{\tilde{\Lambda}_2}\left(\frac{r_o}{h\hat{\vartheta}_1}\right)(\hat{\mathbf{I}}_{mj}^{0010;1})_{13}(\hat{\mathbf{J}}_{nk}^{10})_{13}, \tag{27}$$

$$\hat{k}^{22}_{mjnk}=\frac{(1-\nu)r_o}{\tilde{\Lambda}_1}\left(\frac{4}{\gamma^2}\right)(\hat{\mathbf{I}}^{0101;-1}_{mj})_{22}(\hat{\mathbf{J}}^{00}_{nk})_{22}$$
$$+\frac{r_o}{\tilde{\Lambda}_2}\left[\frac{1}{\hat{\vartheta}_1^2}(\hat{\mathbf{I}}^{1010;1}_{mj})_{22}-\frac{1}{\hat{\vartheta}_1}(\hat{\mathbf{I}}^{0010;0}_{mj})_{22}\right.$$
$$\left.+\frac{1}{\hat{\vartheta}_1}(\hat{\mathbf{I}}^{1000;0}_{mj})_{22}+(\hat{\mathbf{I}}^{0000;-1}_{mj})_{22}\right](\hat{\mathbf{J}}^{00}_{nk})_{22}$$
$$+\frac{r_o}{\tilde{\Lambda}_2}\left(\frac{1}{h^2}\right)(\hat{\mathbf{I}}^{0000;1}_{mj})_{22}(\hat{\mathbf{J}}^{11}_{nk})_{22}, \tag{28}$$

$$\hat{k}^{23}_{mjnk}=\frac{\nu r_o}{\tilde{\Lambda}_1}\left(\frac{2}{\gamma h}\right)(\hat{\mathbf{I}}^{0100;0}_{mj})_{23}(\hat{\mathbf{J}}^{01}_{nk})_{23}+\frac{r_o}{\tilde{\Lambda}_2}\left(\frac{2}{\gamma h}\right)$$
$$\times(\hat{\mathbf{I}}^{0001;0}_{mj})_{23}(\hat{\mathbf{J}}^{10}_{nk})_{23}, \tag{29}$$

$$\hat{k}^{33}_{mjnk}=\frac{(1-\nu)r_o}{\tilde{\Lambda}_1}\left(\frac{1}{h^2}\right)(\hat{\mathbf{I}}^{0000;1}_{mj})_{33}(\hat{\mathbf{J}}^{11}_{nk})_{33}$$
$$+\frac{r_o}{\tilde{\Lambda}_2}\left[\frac{1}{\hat{\vartheta}_1^2}(\hat{\mathbf{I}}^{1010;1}_{mj})_{33}+\frac{4}{\gamma^2}(\hat{\mathbf{I}}^{0101;-1}_{mj})_{33}\right](\hat{\mathbf{J}}^{00}_{nk})_{33}, \tag{30}$$

and the elements in the mass submatrix $\hat{\mathbf{m}}^{\alpha\beta}$ are given by

$$\hat{m}^{11}_{mjnk}=(\hat{\mathbf{I}}^{0000;1}_{mj})_{11}(\hat{\mathbf{J}}^{00}_{nk})_{11}, \tag{31}$$

$$\hat{m}^{22}_{mjnk}=(\hat{\mathbf{I}}^{0000;1}_{mj})_{22}(\hat{\mathbf{J}}^{00}_{nk})_{22}, \tag{32}$$

$$\hat{m}^{33}_{mjnk}=(\hat{\mathbf{I}}^{0000;1}_{mj})_{33}(\hat{\mathbf{J}}^{00}_{nk})_{33}, \tag{33}$$

where

$$\tilde{\Lambda}_1=(1-2\nu)\tilde{\Lambda}_2, \tag{34}$$

$$\tilde{\Lambda}_2=(1+\nu), \tag{35}$$

$$\hat{\vartheta}_0(\overline{x_1})=\tfrac{1}{2}(r_o+r_i)+\hat{\vartheta}_1\overline{x_1}, \tag{36}$$

$$\hat{\vartheta}_1(\overline{x_1})=\tfrac{1}{2}(r_o-r_i), \tag{37}$$

and

$$(\hat{\mathbf{I}}^{defg;Z}_{mj})_{\alpha\beta}=\int_{-1}^{1}\int_{-1}^{1}\frac{\partial^{d+e}\phi^{\alpha}_m(\overline{x_1},\overline{x_2})}{\partial\overline{x_1}^d\partial\overline{x_2}^e}\frac{\partial^{f+g}\phi^{\beta}_j(\overline{x_1},\overline{x_2})}{\partial\overline{x_1}^f\partial\overline{x_2}^g}$$
$$\times\hat{\vartheta}_0^Z(\overline{x_1})d\overline{x_1}\,d\overline{x_2}, \tag{38}$$

$$(\hat{\mathbf{J}}^{rs}_{nk})_{\alpha\beta}=\int_{-1/2}^{1/2}\frac{\partial^r\psi^{\alpha}_n(\overline{x_3})}{\partial\overline{x_3}^r}\frac{\partial^s\psi^{\beta}_k(\overline{x_3})}{\partial\overline{x_3}^s}d\overline{x_3}, \tag{39}$$

in which $\langle\alpha;\beta\rangle=\langle1,2,3;1,2,3\rangle$.

To be consistent with the frequency parameter generally defined in the literature, the eigenvalue in Eq. (21) is nondimensionalized in the following form:

$$\lambda=\frac{\omega r_o^2}{\pi^2}\sqrt{\frac{\rho h}{D}}, \tag{40}$$

where D is the flexural rigidity of plate.

IV. CONVERGENCE AND COMPARISON STUDIES

The convergence characteristics of the frequency parameters λ for annular sector plates with SSSS (simply supported), SFSF, CSCS, and CCCC (clamped) boundary conditions are presented in Table I. Annular sector plates with $h/r_o=0.20$ having different sector angles ($\gamma=30°$, 45°, and 90°) are examined. The rate of convergence of λ is found to be consistent with the upper-bound characteristics of the Ritz method. It is observed that as the order of polynomial P, and the number of terms N, assumed in the amplitude functions increase, λ tends to converge monotonically from above (upper-bound convergence characteristics). For the simply supported (SSSS) annular plate, it is noted that convergence to at least four significant figures is achieved with two-dimensional polynomials $P=9$ and one-dimensional polynomial $N=4$. For the fully clamped (CCCC) annular sector plate, reliable λ with accuracy up to at least three significant figures are obtained also with $P=9$ and $N=4$. Likewise, for the annular sector plates with FSFS and CSCS boundary conditions, reasonably accurate λ are achieved when $P=9$ and $N=4$.

A series of comparison tests is further carried out to examine the discrepancies between the present three-dimensional elasticity solutions, the classical thin-plate solutions, and the Mindlin solutions for sector plates. Comparisons of the frequency solutions obtained from the various approaches for the hard, simply supported annular sector plates are presented in Table II. At a relative thickness ratio of $h/r_o=0.01$, the present solutions are compared with the Mindlin plate solutions of Xiang *et al.* (1993). It is observed that when h/r_o is small, λ obtained from all theories are in good agreement. At $h/r_o=0.10$, the Mindlin plate theory provides a more accurate description of the dynamic behaviors of annular sector plate than the classical thin-plate theory. This can be deduced from the good correlation shown in Table II between the Mindlin plate solutions and the present three-dimensional results. In this table, λ obtained based on the Mindlin plate theory are extracted from Xiang *et al.* (1993).

Further comparisons of λ for thick annular sector plates with SSSS and CSCS boundary conditions are given in Tables III and IV. The Mindlin solutions of Mizusawa (1991) are listed, together with the present three-dimensional frequency results. The reported Mindlin solution assumed a shear correction factor of $\kappa^2=5/6$. It is deduced from Table V that for moderately thick ($h/r_o=0.10$) annular sector plates, the Mindlin plate theory predictions agree quite well with the antisymmetric thickness modes of the three-dimensional frequency values. At a higher thickness ratio ($h/r_o=0.20$) the discrepancies increase, particularly at the higher modes. Besides, it is also found that the vibration spectrum for thick annular sector plates is precipitated with several symmetric thickness modes which are not found in the reported Mindlin solutions. The ability to detect these symmetric thickness modes is an important attribute of the three-dimensional elasticity method.

TABLE I. Convergence of λ for annular sector plates with different boundary conditions having $r_i/r_o = 0.40$, and $h/r_o = 0.20$.

γ	Terms (P, N)	Mode sequence number 1	2	3	4	5	6
		(a) Annular sector plate with SSSS boundary					
30°	(5, 3)	5.5975[a]	6.7042	7.2414[a]	10.837	10.965[a]	12.157
30°	(6, 4)	5.5974[a]	6.7041	7.2413[a]	10.835	10.963[a]	12.155
30°	(6, 5)	5.5973[a]	6.7040	7.2412[a]	10.834	10.962[a]	12.155
30°	(7, 4)	5.5973[a]	6.7040	7.2412[a]	10.834	10.962[a]	12.154
30°	(8, 4)	5.5973[a]	6.7040	7.2412[a]	10.834	10.961[a]	12.153
30°	(9, 4)	5.5973[a]	6.7040	7.2412[a]	10.834	10.961[a]	12.153
		(b) Annular sector plate with SFSF boundary					
45°	(5, 3)	1.8893	4.3564[a]	5.0387	5.7106	6.1943[a]	7.4098[a]
45°	(6, 4)	1.8892	4.3562[a]	5.0385	5.7105	6.1941[a]	7.4097[a]
45°	(6, 5)	1.8890	4.3561[a]	5.0384	5.7105	6.1940[a]	7.4096[a]
45°	(7, 4)	1.8890	4.3561[a]	5.0384	5.7009	6.1940[a]	7.4096[a]
45°	(8, 4)	1.8890	4.3561[a]	5.0384	5.7008	6.1940[a]	7.4096[a]
45°	(9, 4)	1.8890	4.3561[a]	5.0384	5.7008	6.1940[a]	7.4096[a]
		(c) Annular sector plate with CSCS boundary					
45°	(5, 3)	5.2660	5.5973[a]	9.2919	9.3135[a]	9.5383	10.963[a]
45°	(6, 4)	5.2645	5.5973[a]	9.2902	9.3123[a]	9.5349	10.961[a]
45°	(6, 5)	5.2643	5.5973[a]	9.2901	9.3121[a]	9.5347	10.961[a]
45°	(7, 4)	5.2640	5.5973[a]	9.2899	9.3120[a]	9.5343	10.961[a]
45°	(8, 4)	5.2635	5.5973[a]	9.2895	9.3117[a]	9.5336	10.961[a]
45°	(9, 4)	5.2634	5.5973[a]	9.2895	9.3116[a]	9.5334	10.961[a]
		(d) Annular sector plate with CCCC boundary					
90°	(5, 3)	4.3804	5.7013	7.2935	7.5776	8.5688	9.4728
90°	(6, 4)	4.3777	5.6975	7.2922	7.5693	8.5646	9.4697
90°	(6, 5)	4.3774	5.6973	7.2922	7.5692	8.5643	9.4695
90°	(7, 4)	4.3765	5.6958	7.2915	7.5664	8.5635	9.4689
90°	(8, 4)	4.3758	5.6952	7.2914	7.5655	8.5624	9.4678
90°	(9, 4)	4.3755	5.6945	7.2911	7.5643	8.5621	9.4672

[a]Symmetric thickness mode.

V. BENCHMARK RESULTS

Vibration mode shapes are presented in Figs. 2–6 for annular sector plates with different boundary conditions. The contour plots for the radial $U_1(\overline{x_1},\overline{x_2},\overline{x_3})$, the circumferential $U_2(\overline{x_1},\overline{x_2},\overline{x_3})$, and out-of-plane $U_3(\overline{x_1},\overline{x_2},\overline{x_3})$ components and the corresponding three-dimensional displacement mode shapes are presented. The mode shapes are computed for annular sector plates with sector angle $\gamma = 90°$; cutout ratio $r_i/r_o = 0.50$, and relative thickness ratio $h/r_o = 0.10$.

Figure 2 shows the vibration modes of an annular sector

TABLE II. Comparison of λ for sector plates with simply supported boundary condition (antisymmetric thickness mode).

γ	Source	Mode sequence number 1	2	3	4	5	6
		(a) Sector plate with $h/r_o = 0.01$					
30°	Xiang *et al.* (1993)	9.9119	18.621	28.161	29.205	41.750	43.702
30°	Present	9.9284	18.637	28.177	29.223	41.810	43.733
90°	Xiang *et al.* (1993)	2.5768	5.7419	7.0874	9.9119	12.314	13.591
90°	Present	2.5965	5.7586	7.1046	9.9295	12.331	13.606
		(b) Sector plate with $h/r_o = 0.10$					
30°	Xiang *et al.* (1993)	8.5548	14.598	20.278	20.855	27.254	28.176
30°	Present	8.6121	14.728	20.504	21.092	27.629	28.570
90°	Xiang *et al.* (1993)	2.4670	5.2389	6.3463	8.5548	10.326	11.232
90°	Present	2.4879	5.2709	6.3852	8.6121	10.401	11.317
		(c) Sector plate with $h/r_o = 0.20$					
30°	Xiang *et al.* (1993)	6.5675	10.221	13.376	13.686	16.891	17.045
30°	Present	6.6604	10.405	13.654	13.974	16.787	17.442
90°	Xiang *et al.* (1993)	2.2166	4.3268	5.1019	6.5675	7.6833	8.2376
90°	Present	2.2422	4.3778	5.1656	6.6604	7.8015	8.3692

TABLE III. Comparison of λ for thick annular sector plates of SSSS and CSCS boundary conditions with $r_i/r_o=0.50$ (antisymmetric thickness mode).

			Mode sequence number					
γ	h/r_o	Source	1	2	3	4	5	6
		(a) Annular sector plate with SSSS boundary condition						
30°	0.10	Mizusawa (1991)	8.973	17.38	20.28	27.92	28.48	⋯
30°	0.10	Present	9.0271	17.554	20.519	28.339	28.923	33.580
30°	0.20	Mizusawa (1991)	6.830	11.77	13.37	17.35	17.66	⋯
30°	0.20	Present	6.9253	12.009	13.662	17.608	17.755	17.899
60°	0.10	Mizusawa (1991)	5.169	8.973	14.07	14.28	17.38	⋯
60°	0.10	Present	5.1874	9.0271	14.183	14.400	17.554	20.520
60°	0.20	Mizusawa (1991)	4.263	6.830	9.891	10.03	11.77	⋯
60°	0.20	Present	4.3049	6.9252	10.068	10.213	12.009	13.663
		(b) Annular sector plate with CSCS boundary condition						
30°	0.10	Mizusawa (1991)	10.53	19.46	21.04	29.27	29.79	⋯
30°	0.10	Present	10.642	19.784	21.333	29.806	30.407	34.158
30°	0.20	Mizusawa (1991)	7.336	12.14	13.53	17.48	17.75	⋯
30°	0.20	Present	7.4806	12.476	13.847	17.608	18.060	18.266
60°	0.10	Mizusawa (1991)	7.815	10.53	15.27	17.01	19.46	⋯
60°	0.10	Present	7.9080	10.643	15.449	17.294	19.784	21.335
60°	0.20	Mizusawa (1991)	5.404	7.336	10.27	10.49	12.14	⋯
60°	0.20	Present	5.5189	7.4806	10.499	10.779	12.476	13.847

plate with SSSS boundary condition. Most of the modes are characterized by transverse bending motions except for the third and fifth modes, which are the thickness symmetric modes dominated by in-plane stretching and shearing motions. Figure 3 depicts the mode shapes for an annular sector plate with free radial and hard, simply supported curvilinear (SFSF) edges. For an annular sector plate with clamped curvilinear (CFCF) edges, the vibration mode shapes are presented in Fig. 4. It is observed that the out-of-plane modes (antisymmetric thickness modes) of both the SFSF and CFCF annular sector plates are quite similar in appearance, except that for the annular sector plate with hard, simply supported radial edges (SFSF), rotational deformations are observed. This is because the hard, simply supported conditions impose constraint only on the circumferential (U_2) and out-of-plane (U_3) displacements. The radial displacement (U_1) at the curvilinear edges is allowed. On the other hand, for the CFCF annular sector plate (clamped curvilinear edges and free radial edges), all components of displacement at the curvilinear edges are restrained. The symmetric thickness modes (which exhibit in-plane shearing and stretching motions) for both the SFSF and CFCF annular sector plates are also found to occur at different mode sequence numbers. The symmetric thickness motions occur at the second and fifth modes for the SFSF plate, and the fifth mode for the CFCF plate. Comparatively, the SFSF annular sector plate has more symmetric thickness modes than the annular sector plate with CFCF boundary condition.

Figures 5 and 6 show the three-dimensional vibration mode shapes of annular sector plates with CSCS and CCCC boundary conditions. For the CSCS annular sector plate (hard, simply supported radial edges) as shown in Fig. 5, it is noticed that the symmetric thickness in-plane mode takes place at the fourth mode sequence. Most of the vibration modes for this plate configuration are dominated by the out-of-plane transverse bending motions with increasing number of half waves in the circumferential direction. Figure 6 shows the deformed mode shapes of a fully clamped (CCCC)

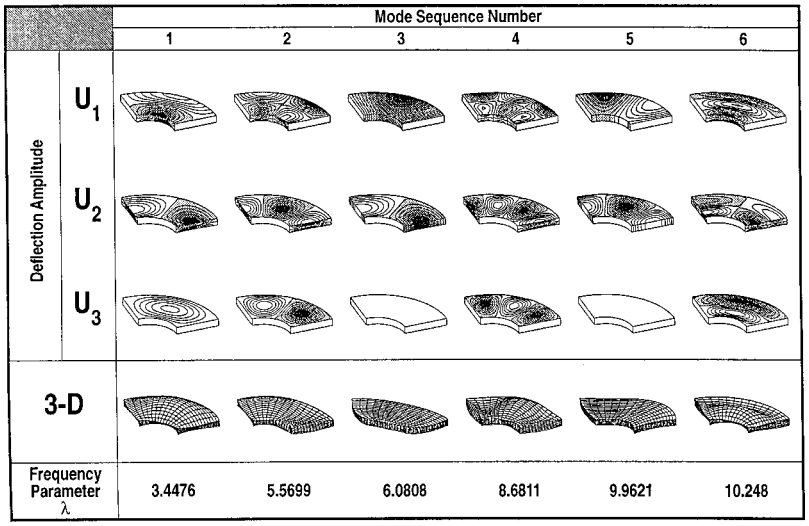


FIG. 2. Deformed mode shapes and frequency parameters of an annular sector plate with SSSS boundary condition ($r_i/r_o=0.50$, $h/r_o=0.10$, $\gamma=90°$).

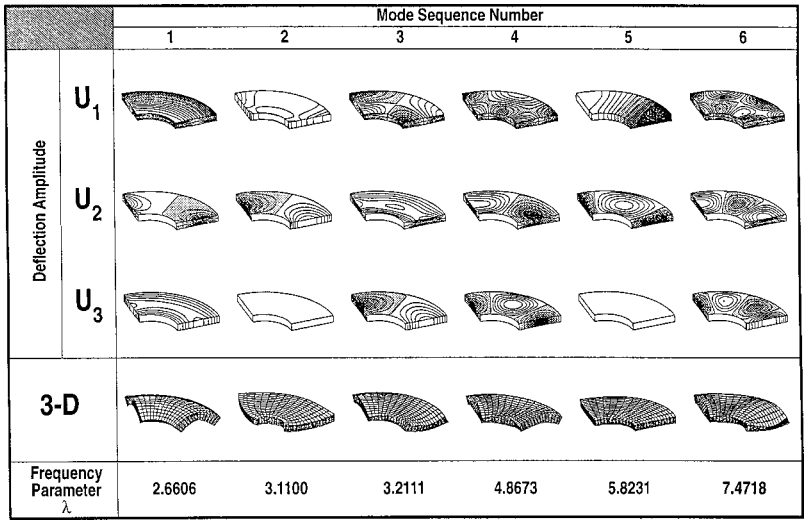


FIG. 3. Deformed mode shapes and frequency parameters of an annular sector plate with SFSF boundary condition ($r_i/r_o=0.50$, $h/r_o=0.10$, $\gamma=90°$).

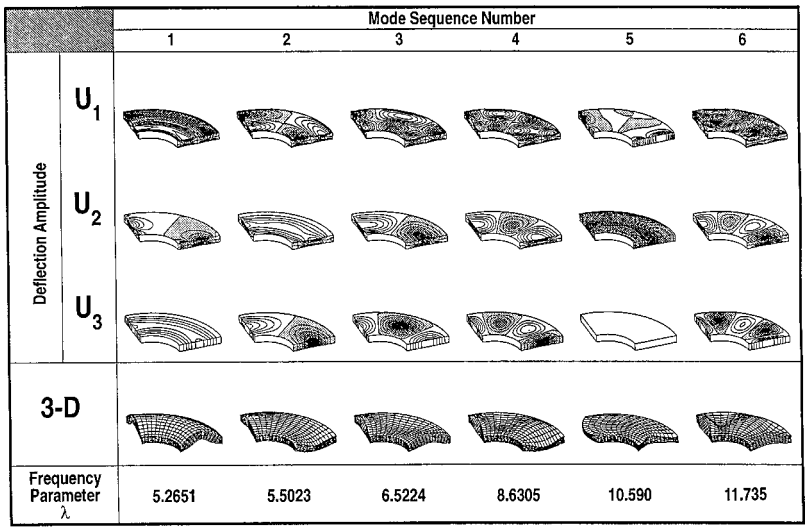


FIG. 4. Deformed mode shapes and frequency parameters of an annular sector plate with CFCF boundary condition ($r_i/r_o=0.50$, $h/r_o=0.10$, $\gamma=90°$).

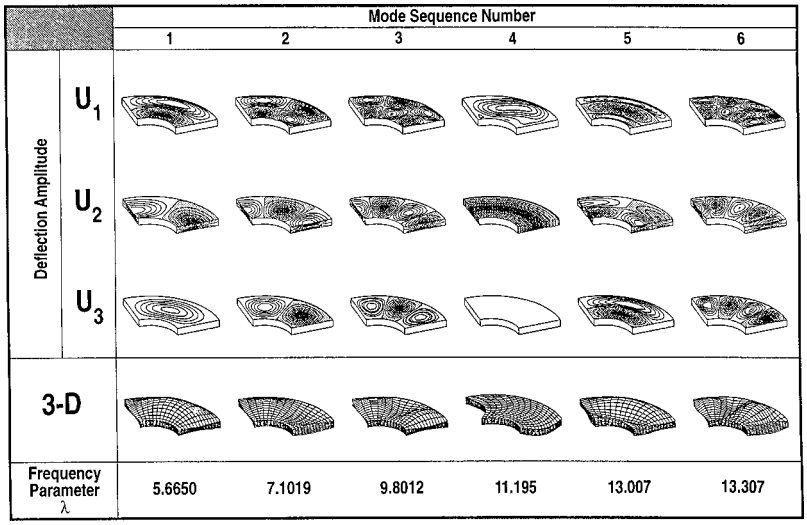


FIG. 5. Deformed mode shapes and frequency parameters of an annular sector plate with CSCS boundary condition ($r_i/r_o=0.50$, $h/r_o=0.10$, $\gamma=90°$).

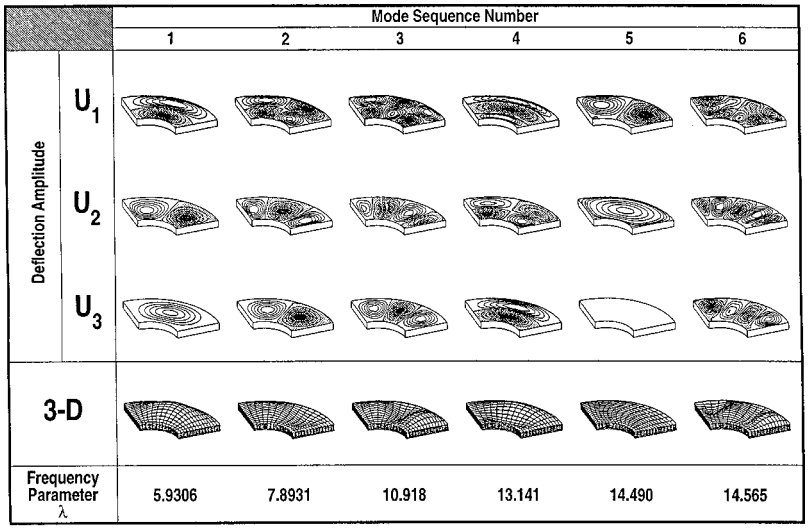


FIG. 6. Deformed mode shapes and frequency parameters of an annular sector plate with CCCC boundary condition ($r_i/r_o=0.50$, $h/r_o=0.10$, $\gamma=90°$).

TABLE IV. Frequency parameters λ for annular sector plates with simply supported radial edges having $r_i/r_o=0.40$.

		Mode sequence number					
γ	h/b	1	2	3	4	5	6
		(a) Annular sector plate with FSFS boundary condition					
30°	0.01	4.7853	12.181	17.762	20.253	30.925	33.152
30°	0.10	4.3891	10.211	12.388[a]	13.729[a]	13.995	15.633
30°	0.20	3.7031	6.1940[a]	6.8625[a]	7.6246	9.9214	10.501[a]
45°	0.01	2.1840	7.0378	8.2552	13.452	17.764	18.280
45°	0.10	2.0845	6.2728	7.2220	8.7122[a]	10.975	12.388[a]
45°	0.20	1.8890	4.3561[a]	5.0384	5.7008	6.1940[a]	7.4096[a]
60°	0.01	1.2096	4.7865	4.9358	10.314	10.804	12.184
60°	0.10	1.1724	4.3891	4.5047	5.6807[a]	8.7973	9.0361
60°	0.20	1.0987	2.8405[a]	3.7031	3.7553	5.8660[a]	6.1940[a]
90°	0.01	0.46621	2.1844	3.1920	4.7867	7.0391	8.2583
90°	0.10	0.45616	2.0845	2.5319[a]	2.9469	4.3891	6.2728
90°	0.20	0.43875	1.2661	1.8890	2.5336	3.7031	4.2969[a]
		(b) Annular sector plate with SSSS boundary condition					
30°	0.01	10.003	19.752	28.042	33.857	43.419	53.099
30°	0.10	8.6811	11.195[a]	14.481[a]	15.481	20.504	21.923[a]
30°	0.20	5.5973[a]	6.7040	7.2412[a]	10.834	10.961[a]	12.153[a]
45°	0.01	6.1138	14.977	15.048	26.285	28.042	28.899
45°	0.10	5.5699	9.9621[a]	11.195[a]	12.289	12.353	18.991[a]
45°	0.20	4.5871	4.9818[a]	5.5973[a]	8.9463	8.9990	9.4959[a]
60°	0.01	4.6891	10.003	13.307	17.947	19.752	27.181
60°	0.10	4.3542	7.8285[a]	8.6811	11.108	11.195[a]	14.314
60°	0.20	3.6952	3.9154[a]	5.5973[a]	6.7040	7.2412[a]	8.2218
90°	0.01	3.6611	6.1138	10.003	12.121	14.977	15.048
90°	0.10	3.4476	5.5699	6.0808[a]	8.6811	9.9621[a]	10.248
90°	0.20	2.9973	3.0424	4.5870	4.9818[a]	5.5973[a]	6.7040
		(c) Annular sector plate with CSCS boundary condition					
30°	0.01	12.084	24.054	30.650	40.982	46.796	57.726
30°	0.10	9.7999	11.195[a]	17.011	21.293	21.923[a]	22.552
30°	0.20	5.5973[a]	7.1143	10.961[a]	11.202	11.281[a]	13.834
45°	0.01	8.5678	17.143	20.124	29.918	30.659	36.897
45°	0.10	7.1019	11.195[a]	13.306	14.449	18.606[a]	20.563
45°	0.20	5.2634	5.5973[a]	9.2895	9.3116[a]	9.5334	10.961[a]
60°	0.01	7.4617	12.084	18.839	20.146	24.055	30.661
60°	0.10	6.2171	9.7999	11.195[a]	13.592	15.222	16.563
60°	0.20	4.5989	5.5973[a]	7.1143	8.2881[a]	8.9303	9.8659[a]
90°	0.01	6.7600	8.5679	12.090	17.150	17.960	20.124
90°	0.10	5.6650	7.1019	9.8012	11.195[a]	13.007	13.307
90°	0.20	4.1722	5.2634	5.5973[a]	7.0733[a]	7.1145	8.4995

[a]Symmetric thickness mode.

TABLE V. Frequency parameters λ for annular sector plates with free and clamped radial edges having $r_i/r_o=0.40$.

		Mode sequence number					
γ	h/b	1	2	3	4	5	6
		(a) Annular sector plate with SFSF boundary condition					
30°	0.01	2.7388	6.4761	10.929	16.531	19.842	24.582
30°	0.10	1.6533[a]	2.5890	5.5342	6.2916[a]	9.3055	12.892
30°	0.20	0.83447[a]	2.2932	3.1462[a]	4.2766	7.0616	8.5447[a]
45°	0.01	2.7681	4.8353	10.966	11.363	13.812	22.394
45°	0.10	2.0269[a]	2.6203	4.3374	7.2970[a]	9.3258	11.258
45°	0.20	1.0211[a]	2.3191	3.5796	3.6484[a]	6.2134	7.0787
60°	0.01	2.7854	4.0850	8.0119	11.045	12.679	14.699
60°	0.10	2.4088[a]	2.6393	3.7442	6.9584	7.3378[a]	9.4097
60°	0.20	1.2113[a]	2.3362	3.1732	3.6692[a]	5.1537[a]	5.4889
90°	0.01	2.8043	3.4396	5.3865	8.6011	11.083	11.846
90°	0.10	2.6606	3.1100[a]	3.2111	4.8673	5.8231[a]	7.4718
90°	0.20	1.5608[a]	2.3558	2.7829	2.9136[a]	4.0289	4.6820[a]
		(b) Annular sector plate with CFCF boundary condition					
30°	0.01	6.1645	8.9782	17.028	21.472	21.959	33.406
30°	0.10	5.1999	6.8881	9.3303[a]	12.364	14.759	15.935
30°	0.20	3.8514	4.6736[a]	4.7089	8.0278	8.7357[a]	8.7866[a]
45°	0.01	6.1982	7.5941	13.405	17.112	19.197	24.636
45°	0.10	5.2310	6.1150	9.9866[a]	10.376	12.424	13.608
45°	0.20	3.8660	4.3711	4.9989[a]	7.3627	8.0720	8.5416[a]
60°	0.01	6.2161	7.0315	10.277	16.694	17.150	18.383
60°	0.10	5.2478	5.7738	8.1525	10.296[a]	12.450	12.685
60°	0.20	3.8747	4.1907	5.1522[a]	5.8880	7.9378[a]	8.0787
90°	0.01	6.2350	6.6011	8.0169	10.802	15.120	17.194
90°	0.10	5.2651	5.5023	6.5224	8.6305	10.590[a]	11.735
90°	0.20	3.8839	4.7411	4.0336	5.2979[a]	6.2429	6.7281[a]
		(c) Annular sector plate with CCCC boundary condition					
30°	0.01	19.035	30.738	41.967	46.316	60.066	67.188
30°	0.10	13.280	19.427	24.325	25.636[a]	26.805	29.625
30°	0.20	8.5269	11.920	12.835[a]	14.364	14.839[a]	16.010
45°	0.01	11.264	22.133	22.666	35.823	38.064	38.360
45°	0.10	8.8580	15.427	15.644	21.831[a]	21.848[a]	22.337
45°	0.20	6.1794	9.8909	9.9787	10.939[a]	10.948[a]	13.067[a]
60°	0.01	8.6418	15.190	19.633	24.619	26.852	36.074
60°	0.10	7.0959	11.482	14.034	16.979	17.997[a]	18.138
60°	0.20	5.1482	7.7658	9.0153[a]	9.1151	10.174[a]	10.905
90°	0.01	7.0824	9.6990	14.080	18.181	19.777	21.002
90°	0.10	5.9306	7.8931	10.918	13.141	14.490[a]	14.565
90°	0.20	4.3755	5.6945	7.2911[a]	7.5643	8.5621	9.4678[a]

[a]Symmetric thickness mode.

annular sector plate. It is interesting to note that for this case, the appearance of the antisymmetric modes resembles the CSCS annular sector plate depicted in Fig. 5. The symmetric thickness mode occurs at the fifth mode of vibration. Comparatively, the frequency parameters λ for the CCCC annular sector plate are the highest among all the annular sector plate configurations examined here.

VI. CONCLUDING REMARKS

The first known elasticity solutions from three-dimensional analysis were presented for free vibrations of annular sector plates. Although the method is applicable to plates of arbitrary (but continuous) boundary conditions, only solutions for selected example problems were given and discussed in this paper. Detailed vibration behaviors of these selected annular sector plates with various thickness ratios and sector angles were carefully investigated. The three-dimensional mode shapes presented encompass the flexural, thickness twist, and thickness shear motions. From the comparison study with the Mindlin solutions, it was found that the thickness twist and shear modes are unable to be predicted by two-dimensional plate theories. Furthermore, the three-dimensional analysis also reveals symmetric thickness modes, which may be important in some practical applications.

Bapu Rao, M. N., Guruswamy, P., and Sampath Kumaran, K. S. (**1977**). ''Finite element analysis of thick annular and sector plates,'' Int. J. Nuclear Eng. Des. **41**, 247–255.

Ben-Amoz, M. (**1959**). ''Note on deflections and flexural vibrations of clamped sectorial plates,'' Trans. ASME, J. Appl. Mech. **26**, 136–137.

Bhat, R. B., (**1997**). ''Nature of stationarity of the natural frequencies at the natural modes in the Rayleigh–Ritz method,'' J. Sound Vib. **203**, 251–263.

Bhat, R. B., Chakraverty, S., and Sitiharu, I. (**1998**). ''Recurrence scheme for the generation of two-dimensional boundary characteristic orthogonal

polynomials to study vibration of plates,'' J. Sound Vib. **216**, 321–327.
Bhattacharya, A. P., and Bhowmic, K. N. (**1975**). ''Free vibration of a sectorial plate,'' J. Sound Vib. **41**, 503–505.
Gorman, D. J. (**1996**). ''Analysis of the free vibration of thick plates with torsional elastic support by the Superposition-Galerkin method,'' Mechanique Industrielle Materiaux **49**, 71–72.
Gorman, D. J. (**1997**). ''Accurate free vibration analysis of shear-deformable plates with torsional elastic edge support,'' J. Sound Vib. **203**, 209–218.
Guruswamy, P., and Yang, T. Y. (**1979**). ''A sector finite element for dynamic analysis of thick plates,'' J. Sound Vib. **62**, 505–516.
Hutchinson, J. R. (**1979**). ''Axisymmetric flexural vibrations of a thick free circular plate,'' Trans. ASME, J. Appl. Mech. **46**, 139–144.
Hutchinson, J. R. (**1984**). ''Vibrations of a thick free circular plate, exact versus approximate solutions,'' Trans. ASME, J. Appl. Mech. **51**, 581–585.
Kim, C. S., and Dickinson, S. M. (**1989a**). ''On the lateral vibration of thin annular and circular composite plates subject to certain complicating effects,'' J. Sound Vib. **130**, 363–367.
Kim, C. S., and Dickinson, S. M. (**1989b**). ''On the free, transverse vibration of annular and circular, thin, sectorial plates subject to certain complicating effects,'' J. Sound Vib. **134**, 407–421.
Laura, P. A. A., Gutierrez, R. H., and Rossi, R. E. (**1999**). ''Transverse vibrations of a circular, annular plate with free edges and two, intermediate concentric circular supports,'' J. Sound Vib. **226**, 1043–1047.
Laura, P. A. A., Gutierrez, R. H., and Rossi, R. E. (**2000**). ''Vibrations of a circular annular plates of cylindrical anisotropy and nonuniform thickness,'' J. Sound Vib. **231**, 246–252.
Liew, K. M., and Lam, K. Y. (**1993**). ''On the use of 2D orthogonal polynomials in the Rayleigh-Ritz method for flexural vibration of annular sector plates of arbitrary shape,'' Int. J. Mech. Sci. **35**, 129–139.
Liew, K. M., Hung, K. C., and Lim, M. K. (**1993**). ''A continuum three-dimensional vibration analysis of thick rectangular plates,'' Int. J. Solids Struct. **30**, 3357–3379.
Liew, K. M., Xiang, Y., and Kitipornchai, S. (**1995**). ''Research on thick plate vibration: A literature survey,'' J. Sound Vib. **189**, 163–176.
Liew, K. M., Wang, C. M., Xiang, Y., and Kitipornchai, S. (**1998**). *Vibration of Mindlin Plates: Programming the P-Ritz Method* (Elsevier Science, Oxford).
Maruyama, K., and Ichinomiya, O. (**1981**). ''Experimental investigation of free vibrations of clamped sector plates,'' J. Sound Vib. **74**, 565–573.
Mizusawa, T. (**1991**). ''Vibration of thick annular sector plates using semi-analytical method,'' J. Sound Vib. **150**, 245–259.
Narita, Y. (**1985**). ''Natural frequencies of free, orthotropic elliptical plates,'' J. Sound Vib. **93**, 593–597.
Qatu, M. S. (**1994**). ''On the validity of nonlinear shear deformation theories for laminated plates and shells,'' Compos. Struct. **27**, 395–401.
Rubin, C. (**1975**). ''Nodal circles and natural frequencies for the isotropic wedge,'' J. Sound Vib. **39**, 523–526.
Srinivasan, R. S., and Thiruvenkatachari, V. (**1985**). ''Free vibration of transverse isotropic annular sector Mindlin plates,'' J. Sound Vib. **101**, 193–201.
Westmann, R. A. (**1962**). ''A note on free vibrations of triangular and sector plates,'' J. Aerosp. Sci. **29**, 1139–1140.
Xiang, Y., Liew, K. M., and Kitipornchai, S. (**1993**). ''Transverse vibration of thick annular sector plates,'' Trans. ASCE J. Eng. Mech. **119**, 1579–1599.

$\mathbb{S}$-matrix theory applied to acoustic scattering by asymmetrically fluid-loaded elastic isotropic plates

H. Franklin
Laboratoire d'Acoustique Ultrasonore et d'Electronique, UMR CNRS 6068, Université du Havre, Place R. Schuman, 76610 Le Havre, France

E. Danila
Physics Department, Galati University, 111, Domneasca Street, Galati, Romania

J.-M. Conoir
Laboratoire d'Acoustique Ultrasonore et d'Electronique, UMR CNRS 6068, Université du Havre, Place R. Schuman, 76610 Le Havre, France

(Received 27 June 2000; revised 13 February 2001; accepted 31 March 2001)

Exact and approximate formalisms describing the interactions of acoustic plane waves with an elastic isotropic plate immersed between two different fluids (asymmetrically fluid-loaded plate) are presented. This constitutes an extension of the Fiorito, Madigosky, and Überall (FMU) [R. Fiorito *et al.*, J. Acoust. Soc. Am. **66**, 181 (1979)] theory, refined later by Freedman [A. Freedman, J. Sound Vib. **82**, 181 (1982); *ibid.* **82**, 197 (1982)]. The method, based upon the multichannel resonant scattering theory derived from quantum physics [A. Bohm, *Quantum Mechanics, Foundations, and Applications* (Springer, New York, 1993)], consists of two parts. First, a 2×2 scattering $\mathbb{S}$-matrix in which the diagonal elements are the two reflection coefficients and the off-diagonal elements are the two transmission coefficients, is built. Second, in order to compare our results with the FMU theory, resonant approximations are given for these coefficients, assuming light fluids when compared to the plate. The approximated coefficients show that the resonance widths are the sum of two independent partial widths, each of them being related to one fluid and to the plate physical properties. Of importance in this extension is the fact that the eigenvalues of the matrix, which reveal all the resonant features of the immersed plate, allow the separation between antisymmetrical and symmetrical modes, contrary to the reflection and transmission coefficients. The eigenvalues also allow the analysis of resonances despite the overlapping phenomenon. For the computations, the exact coefficients and eigenvalues, rather than their approximate forms, are used. This allows us to check the validity of the exact part of the theory under any fluid loading and not especially for light fluid loading. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1373636]

PACS numbers: 43.40.Fz [CBB]

I. INTRODUCTION

The subject of propagating waves on fluid-loaded plates and shells has become of great interest. Several papers have dealt recently with the subject of an elastic plate loaded with the same fluid[1–4] or with different fluids on each side.[5–7] The method used in these works is based upon an analysis of the dispersion curves, except Ref. 2 which is based upon the phase derivative properties of the reflection coefficient. The analysis of propagating vibration modes in plate structures is simpler than the analysis of vibration modes in cylindrical or spherical shells. Knowledge of the behavior of modes in an asymmetrically fluid-loaded plate is helpful for the understanding of the waves generated in shells. For a cylindrical or a spherical shell loaded with different fluids inside and outside, it is possible to extract the parameters of the inside fluid from the resonant scattering, as related to the dispersion of propagating plate waves. However, despite these numerous works, some questions have not yet received attention. For instance, there is lack of a generalized resonance theory which can account for the interactions of acoustic plane waves with an asymmetrically fluid-loaded elastic plate.

The resonance scattering theory of acoustic waves interacting with a symmetrically fluid-loaded elastic isotropic plate S (we will call this structure $F1/S/F1$, $F1$ standing for fluid) has been made by Fiorito, Madigosky and Überall[8] (FMU) and refined later by Freedman[9,10] under the assumption of light fluid loading. When a resonance occurs, the reflection coefficient modulus reaches zero (the curve has the shape of a trough), while the transmission coefficient modulus equals one (the curve has the shape of a peak). Then, the resonance scattering theory gives explicit expressions for the individual resonance amplitudes in terms of the Breit–Wigner formulas. From that, the reflection (or the transmission) coefficient is expressed as a symbolic summation of individual Breit–Wigner formulas. This summation allowed these authors to interpret the overlapping phenomenon between resonances (responsible for differences observed between the Breit–Wigner formulas and the curves of the reflection coefficient) that occurred particularly at angles near the Rayleigh angle. In this paper an alternative method is presented which does not use the symbolic summation, but introduces the properties of the scattering matrix for interpre-

tation of the overlapping. At the same time, an extension of the domain of applicability of the FMU's theory to asymmetrically fluid-loaded plates is proposed.

Let us assume that the resonance widths and positions for the structure are known. If one replaces exclusively the fluid on the lower face by another fluid, say $F2$ (we will call this second structure $F1/S/F2$ or asymmetrically fluid-loaded plate), the reflection and the transmission coefficients moduli show an enlargement or a shrinking of the widths. In all cases, a diminution of the amplitude of the troughs and peaks occurs with no significant shift in the resonance positions, if the fluids are light. One can then consider that a resonance width of the $F1/S/F2$ structure is the sum of two partial widths. Each of them results from the reemission of the acoustic energy stored by the plate into one of the fluids (if the fluids are identical, which is the case of the $F1/S/F1$ structure, the two partials widths are equal and their sum equals the width expressed by FMU).

Each of the two possibilities of reemission by reflection into one of the fluids defines a channel. From this consideration and from the energy flow conservation through the plate, it is possible to construct a scattering $\mathbb{S}$-matrix for the $F1/S/F2$ structure. The diagonal matrix elements are the two reflection coefficients constituting the two channels. The off-diagonal elements are the transmission coefficients. The method proposed here extends and shows for the analytical expressions, in the same validity domain of the FMU formalism (that is for light fluids loading), several new aspects such as: (1) the explanation, by the introduction of the partial width concept, of the broadening, the shrinking and the shortening mechanisms of the resonance curves due to the differences of the acoustical parameters of the loading fluids; (2) the explanation of the independence between partial widths for the retained approximations (summation of the partial widths gives the resonance width); (3) when the fluids are different, it is no longer possible to separate, in the reflection and transmission coefficients, the contribution of symmetrical modes from that of antisymmetrical modes. However, it is still possible to distinguish nearly symmetrical and nearly antisymmetrical modes from the $\mathbb{S}$-matrix eigenvalues study, even when the fluids are not actually light; (4) the ability of the eigenvalues to reveal in a powerful way the resonant feature of the reflection and transmission coefficients. Obviously, in cases (1) and (2), if the assumption of light fluids fails, the interactions between the two fluids must be taken into account: the discrimination between partial widths is not possible and the two channel theory is no longer valid. Then, only the exact forms of the reflection and transmission coefficients must be used for the evaluation of the resonant widths.

In Sec. II, the scattering $\mathbb{S}$-matrix of a $F1/S/F2$ structure is constructed. The eigenvalues are then expressed in terms of the reflection and transmission coefficients. In Sec. III, exact forms of these coefficients are furnished. Assuming light fluid loading, resonance decompositions in the frequency variable are given for the reflection and transmission coefficients and for the eigenvalues. It is shown that all resonant quantities can be expressed with the help of two different partial widths. Some particular structures (the $F1/S/$Vacuum and the $F1/S/F1$) are briefly analyzed. Section IV deals with exact forms and their resonance decompositions in the angular variable. In particular, the way the eigenvalues can be used to describe and to evaluate the main characteristics of the Rayleigh type modes is shown.

In the paper, the numerical computations are performed using the exact theory. The fluids then do not need to be light. We choose water and glycerine. It should be pointed out that use of those fluids in the approximate formulas would give poor results, as the light fluids assumption would not hold.

II. THE $\mathbb{S}$-MATRIX CONSTRUCTION

A. Definitions and numerical values for the physical parameters

Consider plane monochromatic waves, with angular frequency ω, propagating through the $F1/S/F2$ structure. The $F1$ (respectively, $F2$) fluid has density ρ_1 (resp. ρ_2), sound wave speed c_1 (resp. c_2) and associate wave number $k_1 = \omega/c_1$ (resp. $k_2 = \omega/c_2$). The S plate has thickness d and density ρ. The velocity of the L longitudinal waves is c_L with associated wave number $k_L = \omega/c_L$. The velocity of the T transverse waves is c_T and the associated wave number is $k_T = \omega/c_T$. The incident (or reflection) and transmission angles are, for the fluids θ_1 and θ_2, and for the plate α (L waves) and β (T waves). The wave number following the Ox direction of interfaces is chosen positive. It is the same in all media and it will be noted $k_x = k_1 \sin\theta_1$. We have the following relations (Snell's laws):

$$k_1 \sin\theta_1 = k_2 \sin\theta_2 = k_L \sin\alpha = k_T \sin\beta, \tag{1a}$$

or identically, if one introduces the dimensionless wave number $\overline{k_x} = \sin\theta_1$,

$$\sin\theta_1 = n_2 \sin\theta_2 = n_T \sin\beta = n_L \sin\alpha, \tag{1b}$$

with $n_2 = c_1/c_2$, $n_T = c_1/c_T$ and $n_L = c_1/c_L$. Throughout the paper we will assume that

$$n_L < n_T < n_2 < 1, \tag{2}$$

which indicates the existence of three critical incidence angles. The following condition is also assumed:

$$\overline{k_x} < n_2, \tag{3}$$

because we are interested only in the phenomena occurring before the third critical incidence angle when the waves are propagating in both fluids. The wave numbers following the Oz direction perpendicular to the interfaces (with increasing values of z from $F1$ to $F2$) are conventionally chosen in the fluids such as

$$k_{z1} = k_1 \cos\theta_1 = k_1 \overline{k_{z1}} > 0, \quad k_{z2} = k_2 \cos\theta_2 = k_1 \overline{k_{z2}} > 0, \tag{4}$$

with

$$\overline{k_{z1}} = (1 - \overline{k_x}^2)^{1/2} \text{ and } \overline{k_{z2}} = (n_2^2 - \overline{k_x}^2)^{1/2}. \tag{5}$$

For the waves in the plate, three regions must be distinguished in which the z components of the wave vectors take the following forms:

Region I—If $\overline{k_x}<n_L$, then $k_{zT}=k_1\overline{k_{zT}}>0$, $k_{zL}=k_1\overline{k_{zL}}>0$, with $\overline{k_{zL}}=(n_L^2-\overline{k_x^2})^{1/2}>0$ and $\overline{k_{zT}}=(n_T^2-\overline{k_x^2})^{1/2}>0$. The L and T waves both propagate in the plate.

Region II—If $n_L<\overline{k_x}<n_T$, the L waves are evanescent while the T waves propagate in the plate ($\overline{k_{zL}}=+i(\overline{k_x^2}-n_L^2)^{1/2}$ is an imaginary positive number, k_{zT} is real positive and given by Eq. (5)),

Region III—If $n_T<\overline{k_x}<n_2$, both the L and T waves are evanescent in the plate ($\overline{k_{zL}}=+i(\overline{k_x^2}-n_L^2)^{1/2}$ and $\overline{k_{zT}}=+i(\overline{k_x^2}-n_T^2)^{1/2}$ are imaginary positive numbers). Waves propagate only in the fluids.

Throughout this paper, numerical results will be given by using:

(1) Water for the $F1$ fluid: $\rho_1=1000\,\mathrm{kg/m^3}$; $c_1=1485\,\mathrm{m/s}$;
(2) Glycerine for the $F2$ fluid: $\rho_2=1260\,\mathrm{kg/m^3}$; $c_2=1920\,\mathrm{m/s}$;
(3) Aluminum plate with: $\rho=2790\,\mathrm{kg/m^3}$; $c_L=6380\,\mathrm{m/s}$; $c_T=3100\,\mathrm{m/s}$.

With the above values of the velocities, the three critical angles are located approximately at 13.46°, 28.62° and 50.66°.

B. Energy flow conservation and $\mathbb{S}$-matrix construction

In a fluid, the acoustic intensity vector, which is defined as the average rate of flow of energy is

$$\boldsymbol{\mathcal{I}}=\langle p\,\boldsymbol{\nu}\rangle_t, \tag{6}$$

where p is the acoustic pressure and $\boldsymbol{\nu}$ the particle velocity (the $\langle\ \rangle_t$ symbol indicates time averaging). In the Oz direction, we may write

$$\mathcal{I}_z=\langle p\,\nu_z\rangle_t=\tfrac{1}{4}(p\,\nu_z^*+p^*\nu_z), \tag{7}$$

where the sign * indicates the complex conjugate. In absence of absorption and other causes of attenuation, the acoustic intensity conservation in the Oz direction holds. Thus, we may write:

$$\mathcal{I}_{z1}=\mathcal{I}_{z2}. \tag{8}$$

Generally speaking, the acoustic pressure in fluids $F1$ and $F2$ is given by:

$$p_1=\rho_1\omega^2[A_1e^{ik_{z1}z}+B_1e^{-ik_{z1}z}]e^{i(k_xx-\omega t)}, \tag{9}$$

$$p_2=\rho_2\omega^2[A_2e^{ik_{z2}z}+B_2e^{-ik_{z2}z}]e^{i(k_xx-\omega t)}, \tag{10}$$

where the A_1, B_1, A_2, B_2 are complex amplitudes. In these equations, according to our convention Eqs. (4), an incident wave in the $F1$ fluid with amplitude A_1 generates a transmitted field in $F2$, so that $B_2=0$. In Eqs. (10), we have kept the part containing B_2 which represents an incident wave in $F2$ coming from $z=+\infty$, because it turns out to be necessary for the matrix construction. Considering now the z component of the linearized Euler equation for perfect fluids,

$$\rho_m\frac{\partial\nu_z}{\partial t}=-\frac{\partial p_m}{\partial z},\quad (m=1,2) \tag{11}$$

one obtains, with the help of Eq. (8):

$$A_1A_1^*+\chi^2B_2B_2^*=B_1B_1^*+\chi^2A_2A_2^*, \tag{12}$$

with

$$\chi^2=\frac{\rho_2k_{z2}}{\rho_1k_{z1}}. \tag{13}$$

Equation (12) states the equality of two vectors lengths: $\mathbf{e}=(A_1,\chi B_2)$ called the input vector and $\mathbf{s}=(B_1,\chi A_2)$ called the output vector. Both of them are defined in a complex two-dimensional vector space supplied with orthogonal axis. It should be noticed that this space is purely abstract, with no relation with the usual physical space. Because of the energy conservation, Eq. (12), the input and output vectors are related by a scattering matrix $\mathbb{S}$ with complex elements S_{mn} $(m,n=1,2)$ as follows:

$$\mathbf{s}=\mathbb{S}\mathbf{e}. \tag{14}$$

This matrix verifies $\mathbb{S}\mathbb{S}^+=\mathbb{I}$ (the sign + designates the Hermitian conjugate or the adjoint, while $\mathbb{I}$ is the identity matrix). In order to determine the physical meaning of the $\mathbb{S}$-matrix elements, let us consider two cases for the components of the input vector $\mathbf{e}$.

(1) The incident wave propagates in the increasing z direction (from $F1$ to $F2$):
$$A_1\neq0,\ \chi B_2=0;\ \chi A_2/A_1=S_{21},\ B_1/A_1=S_{11}. \tag{15}$$
The S_{11} (resp. S_{21}) element represents the plate reflection coefficient (resp. transmission coefficient multiplied by χ).
(2) The incident wave propagates in the decreasing z direction (from $F2$ to $F1$):
$$A_1=0,\ \chi B_2\neq0;\ \chi A_2/\chi B_2=S_{22},\ B_1/\chi B_2=S_{12}. \tag{16}$$

The S_{22} (resp. S_{12}) element represents the plate reflection coefficient (resp. transmission coefficient divided by χ). If one introduces now

$$r_1=B_1/A_1,\ t_2=B_1/B_2,\ t_1=A_2/A_1,\ r_2=A_2/B_2, \tag{17}$$

where r_1 (resp. r_2) is the reflection coefficient in the fluid $F1$ (resp. $F2$) and t_1 (resp. t_2) is the transmission coefficient from the fluid $F1$ to the fluid $F2$ (resp. from the fluid $F2$ to the fluid $F1$), the $\mathbb{S}$-matrix may be written in a first step as

$$\mathbb{S}=\begin{pmatrix} r_1 & \dfrac{1}{\chi}t_2 \\ \chi t_1 & r_2 \end{pmatrix}. \tag{18}$$

C. $\mathbb{S}$-matrix simplifications

A first simplification is provided by means of the invariance under time reversal principle. Under this operation, the pressures undergo the following transformation: $p(x,z,t)\rightarrow p^*(x,z,-t)$ and become

$$p_1^*=\rho_1\omega^2[A_1^*e^{-ik_{z1}z}+B_1^*e^{ik_{z1}z}]e^{-i(k_xx+\omega t)}, \tag{19}$$

$$p_2^*=\rho_2\omega^2[A_2^*e^{-ik_{z2}z}+B_2^*e^{ik_{z2}z}]e^{-i(k_xx+\omega t)}. \tag{20}$$

The invariance under time reversal implies that the pressures p_1^* and p_2^* are also solutions of the wave equations in each

fluid. As a consequence of the sign exchanges, the incoming and outgoing waves are exchanged. From that, it is found that the $\mathbb{S}$-matrix links the complex conjugate vectors $\mathbf{e}^*$ and $\mathbf{s}^*$ as follows

$$\mathbf{e}^* = \mathbb{S}\mathbf{s}^*. \tag{21}$$

The complex conjugate of Eq. (21) gives $\mathbf{e} = \mathbb{S}^*\mathbf{s}$, that is, by inverting

$$\mathbf{s} = (\mathbb{S}^*)^{-1}\mathbf{e}. \tag{22}$$

If this relation is identified with Eq. (14), it comes out

$$\mathbb{S}^* = \mathbb{S}^{-1}. \tag{23}$$

The $\mathbb{S}$ matrix being unitary, one has $\mathbb{S}^{-1} = \mathbb{S}^+$. It follows that $S_{21} = S_{12}$, or

$$\chi t_1 = (1/\chi) t_2, \tag{24}$$

and then

$$\mathbb{S} = \begin{pmatrix} r_1 & \chi t_1 \\ \chi t_1 & r_2 \end{pmatrix}. \tag{25}$$

A second simplification comes from the energy flow conservation [cf. Eq. (12)]. On one hand, if we put $B_2 = 0$, then $|r_1|^2 = 1 - \chi^2 |t_1|^2$. On the other hand, if we put $A_1 = 0$, then $|r_2|^2 = 1 - (1/\chi^2)|t_2|^2$. Using Eq. (24), it may be deduced that

$$|r_1| = |r_2|. \tag{26}$$

So, the reflection coefficients r_1 and r_2 only differ by a phase factor, because of our assumption that the incidence angle is less than the third critical angle. A third simplification occurs when the fluids $F1$ and $F2$ are identical. In this case, there is a space reflection symmetry which implies an invariance for the transformation $(x,z) \to (-x,-z)$. As a consequence, $r_1 = r_2$ and from Eq. (24) $t_1 = t_2$, with $\chi = 1$.

D. $\mathbb{S}$-matrix eigenvalues

A straightforward calculation gives the $\mathbb{S}$-matrix eigenvalues λ and μ

$$\lambda = \tfrac{1}{2}[r_1 + r_2 - 2\chi t_1 \Delta]; \quad \mu = \tfrac{1}{2}[r_1 + r_2 + 2\chi t_1 \Delta], \tag{27}$$

with $\Delta = \sqrt{((r_2 - r_1)/2\chi t_1)^2 + 1}$. The expression inside the square root is real and positive and $|\lambda| = |\mu| = 1$ because of the energy flow conservation. In the case of identical fluids, one has

$$\lambda = r_1 - t_1, \quad \mu = r_1 + t_1. \tag{28}$$

The eigenvalues exhibit important properties of the acoustic scattering which will be detailed in the following. However, before analyzing these properties, it is necessary to obtain their resonant forms.

III. RESONANCE DECOMPOSITION IN THE FREQUENCY VARIABLE

A. Exact forms of the reflection and transmission coefficients

The reflection and transmission coefficients of the $F1/S/F2$ structure are given in Brekhovskikh.[11] We use the following expressions, as they will prove useful for the decomposition into resonance terms

$$r_1 = \frac{(Ca - i\tau_1)(Cs - i\tau_2) + (Ca + i\tau_2)(Cs + i\tau_1)}{(Ca + i\tau_1)(Cs - i\tau_2) + (Ca + i\tau_2)(Cs - i\tau_1)}, \tag{29}$$

$$r_2 = \frac{(Ca + i\tau_1)(Cs + i\tau_2) + (Ca - i\tau_2)(Cs - i\tau_1)}{(Ca + i\tau_1)(Cs - i\tau_2) + (Ca + i\tau_2)(Cs - i\tau_1)}, \tag{30}$$

$$t_1 = \frac{2i\tau_1 (Ca + Cs) k_{z1}/k_{z2}}{(Ca + i\tau_1)(Cs - i\tau_2) + (Ca + i\tau_2)(Cs - i\tau_1)}. \tag{31}$$

In the above expressions,

$$Ca = (n_T^2 - 2\overline{k_x^2})^2 \tan(\overline{k_{zL}} k_1 d/2) + 4\overline{k_x^2}\,\overline{k_{zL}}\,\overline{k_{zT}} \tan(\overline{k_{zT}} k_1 d/2), \tag{32}$$

$$Cs = (n_T^2 - 2\overline{k_x^2})^2 \cot(\overline{k_{zL}} k_1 d/2) + 4\overline{k_x^2}\,\overline{k_{zL}}\,\overline{k_{zT}} \cot(\overline{k_{zT}} k_1 d/2), \tag{33}$$

$$\tau_1 = \frac{\rho_1}{\rho} n_T^4 \frac{\overline{k_{zL}}}{\overline{k_{z1}}} \quad \text{and} \quad \tau_2 = \frac{\rho_2}{\rho} n_T^4 \frac{\overline{k_{zL}}}{\overline{k_{z2}}}. \tag{34}$$

Equations (29)–(31) generalize those obtained by FMU for the $F1/S/F1$ structure. In particular, they show the mixing of symmetrical $(Cs \pm i\tau_m)$ and antisymmetrical $(Ca \pm i\tau_m)$ functions $(m = 1,2)$ when the plate is loaded with different fluids.

B. Resonance decomposition in the frequency variable

Let us now consider the resonant decomposition of these equations, which will allow us to incorporate them into the two channel theory (see the Appendix) and to extend previous works on resonant fluid-loaded plates.

When the incident angle is constant, the functions Ca and Cs depend on the frequency variable $X = k_1 d$ while the angular variable $\overline{k_x}$ (or θ_1) is a fixed parameter. Let $p \geqslant 0$ be an integer denoting the order of the plate mode. Following the terminology of Pitts *et al.*,[12] if $p = 0$ the mode is called Rayleigh–Lamb type mode (the well-known A_0 and S_0 modes) and if $p \geqslant 1$ the mode is called Lamb type mode. The equations $Ca(X_{p,a}) = 0$ and $Cs(X_{p,s}) = 0$ determine the resonance frequencies $X_{p,a}$ and $X_{p,s}$ of the antisymmetrical and symmetrical modes of the plate in vacuum. When the fluids are light comparatively to the plate, it is assumed that $|\tau_m| \ll 1$ $(m = 1,2)$. This condition is not satisfied for angles such that $\overline{k_x} \approx n_2$ or $\overline{k_x} \approx 1$, that is when $\overline{k_{z2}}$ or $\overline{k_{z1}}$ approaches zero and, therefore, the approximations considered here would not give good results under such situations. However, in the cases where the condition is satisfied, the light fluid loading implies that the modes are slightly perturbed by the fluids.

Hence, in Eqs. (27)–(29), in the vicinity of the resonance frequencies $X_{p,a,s}$, an expansion up to the first order of the functions $Ca(X)$ and $Cs(X)$ is made. The reflection and transmission coefficients can then be expressed as (index a if antisymmetrical, index s if symmetrical):

$$r_1 \approx r_{1,p,a,s} = \frac{X - X_{p,a,s} - i\Gamma^-_{p,a,s}/4}{X - X_{p,a,s} + i\Gamma^+_{p,a,s}/4}, \tag{35}$$

$$r_2 \approx r_{2,p,a,s} = \frac{X - X_{p,a,s} + i\Gamma^-_{p,a,s}/4}{X - X_{p,a,s} + i\Gamma^+_{p,a,s}/4}, \tag{36}$$

$$\chi t_1 = \chi t_{1,p,a,s} = \frac{\pm i\Gamma^{1/2}_{1,p,a,s}\Gamma^{1/2}_{2,p,a,s}/2}{X - X_{p,a,s} + i\Gamma^+_{p,a,s}/4}. \tag{37}$$

They satisfy the energy conservation law

$$|r_{1,p,a,s}|^2 + |\chi t_{1,p,a,s}|^2 = 1. \tag{38}$$

In these expressions,

$$\Gamma_{m,p,a,s} = \pm 2\tau_{m,p,a,s} \Big/ \left(\frac{dCa,s}{dX}(X_{p,a,s}) \right), \quad (m = 1,2) \tag{39}$$

and

$$\Gamma^+_{p,a,s} = \Gamma_{1,p,a,s} + \Gamma_{2,p,a,s}, \quad \Gamma^-_{p,a,s} = \Gamma_{1,p,a,s} - \Gamma_{2,p,a,s}. \tag{40}$$

The lower sign $(-)$ in Eqs. (37) and (39) corresponds to index s. In these resonance decompositions, the resonance shift due to fluid loading is neglected in both the numerator and the denominator because it is a second order phenomenon. For the retained approximations, a separation occurs between antisymmetrical and symmetrical modes as shown in Eq. (39). In the first three regions defined in Sec. II A, the above resonance forms do not change fundamentally. Only the widths are to be modified in order to account for the expressions of τ_1, τ_2, Ca, Cs, $Ca_{,X}$ and $Cs_{,X}$ in the different regions.

The right hand sides of Eqs. (35)–(37) are similar to the matrix elements of Eq. (A6): thus, the resonances of a $F1/S/F2$ structure enter into the frame of the two channel theory presented in the appendix. It is important to note also that the resonance functions are different from that given by FMU. For example, contrary to the result given by these authors, it may be seen that, for $X = X_{p,a,s}$, the reflection coefficients moduli never reach zero, except for the case of the equality of the two partial widths. In the same way, it is the only case where the resonant transmission coefficient modulus is equal to one. Referring to Eq. (39), it can be seen that the partial width $\Gamma_{1,p,a,s}$ (resp. $\Gamma_{2,p,a,s}$) depends only on the physical parameters of both the plate and the $F1$ (resp. $F2$) fluid. Thus, the two partial widths are independent within the bounds of the approximation considered here. The scattering of the system obeys to the two channel theory, with the width related to each fluid defining one channel. At a resonance frequency one has:

$$|r_{1,p,a,s}(X_{p,a,s})|^2 = |r_{2,p,a,s}(X_{p,a,s})|^2 \approx \left| \frac{\Gamma^-_{p,a,s}}{\Gamma^+_{p,a,s}} \right|^2 \leqslant 1, \tag{41}$$

$$|\chi t_{1,p,a,s}(X_{p,a,s})|^2 \approx \left| \frac{2\Gamma^{1/2}_{1,p,a,s}\Gamma^{1/2}_{2,p,a,s}}{\Gamma^+_{p,a,s}} \right|^2 \leqslant 1. \tag{42}$$

Actually, the above width ratios depend only on the fluid parameters because $Ca_{,X}$ and $Cs_{,X}$ do not appear in the expressions. The characteristic impedances for the fluids are defined as $Z_m = \rho_m c_m$, $(m = 1,2)$, and Eqs. (41) and (42) become

$$\left| \frac{\Gamma^-_{p,a,s}}{\Gamma^+_{p,a,s}} \right|^2 = \left| \frac{Z_1 \cos\theta_2 - Z_2 \cos\theta_1}{Z_1 \cos\theta_2 + Z_2 \cos\theta_1} \right|^2 = |r_{F1/F2}|^2 = |r_{F2/F1}|^2, \tag{43}$$

$$\left| \frac{2\Gamma^{1/2}_{1,p,a,s}\Gamma^{1/2}_{2,p,a,s}}{\Gamma^+_{p,a,s}} \right|^2 = \frac{4Z_1 Z_2 \cos\theta_1 \cos\theta_2}{(Z_1 \cos\theta_2 + Z_2 \cos\theta_1)^2} = |t_{F1/F2} t_{F2/F1}|, \tag{44}$$

Eq. (43) represents the squared modulus of the $F1/F2$ (or the $F2/F1$) interface reflection coefficient and is often defined as the power reflection coefficient. Eq. (44) represents the power transmission coefficient of the $F1/F2$ (or the $F2/F1$) interface. Because of the assumption made in Eq. (2), a critical incidence angle occurs in the curves of these coefficients if one considers an incident wave from the $F1$ fluid onto the plate.

C. Properties of the eigenvalues

With the approximate forms obtained for the coefficients r_1, r_2 and χt_1, the eigenvalue λ is found to be

$$\lambda_{p,a,s} \approx \frac{X - X_{p,a,s} \mp i\Gamma^+_{a,s}/4}{X - X_{p,a,s} + i\Gamma^+_{a,s}/4}, \tag{45}$$

the upper (resp. lower) sign being related to the index a (resp. s). The eigenvalue μ takes a similar form, apart from the fact that the sign $\mp$ has to be replaced by $\pm$. From Eq. (45), it can be seen that λ detects antisymmetrical modes while μ detects symmetrical ones. As a proof, when antisymmetrical modes are encountered

$$\lambda_{p,a} \approx \frac{X - X_{p,a} - i\Gamma^+_a/4}{X - X_{p,a} + i\Gamma^+_a/4} \text{ and } \mu_{p,a} \approx 1, \tag{46}$$

while when symmetric modes are encountered

$$\lambda_{p,s} \approx 1 \text{ and } \mu_{p,s} \approx \frac{X - X_{p,s} - i\Gamma^+_s/4}{X - X_{p,s} + i\Gamma^+_s/4}. \tag{47}$$

This discrimination between the two types of modes, which is a consequence of the two channel theory [Eqs. (46)–(47) are equivalent to Eqs. (A3)], can be seen numerically by plotting the square modulus of the two exact transition terms

$$T_\lambda = \frac{1}{2i}(\lambda - 1) \text{ and } T_\mu = \frac{1}{2i}(\mu - 1), \tag{48}$$

for a fixed incidence angle θ_1, and by comparing their curves to $|\chi t_1|^2$. In Fig. 1 which corresponds to $\theta_1 = 5°$, the curves of $|T_\lambda|^2$ (dotted) and $|T_\mu|^2$ (dashed) show differences in the resonant peaks of $|\chi t_1|^2$ (solid). The curves have the same widths and the same resonance positions but not the same amplitude. The modulus of the transition terms always

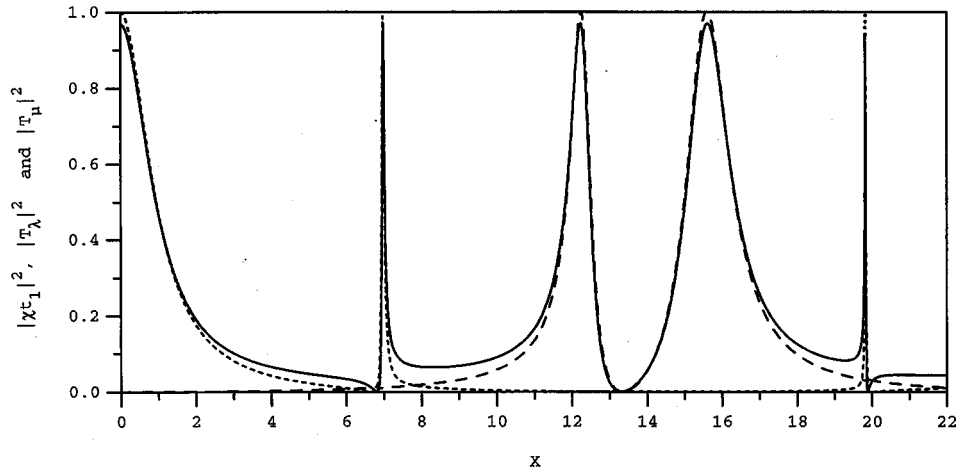


FIG. 1. $|\chi t_1|^2$ (solid), $|T_\lambda|^2$ (dotted) and $|T_\mu|^2$ (dashed) vs $X=k_1 d$ at $\theta_1=5°$ showing the separation between modes in the frequency variable.

equals one at resonance frequencies. It must be noticed that there is no reason for the peak positions and widths of the transition terms to be confused with those of the transmission coefficients. A case which contradicts the above example will be given in the following (Fig. 2).

In the case of identical fluids, it is easy to express the exact form of the eigenvalues. Eqs. (28) then become

$$\lambda=\frac{(Ca-i\tau_1)}{(Ca+i\tau_1)} \text{ and } \mu=\frac{(Cs+i\tau_1)}{(Cs-i\tau_1)}. \tag{49}$$

As a consequence, whatever the value of the frequency X, antisymmetrical and symmetrical parts are clearly separated and independent. In FMU theory, the reflection coefficient is indeed a sum of an antisymmetrical part and a symmetrical part, but it is multiplied by a coefficient which contains both antisymmetrical and symmetrical functions; see Eqs. (1b) of Ref. 1. It is now evident that the right decomposition is intrinsically an eigenvalue problem since one can write $r_1=(\mu+\lambda)/2$ and $t_1=(\mu-\lambda)/2$.

In Fig. 2, corresponding to angle $\theta_1=29°$ near a critical angle, the curve of the transition term $|T_\mu|^2$ (dotted) indicates that the transmission coefficient (solid) describes a symmetrical mode (the Rayleigh type mode S_0). In addition, both an enlargement of the width and a shift of the maximum position of $|T_\mu|^2$ with regard to the transmission coefficient can be observed. These two phenomena are probably a consequence of the fact that $|T_\lambda|^2$ (dashed) is not negligible in the whole frequency range represented in Fig. 2. Thus, the Rayleigh mode S_0 is subject to antisymmetrical effects, i.e., there is an overlapping between the S_0 (μ eigenvalue) and the antisymmetrical (λ eigenvalue) modes. In order to determine which one, $|T_\mu|^2$ or $|\chi t_1|^2$, gives the actual resonance position, it would be necessary to search for the roots of the transmission coefficient. Actually, this case does not exactly verify the two channel theory, in which the poles of both T_μ and χt_1 must be the same (see the Appendix). Nevertheless, in many other cases the shift and the width enlargement can be neglected and then the theory can be applied with no significant errors. It should be noted that the problem encountered in Fig. 2 occurs also for the $F1/S/F1$ structure.

D. Analysis of two particular structures

Now, let us briefly consider some implications of formulas (35)–(37) for two particular structures:

(1) The $F1/S/$Vacuum structure: $\rho_2=0$; then $\Gamma_{2,p,a,s}=0$, $\Gamma^{+}_{p,a,s}=\Gamma_{1,p,a,s}$ and $\Gamma^{-}_{1,p,a,s}=\Gamma_{1,p,a,s}$ everywhere. The approximate formulas become

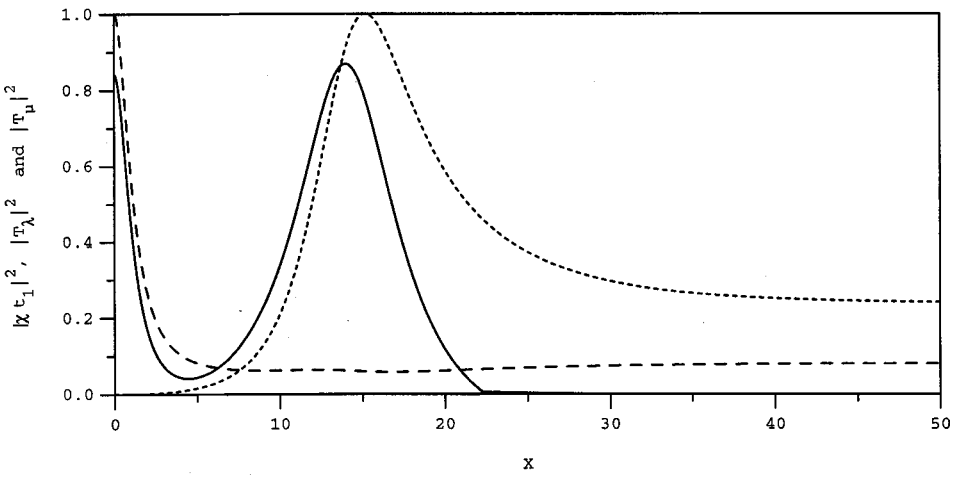


FIG. 2. $|\chi t_1|^2$ (solid), $|T_\lambda|^2$ (dotted) and $|T_\mu|^2$ (dashed) vs $X=k_1 d$ at $\theta_1=29°$ showing a shift of the peaks positions.

$$r_{1,p,a,s}=\frac{X-X_{p,a,s}-i\Gamma_{1,p,a,s}/4}{X-X_{p,a,s}+i\Gamma_{1,p,a,s}/4};$$
$$r_{2,p,a,s}=1; \quad \chi t_{1,p,a,s}=0. \tag{50}$$

This situation corresponds to the removal of the channel in fluid $F2$ (the mixing angle is $\zeta=0$). We obtain $|r_{1,p,a,s}(X)|^2=1$ from Eq. (50) and this corresponds to total reflection. The width of the resonance curve is $\Gamma_{1,p,a,s}/2$.

(2) The $F1/S/F1$ structure: one has $\rho_2=\rho_1$, $c_2=c_1$; then $\Gamma_{1,p,a,s}=\Gamma_{2,p,a,s}$. It follows that $\Gamma^+_{p,a,s}=2\Gamma_{1,p,a,s}$ and $\Gamma^-_{p,a,s}=0$. The approximate formulas become

$$r_{1,p,a,s}=r_{2,p,a,s}=\frac{X-X_{p,a,s}}{X-X_{p,a,s}+i\Gamma_{1,p,a,s}/2},$$
$$t_{1,p,a,s}=\frac{\pm i\Gamma_{1,p,a,s}/2}{X-X_{p,a,s}+i\Gamma_{1,p,a,s}/2}. \tag{51}$$

The reflection and transmission coefficients in Eqs. (51) look like those of FMU, except that here we have neglected the shifts of the reflection coefficient zeros. At the resonance frequencies, one has

$$|r_{1,p,a,s}(X_{p,a,s})|^2=|r_{2,p,a,s}(X_{p,a,s})|^2=0;$$
$$|t_{1,p,a,s}(X_{p,a,s})|^2=1. \tag{52}$$

The two channels are opened and describe identical processes in the upper and lower fluids (the mixing angle is $\zeta=\pi/4$). The width of the resonance is $\Gamma_{1,p,a,s}$. The resonance energy dissipates twice as rapidly as for the $F1/S/$Vacuum structure.

From the above analysis, we remark that, for the two structures $F1/S/$Vacuum and $F1/S/F1$, only the reflection coefficient is needed for the width evaluation. In these cases, the conditions $\Gamma_{2,p,a,s}=0$ ($F1/S/$Vacuum) and $\Gamma_{2,p,a,s}=\Gamma_{1,p,a,s}$ ($F1/S/F2$) make easier the width calculation. In the more complicated $F1/S/F2$ structure, the study of the reflection coefficient is not sufficient to obtain simultaneously the values $\Gamma_{1,p,a,s}$ and $\Gamma_{2,p,a,s}$. It becomes necessary to use, for example, the diagonal elements of the transition matrix given in the appendix [cf. Eq. (A8)]. These furnish the half width $\Gamma^+_{p,a,s}$ and the peak heights $\Gamma_{m,p,a,s}/\Gamma^+_{p,a,s}$, $(m=1,2)$ of the resonances, and then $\Gamma_{1,p,a,s}$ and $\Gamma_{2,p,a,s}$ are given by Eqs. (40).

IV. RESONANCE DECOMPOSITION IN THE ANGULAR VARIABLE

A. Resonance decomposition

With the frequency $X=k_1d$ being fixed, Ca and Cs depend on the angular variable $\overline{k_x}$. The equations $Ca(\overline{k}_{x,p,a})=0$ and $Cs(\overline{k}_{x,p,s})=0$ determine now the resonance angles $\overline{k}_{x,p,a}$ and $\overline{k}_{x,p,s}$ of the antisymmetrical and symmetrical Lamb modes of the free plate. An expansion up to the first order of the functions $Ca(\overline{k_x})$ and $Cs(\overline{k_x})$ in the vicinity of $\overline{k}_{x,p,a}$ and $\overline{k}_{x,p,s}$ leads to resonance forms like those in Sec. III, except for changes in the signs (with respect to the FMU sign convention):

$$r_1\approx r_{1,p,a,s}=\frac{\overline{k_x}-\overline{k}_{x,p,a,s}+i\gamma^-_{p,a,s}/4}{\overline{k_x}-\overline{k}_{x,p,a,s}-i\gamma^+_{p,a,s}/4}, \tag{53}$$

$$r_2\approx r_{2,p,a,s}=\frac{\overline{k_x}-\overline{k}_{x,p,a,s}-i\gamma^-_{p,a,s}/4}{\overline{k_x}-\overline{k}_{x,p,a,s}-i\gamma^+_{p,a,s}/4}, \tag{54}$$

$$\chi t_1\approx \chi t_{1,p,a,s}=\frac{\mp i\gamma^{1/2}_{1,p,a,s}\gamma^{1/2}_{2,p,a,s}/2}{\overline{k_x}-\overline{k}_{x,p,a,s}-i\gamma^+_{p,a,s}/4}. \tag{55}$$

From Eqs. (53)–(55), the phase expressions are easily obtained and are not reported here. In the above expressions,

$$\gamma_{m,p,a,s}=\mp 2\tau_{m,p,a,s}\Big/\frac{dCa,s}{d\overline{k_x}}(\overline{k}_{x,p,a,s}), \quad (m=1,2), \tag{56}$$

and

$$\gamma^+_{p,a,s}=\gamma_{1,p,a,s}+\gamma_{2,p,a,s}, \quad \gamma^-_{p,a,s}=\gamma_{1,p,a,s}-\gamma_{2,p,a,s}. \tag{57}$$

The lower sign $(+)$ in Eqs. (55)–(56) corresponds to index s. As in the preceding section, the shift of the resonance is

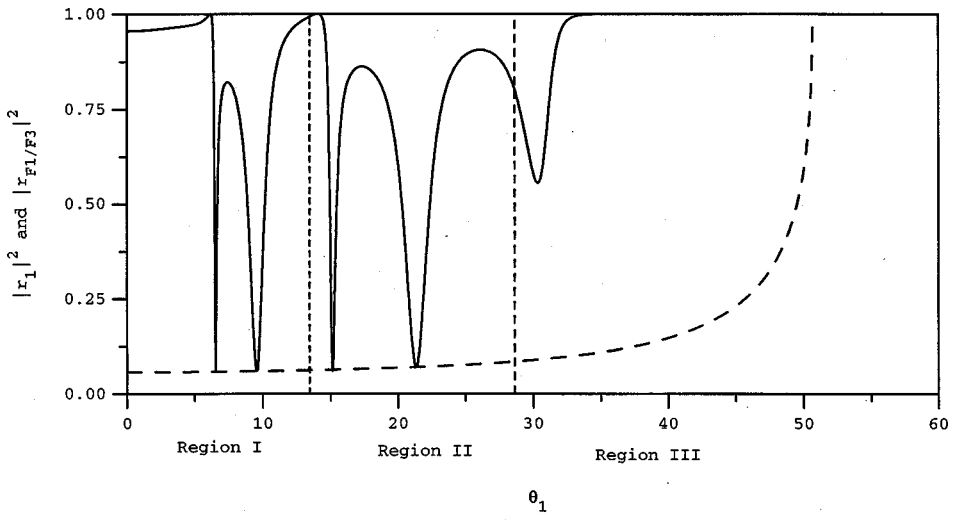


FIG. 3. $|r_1|^2$ (solid) and $|r_{F1/F2}|^2$ (dashed) vs θ_1 for $X=20$ showing the relation between the two functions. In region III $|r_1|^2$ has a specific behavior.

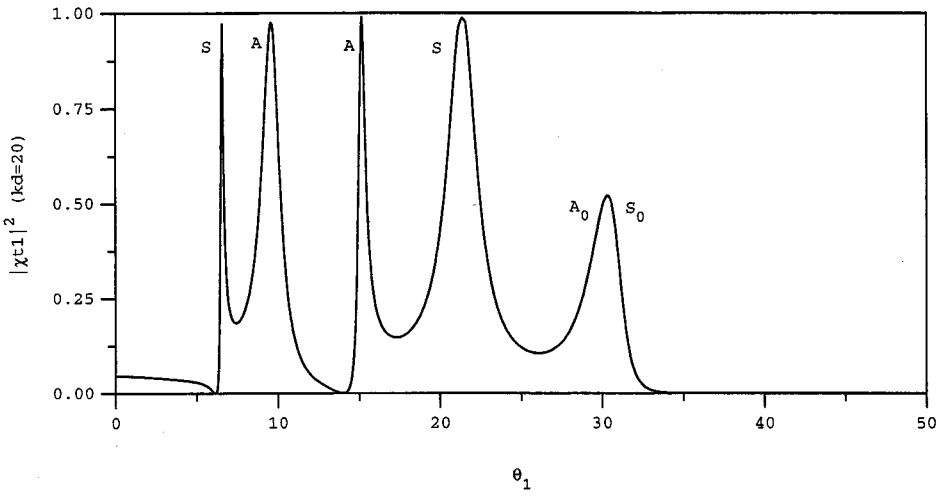


FIG. 4. $|\chi t_1|^2$ vs θ_1 for $X=20$.

neglected, and the separation between antisymmetrical and symmetrical modes appears again in these resonance decompositions. In the present case, Eqs. (41)–(44) are again valid if Γ is replaced by γ.

With the approximate forms obtained above for the r_1, r_2 and χt_1 coefficients, we have

$$\lambda_{p,a} \approx \frac{\overline{k_x} - \overline{k}_{x,p,a} + i\gamma^+_{p,a}/4}{\overline{k_x} - \overline{k}_{x,p,a} - i\gamma^+_{p,a}/4} \quad \text{and} \quad \mu_{p,a} \approx 1, \tag{58}$$

$$\lambda_{p,s} \approx 1 \quad \text{and} \quad \mu_{p,s} \approx \frac{\overline{k_x} - \overline{k}_{x,p,s} + i\gamma^+_{p,s}/4}{\overline{k_x} - \overline{k}_{x,p,s} - i\gamma^+_{p,s}/4}. \tag{59}$$

Thus, in the angular variable, the separation between antisymmetrical and symmetrical modes by the eigenvalues holds again.

B. Analysis of the overlapping resonances

Figure 3 gives the shape of $|r_1|^2$ for a fixed value of the frequency ($X=20$), the angle being the variable. Two cases appear. (1) The resonances lying in regions I and II all present troughs of the same amplitude value as those for the squared modulus of the $F1/F2$ reflection coefficient. This shows that the approximations given by Eqs. (43)–(44) and (53)–(55) are valid. (2) In region III, the approximation given by Eqs. (43)–(44) and (53)–(55) are not valid because of the overlapping of the Rayleigh type modes.

The same phenomenon described above occurs in region III when $|\chi t_1|^2$ is plotted (Fig. 4). If the squared moduli of the transition terms $|T_\lambda|^2$ and $|T_\mu|^2$ are plotted (Fig. 5) the selection between antisymmetrical and symmetrical modes is clear. This confirms the analysis of Sec. III A on the separation between these two types of modes by use of the eigenvalues. The interesting case lies again in region III, where two resonance curves overlap near the peak observed in Fig. 4 (around 30°). Each curve describes a Rayleigh type mode as indicated by the resonance positions which are slightly different from the only one position given by $|\chi t_1|^2$. It can be shown that this peak in $|\chi t_1|^2$ in the third region tends to vanish for increasing values of X (see Fig. 6, for $X=50$), while the transition terms curves always indicate the position and width of two modes (see Fig. 7 and discussion below).

Rigorously, at very high frequencies, the resonance decomposition must be reconsidered separately in each of the three regions.

In regions I and II, only Lamb type modes ($p \geqslant 1$) are encountered. It is straightforward to show that

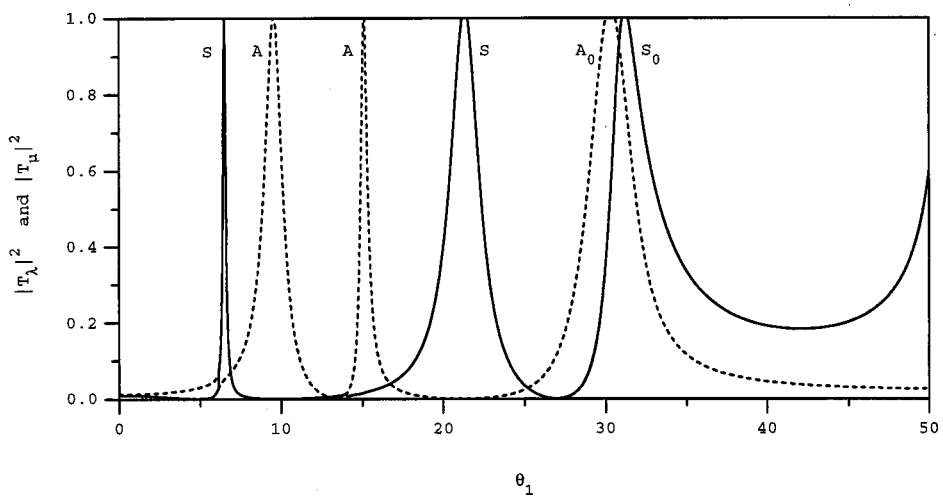


FIG. 5. $|T_\lambda|^2$ (solid) and $|T_\mu|^2$ (dotted) vs θ_1 for $X=20$ showing the separation between antisymmetrical and symmetrical modes in the angular variable. Compare to Fig. 4.

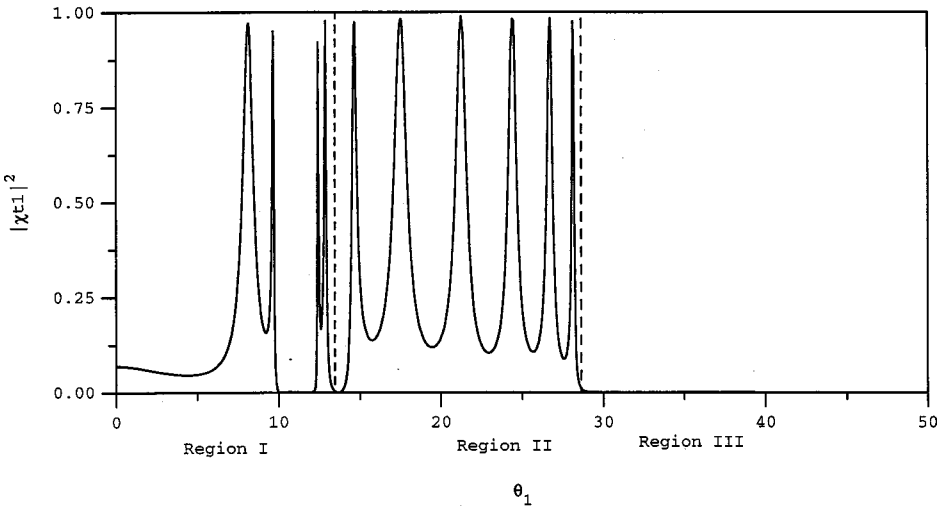


FIG. 6. $|\chi t_1|^2$ vs θ_1 for $X=50$. In region III, the transmission coefficient modulus is null.

$$\lim_{X\to+\infty} \gamma_{m,p,a,s}(X)=0, \quad (m=1,2), \tag{60}$$

which means that the resonances widths tend to zero. In addition, the number of modes increases with frequency and their accumulation point is $\overline{k_x}=n_T$.

In region III, only Rayleigh type modes ($p=0$) exist. The approximate expressions given in Sec. IV A cannot be used. One must first evaluate the exact reflection and transmission coefficients limits for $X\to+\infty$, and, only after that, one can deduce therefrom their resonant forms. When $X\to+\infty$, one has $Ca(X)\to -iR$ and $Cs(X)\to +iR$, where

$$R\equiv R(\overline{k_x})=(n_T^2-2\overline{k_x^2})^2-4\overline{k_x^2}(\overline{k_x^2}-n_L^2)^{1/2}(\overline{k_x^2}-n_T^2)^{1/2} \tag{61}$$

is known as the Rayleigh function. It follows that

$$\lim_{X\to+\infty} r_1=\frac{R-i\tau_1''}{R+i\tau_1''}=r_{F1/S} \quad \text{and} \quad \lim_{X\to+\infty} r_2=\frac{R-i\tau_2''}{R+i\tau_2''}=r_{F2/S}, \tag{62}$$

with

$$\tau_m''=\frac{\rho_m}{\rho}n_T^4\frac{(\overline{k_x^2}-n_L^2)^{1/2}}{(1-\overline{k_x^2})^{1/2}}, \quad (m=1,2). \tag{63}$$

$r_{F1/S}$ (resp. $r_{F2/S}$) represents the $F1/S$ (resp. $F2/S$) reflection coefficient. We have also

$$\lim_{X\to+\infty} \chi t_1=0. \tag{64}$$

Equations (62) and (64) mean that, at very high frequencies, the interfaces of the $F1/S/F2$ structure become uncoupled (S can be seen now as a semi-infinite medium). Equation (64) is then consistent with two facts: (1) the system behaves like two independent F/S subsystems; (2) a generalized Rayleigh wave radiates energy from the solid into the fluid, that is, there is no transmission phenomenon. The above limits do not indicate if, among the two Rayleigh type modes, it is the A_0 or the S_0 which is linked to the $F1/S$ or the $F2/S$ Rayleigh wave. The answer to this question is provided by the behavior of the eigenvalues. One finds that

$$\lim_{X\to+\infty} \lambda=r_{F1/S} \quad \text{and} \quad \lim_{X\to+\infty} \mu=r_{F2/S}. \tag{65}$$

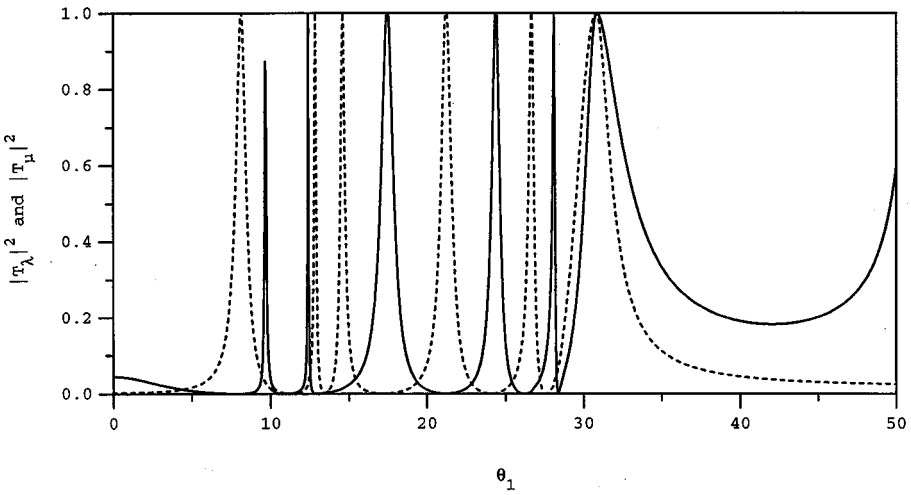


FIG. 7. $|T_\lambda|^2$ (solid) and $|T_\mu|^2$ (dotted) vs θ_1 for $X=50$. In region III, the transition terms indicate an antisymmetrical and a symmetrical mode, contrary to Fig. 6.

So, A_0 (λ eigenvalue) tends toward the Rayleigh wave of the $F1/S$ interface, and S_0 (μ eigenvalue) tends toward the Rayleigh wave of the $F2/S$ interface.

The resonant expansion of $r_{F1/S}$ and $r_{F2/S}$ in the neighborhood of the Rayleigh wave number $\overline{k_R}=\sin\theta_R$ [$\overline{k_R}$ is the real root of the equation $R(\overline{k_x})=0$, and θ_R the Rayleigh angle] can be written[8,13]

$$r_{Fm/S}\approx\frac{\overline{k_x}-\overline{k_R}+i\,\gamma_{m,0}/2}{\overline{k_x}-\overline{k_R}-i\,\gamma_{m,0}/2}, \tag{66}$$

with

$$\gamma_{m,0}=-2\,\tau''_m\Bigg/\left(\frac{dR}{d\overline{k_x}}(\overline{k_R})\right)\quad(m=1,2). \tag{67}$$

As previously discussed, $|\chi t_1|^2$ does not provide an estimation of the resonance positions and widths near the Rayleigh angle θ_R. Because of the overlapping between the A_0 and S_0 modes, the amplitude of the peak in region III tends to zero for increasing values of X (Fig. 6). The estimation of both the resonance position and the width at such incidence angles becomes possible if the eigenvalues are used. As shown in Fig. 7, for $X=50$, the two transition terms $|T_\lambda|^2$ and $|T_\mu|^2$ present separately one peak for each Rayleigh type mode and then overlap. The resonance positions are almost the same, as stated by Eq. (66) and located at $\overline{k_x}\approx0.513$, that is at $\theta_1\approx30.8°$ which is the Rayleigh angle. The widths are determined from the curves that give $\gamma_{1,0}\approx0.037$ and $\gamma_{2,0}\approx0.050$. In comparison, a computation of the complex root of the denominator of $r_{F1/S}$ (resp. $r_{F2/S}$), Eqs. (62), gives $\overline{k_x}=0.510+i0.015$ (resp. $\overline{k_x}=0.504+i0.027$).

As can be seen in Figs. 5 and 7, in region III, the curves for transition do not have the shape of Breit–Wigner curves, especially in the case of $|T_\mu|^2$. Nevertheless, the width of this curve provides a good agreement with the exact calculations of the reflection coefficient poles. As the value of the incident angle goes beyond the Rayleigh angle, approaching the third critical angle, the curve increases again, indicating the presence of a resonance phenomenon. However, this resonance (infinite peak), occurring after the third critical angle, does not enter the frame of the two channel theory as discussed here and has not been investigated in this paper.

V. CONCLUSION

We have presented both exact and approximate formalisms for the interaction of acoustic plane waves with asymmetrically fluid-loaded elastic isotropic plates ($F1/S/F2$ structures). This constitutes an extension of the FMU resonance theory of symmetrically fluid-loaded plates ($F1/S/F1$ structures).

(1) A 2×2 scattering $\mathbb{S}$-matrix is constructed; its diagonal elements are the reflection coefficients of the system and the off-diagonal terms are the transmission coefficients.
(2) Assuming light fluid loading, resonant approximations of the matrix elements, in the frequency or the angular variable, show that the partial widths concept of the two channel theory allows a fruitful description of the scattering and the FMU results appear as particular cases. In particular, the approximations show the independence of the two partial widths due to the light fluids loading. If the fluids are not light, the approximations are no longer valid because they do not take into account the coupling that exists between the fluids.
(3) The eigenvalues of the $\mathbb{S}$-matrix play an important role as shown by the numerical computations with fluids (water a and glycerine) which cannot be considered as light when compared to the aluminum plate: (a) they allow a discrimination between antisymmetrical and symmetrical modes, (b) they give a decomposition for the reflection coefficient (in the case of identical fluids) which is different from the FMU results, (c) they provide a new way to analyze the overlapping phenomenon and to characterize the Rayleigh modes.
(4) In the case of the symmetrically loaded plate, the use of the two channel resonance formalism allows, via a study of the eigenvalues, the description of all the resonances, with no need for a symbolic summation as can be found in FMU or Freedman works.

As the method proposed in this paper describes correctly all the resonances of the immersed plate (even the overlapping phenomenon), it will be suitable for further analysis of more complicated structures, for instance, (i) a stratified or an anisotropic elastic immersed plate, or a solid layer sandwiched between dissimilar materials, (ii) the same as the preceding cases, with viscoelastic effects included.[14,15] In this case, the energy conservation, self-contained in the unitarity of the scattering matrix no longer holds, and it would be necessary to take into account the energy absorption by the plate to deal with matrices of upper dimensions, (iii) cylindrical or spherical shells loaded with different fluids inside and outside, in order to explain the repulsion phenomenon between the lower modes, and, possibly, extract the parameter of the inside fluid from the knowledge of the relation between resonant scattering and dispersion of propagating plate waves.

APPENDIX: THE TWO CHANNELS RESONANCE FORMALISM

For the purpose of resonance analysis, the $\mathbb{S}$-matrix must be diagonalized. We then write

$$\mathbb{S}=\mathbb{R}\mathbb{S}^d\mathbb{R}^t, \tag{A1}$$

with

$$\mathbb{S}^d=\begin{pmatrix}\lambda & 0\\ 0 & \mu\end{pmatrix}\quad\text{and}\quad\mathbb{R}=\begin{pmatrix}\cos\zeta & -\sin\zeta\\ \sin\zeta & \cos\zeta\end{pmatrix}. \tag{A2}$$

$\mathbb{S}^d$ is a diagonal matrix. $\mathbb{R}$ is an orthogonal real matrix and $\mathbb{R}\mathbb{R}^t=\mathbb{I}$ (t indicates transposition). ζ is called the mixing angle. The functions $\lambda=e^{i\delta_\lambda}$ and $\mu=e^{i\delta_\mu}$ are the eigenvalues of $\mathbb{S}$. The functions $\delta_\lambda=\delta_\lambda(w)$ and $\delta_\mu=\delta_\mu(w)$ are the real eigenphases. They depend on the real dimensionless variable w which can represent, in the matrix elements Eq. (25), either the frequency variable $X=k_1d$, or the angular variable $\overline{k_x}=\sin\theta_1$. A resonance corresponds to the case

where one of the eigenphases, δ_λ for example, varies rapidly as a function of w around a value w_p, while the other eigenphase δ_μ stays constant (practically equal to zero). In these conditions, in the vicinity of a resonance position w_p, we get

$$\lambda \approx \frac{w-w_p-i\Gamma^+/4}{w-w_p+i\Gamma^+/4} \quad \text{and} \quad \mu \approx 1. \tag{A3}$$

The first equation of (A3) has the standard resonance form and is called a Breit–Wigner approximation. The term $w_p - i\Gamma^+/4$ represents the pole of a reflection or transmission coefficient. Let us introduce the partial widths Γ_1 and Γ_2 defined as

$$\Gamma^+ \cos^2 \zeta = \Gamma_1 \quad \text{and} \quad \Gamma^+ \sin^2 \zeta = \Gamma_2. \tag{A4}$$

This leads to the relation

$$\Gamma^+ = \Gamma_1 + \Gamma_2, \tag{A5}$$

which expresses the width conservation law and, by extension, the energy conservation law of the resonance. The mixing angle ζ depends on the relative intensity of the partial widths. For example, if the two widths have the same intensity then $\zeta = \pi/4$. With the help of Eqs. (A3), (A4) and (A5), the $\mathbb{S}$ matrix can be written in the vicinity of a resonance:

$$\mathbb{S} = \begin{pmatrix} \dfrac{w-w_p-i\Gamma^-/4}{w-w_p+i\Gamma^+/4} & \dfrac{-i\Gamma_1^{1/2}\Gamma_2^{1/2}/2}{w-w_p+i\Gamma^+/4} \\ \dfrac{-i\Gamma_1^{1/2}\Gamma_2^{1/2}/2}{w-w_p+i\Gamma^+/4} & \dfrac{w-w_p+i\Gamma^-/4}{w-w_p+i\Gamma^+/4} \end{pmatrix}, \tag{A6}$$

with $\Gamma^- = \Gamma_1 - \Gamma_2$. Of frequent use in resonance theory is the transition matrix $\mathbb{T}$ defined as

$$\mathbb{T} = \frac{1}{2i}(\mathbb{S} - \mathbb{I}). \tag{A7}$$

Its elements take the form

$$(\mathbb{T})_{mn} = \frac{-i\Gamma_m^{1/2}\Gamma_n^{1/2}/2}{w-w_p+i\Gamma^+/4}, \quad (m,n=1,2), \tag{A8}$$

called the Breit–Wigner formula for a resonance in the multichannel theory.[16] This from expresses that the same resonance appears in the matrix elements with an intensity depending on the partial widths Γ_1 and Γ_2. Thus, at the resonance where $w = w_p$, the modulus of $(\mathbb{T})_{mm}$ equals Γ_m/Γ^+ and allows us to determine Γ_m as long as the total width Γ^+ is known.

Let us apply now to the $\mathbb{S}$ matrix the so-called Cayley transform

$$\mathbb{K} = i(\mathbb{I} - \mathbb{S})(\mathbb{I} + \mathbb{S})^{-1}, \tag{A9}$$

in which $\mathbb{K}$ is an Hermitian (or self adjoint) matrix. This matrix can be transformed into a diagonal one, $\mathbb{K}^d$, trough the relation $\mathbb{K} = \mathbb{R}\mathbb{K}^d\mathbb{R}^t$. We obtain

$$\mathbb{K}^d = \begin{pmatrix} \tan(\delta_\lambda/2) & 0 \\ 0 & \tan(\delta_\mu/2) \end{pmatrix}, \tag{A10}$$

and, from Eqs. (A3),

$$\delta_\lambda = 2a \tan \frac{\Gamma^+/4}{w_p - w} \quad \text{and} \quad \delta_\mu \approx 0. \tag{A11}$$

It is instructive to compare these eigenphases to the phases of the diagonal elements S_{mm} $(m=1,2)$ of Eq. (A6). If we write $S_{mm} = |S_{mm}| e^{i\varphi_{mm}}$, it is straightforward to show that

$$\varphi_{mm} = -a \tan \frac{\Gamma^+/4}{w-w_p} \mp a \tan \frac{\Gamma^-/4}{w-w_p}, \tag{A12}$$

where the $-$ sign in front of the second term of the right member corresponds to $m=2$. The phases φ_{mm} depend on both Γ^+ and Γ^- while the eigenphase δ_λ depends only on Γ^+. In addition, Eq. (A12) shows that the phase difference $\varphi_{11} - \varphi_{22}$ is equal to $2a \tan[\Gamma^-/4)/(w_p - w)]$.

[1] See, e.g., D. E. Chimenti and S. Rokhlin, ''Relationship between leaky Lamb modes and reflection coefficient zeros for a fluid coupled elastic layer,'' J. Acoust. Soc. Am. **88**, 1603–1611 (1990); S. I. Rokhlin, D. E. Chimenti, and A. H. Nayfeh, ''On the topology of the complex wave spectrum in a fluid-coupled elastic layer,'' *ibid.*, **85**, 1074–1080 (1989).

[2] O. Lenoir, J. Duclos, J.-M. Conoir, and J.-L. Izbicki, ''Study of Lamb waves based upon the frequency and angular derivatives of the phase of the reflection coefficient,'' J. Acoust. Soc. Am. **94**, 330–343 (1993).

[3] A. Freedman, ''Effects of fluid-loading on Lamb modes spectra,'' J. Acoust. Soc. Am. **99**, 3488–3496 (1996).

[4] J.-P. Sessarego, J. Sageloli, C. Gazanhes, and H. Überall, ''Two Scholte–Stoneley waves on doubly fluid-loaded plates and shells,'' J. Acoust. Soc. Am. **101**, 135–142 (1994).

[5] H. Überall, B. Hosten, M. Deschamps, and A. Gérard, ''Repulsion of phase-velocity dispersion curves and the nature of plate vibrations,'' J. Acoust. Soc. Am. **96**, 908–917 (1994).

[6] X. L. Bao, H. Franklin, P. K. Raju, and H. Überall, ''The splitting of dispersion curves for plates fluid-loaded on both sides,'' J. Acoust. Soc. Am. **102**, 1246–1248 (1997).

[7] X. L. Bao, H. Franklin, P. K. Raju, H. Überall, and O. Poncelet, ''Fluid-borne and Lamb-type waves on elastic plates in contact with two different fluids,'' Acustica **84**, 823–829 (1998).

[8] R. Fiorito, W. Madigosky, and H. Überall, ''Resonance theory of acoustic waves interacting with an elastic plate,'' J. Acoust. Soc. Am. **66**, 1857–1866 (1979).

[9] A. Freedman, ''On the Overlapping resonances concept of acoustic transmission through an elastic plate, I: An examination of properties,'' J. Sound Vib. **82**, 181–195 (1982).

[10] A. Freedman, ''On the Overlapping resonances concept of acoustic transmission through an elastic plate, II: Numerical examples and physical implications,'' J. Sound Vib. **82**, 197–213 (1982).

[11] L. M. Brekhovskikh, *Waves in Layered Media* (Academic, New York, 1980).

[12] L. E. Pitts, T. J. Plona, and W. G. Mayer, ''Theoretical similarities of Rayleigh and Lamb modes of vibration,'' J. Acoust. Soc. Am. **60**, 374–377 (1976).

[13] J. Duclos, J.-L. Izbicki, O. Lenoir, and J.-M. Conoir, ''Resonant formalism for the liquid-solid interface Rayleigh mode,'' Acta Acust. (Beijing) **2**, 375–378 (1994).

[14] B. Hosten, ''Bulk heterogeneous plane waves propagation through viscoelastic plates and stratified media with large values of frequency domain,'' Ultrasonics **29**, 445–449 (1991).

[15] A. I. Lavrentyev and S. I. Rokhlin, ''Anomalous attenuation effect on reflectivity of an ultrasonic wave from a thin layer between dissimilar materials,'' J. Acoust. Soc. Am. **101**, 3405–3414 (1997).

[16] A. Bohm, *Quantum Mechanics, Foundations and Applications* (Springer, New York 1993), Chap. XX.

Three-dimensional vibration analysis of a homogeneous transversely isotropic thermoelastic cylindrical panel

J. N. Sharma[a)]
Department of Mathematics, Regional Engineering College Hamirpur (H.P.)—177 005, India

(Received 26 October 1999; accepted for publication 30 March 2001)

In this paper, based on three-dimensional thermoelasticity, an exact analysis of the free vibrations of a simply supported, homogeneous, transversely isotropic, cylindrical panel is presented. Three displacement potential functions are introduced so that the equations of motion and heat conduction are uncoupled and simplified. It is noticed that a purely transverse mode is independent of temperature change and the rest of the motion. The equations for free vibration problems are further reduced to four second-order ordinary differential equations, after expanding the potential and temperature functions with an orthogonal series. A modified Bessel function solution with complex arguments is directly used for complex eigenvalues. To clarify the developed method and compare the results to the existing ones, numerical examples are presented. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1378350]

PACS numbers: 43.40.Ey [PJR]

I. INTRODUCTION

The cylindrical panels are frequently used as structural components and their vibration characteristics are obviously important for practical design. A large number of previous studies on the vibration of isotropic curved panels have been performed.[1–7] For the last several decades, anisotropic materials have been widely used in many areas because of their excellent static and dynamic behavior and low strength-to-weight ratio. Many researchers[8–12] have devoted themselves to the vibration analysis of anisotropic cylindrical shell and/or panels, basing their works mainly on two-dimensional shell theories. Soldatos and Hadhgeorgiou[13] used an iterative approach to predict the frequencies of isotropic cylindrical shells and panels based on the governing equations of three-dimensional elasticity. Leissa[14] studied free vibrations of thick hollow cylinders by the Ritz method. Jing[15] employed the perturbation method to study three-dimensional vibrations of fiber reinforced composite laminated cylindrical shells, while Fan and Ding[16] and Ye and Soldatos[17] used a state space method to analyze laminated orthotropic cylindrical shells and cross-ply cylindrical panels, respectively. Mirsky[18] was the first to exactly study axisymmetric free vibrations of orthotropic cylindrical shells with infinite length based on three-dimensional elastic theory by employing the Frobenius method, and he[19] also obtained a Bessel function solution for simply supported transversely isotropic cylindrical shells. However, he suggested shell theory[20] should be used for the case when complex eigenvalues emerge. Chau[21] employed a similar method[19] to study the free vibration problems of simply supported transversely isotropic cylinders, but he avoided the complex eigenvalue case by considering the specific transversely isotropic material sapphire. Chen *et al.*[22] presented an exact analysis of the free vibrations of a simply supported, transversely isotropic cylindrical panel by using a modified Bessel function solution with complex arguments for the case of complex eigenvalues, and clarified the correctness and effectiveness of their method by presenting numerical examples and comparing their results with existing papers.

Thick composite cylindrical shells can be used in applications involving aerospace, offshore, and submarine structures, pressure vessels, civil engineering structures, chemical pipes, and even automotive suspension components. These structures can be easily exposed to a variety of temperature fields in different environments. In high-temperature applications, thermal stresses, which are induced from heat temperature buildup and cooling processes, may rise above the ultimate strength and leads to unexpected failures. Thus, the importance of thermal stresses in causing structural damage and changes in a functionality of the structure is well recognized whenever thermal environments are involved. Therefore, the capability to predict elastodynamic stresses induced by sudden thermal loading in composite structures is essential for the proper and safe design and the knowledge of its response during service in these severe thermal applications.

In the case of suddenly applied thermal loading, thermal deformation and the role of inertia become larger. Since thermal stresses change very rapidly, the static analysis cannot capture its behavior. This dynamic thermoelastic stress response is significant and leads to the propagation of elastic stress waves in the solid.

The theory of elastic wave propagation in anisotropic solids is well known.[23,24] Chadwick and Seet[25] studied the propagation of plane harmonic waves in transversely isotropic and Chadwick[26] in homogeneous anisotropic heat conducting solids. Singh and Sharma[27,28] and Sharma and Sidhu[29] studied the propagation of plane harmonic thermoelastic waves in homogeneous, transversely isotropic, cubic crystals and anisotropic materials in the context of generalized thermoelasticity.

In the present paper, we choose three displacement func-

a)Electronic mail: jns@patra.recham.ernet.in

 0001-4966/2001/110(1)/254/6/$18.00

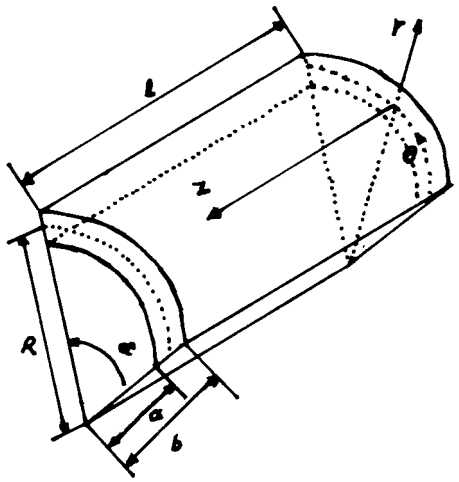


FIG. 1. Geometry of the problem.

tions to represent three displacement components on the basis of three-dimensional thermoelasticity for transversely isotropic media. The fundamental equations are then simplified and the free vibration solution for a simply supported cylindrical panel is obtained by using modified Bessel functions with complex arguments. It can be verified that the resulting frequency equations and expressions of stresses, temperature change, and displacements are all in real forms. Numerical examples are presented and compared to relevant publications.

II. FORMULATION AND SOLUTION

We consider a homogeneous transversely isotropic, thermally conducting elastic cylindrical panel of length ''L'' having inner and outer radii ''a'' and ''b'' respectively, with central angle α at uniform temperature T_0 in the undisturbed state initially. The geometry of the problem is shown in Fig. 1. The governing field equations of motion and heat conduction in the absence of body forces and heat sources are

$$\begin{aligned}
&\sigma_{rr,r}+\frac{1}{r}\sigma_{r\theta,\theta}+\sigma_{rz,z}+\frac{\sigma_{rr}-\sigma_{\theta\theta}}{r}=\rho\ddot{u}_r,\\
&\sigma_{r\theta,r}+\frac{1}{r}\sigma_{\theta\theta,\theta}+\sigma_{\theta z,z}+\frac{2\sigma_{r\theta}}{r}=\rho\ddot{u}_\theta,\\
&\sigma_{rz,r}+\frac{1}{r}\sigma_{\theta z,\theta}+\sigma_{zz,z}+\frac{\sigma_{rz}}{r}=\rho\ddot{u}_z,\\
&K_1\left(T_{,rr}+\frac{1}{r}T_{,r}+\frac{1}{r^2}T_{,\theta\theta}\right)+K_3T_{,zz}-\rho C_e\dot{T}\\
&\quad=T_0\frac{\partial}{\partial t}[\beta_1(e_{rr}+e_{\theta\theta})+\beta_3e_{zz}],
\end{aligned}\tag{1}$$

where

$$\begin{aligned}
&\sigma_{rr}=c_{11}e_{rr}+c_{12}e_{\theta\theta}+c_{13}e_{zz}-\beta_1T,\quad \sigma_{r\theta}=(c_{11}-c_{12})e_{r\theta},\\
&\sigma_{\theta\theta}=c_{12}e_{rr}+c_{11}e_{\theta\theta}+c_{13}e_{zz}-\beta_1T,\quad \sigma_{\theta z}=c_{44}e_{\theta z},\\
&\sigma_{zz}=c_{13}e_{rr}+c_{13}e_{\theta\theta}+c_{33}e_{zz}-\beta_3T,\quad \sigma_{rz}=c_{44}e_{rz},
\end{aligned}\tag{2}$$

$$\begin{aligned}
&e_{rr}=\frac{\partial u_r}{\partial r},\quad e_{\theta\theta}=\frac{1}{r}\frac{\partial u_\theta}{\partial\theta}+\frac{u_r}{r},\quad e_{zz}=\frac{\partial u_z}{\partial z},\\
&e_{r\theta}=\frac{1}{2}\left(\frac{1}{r}\frac{\partial u_r}{\partial\theta}+\frac{\partial u_\theta}{\partial r}-\frac{u_\theta}{r}\right),\quad e_{rz}=\frac{1}{2}\left(\frac{\partial u_r}{\partial z}+\frac{\partial u_z}{\partial r}\right),\\
&e_{\theta z}=\frac{1}{2}\left(\frac{\partial u_\theta}{\partial z}+\frac{1}{r}\frac{\partial u_z}{\partial\theta}\right),
\end{aligned}\tag{3}$$

$$\beta_1=(c_{11}+c_{12})\alpha_1+c_{13}\alpha_3,\quad \beta_3=2c_{13}\alpha_1+c_{33}\alpha_3.$$

Here $\vec{u}=(u_r,u_\theta,u_z)$ is the displacement vector, $T(r,\theta,z,t)$ is the temperature change; c_{11}, c_{12}, c_{13}, c_{33}, and c_{44} are five independent isothermal elasticities; α_1, α_3 and K_1, K_3 are, respectively, the coefficients of linear thermal expansion and thermal conductivities along and perpendicular to the axis of symmetry; ρ and C_e are the density and specific heat at constant strain, respectively. The comma notation is used for spatial derivatives and the superposed dot represents time differentiation.

It can be proved thermodynamically[25] that $K_1>0$, $K_3>0$, and, of course, $\rho>0$, $T_0>0$. We assume in addition that $C_e>0$ and that the isothermal linear elasticities are components of a positive definite fourth-order tensor. The necessary and sufficient conditions for the satisfaction of the latter requirement are

$$\begin{aligned}
&c_{11}>0,\quad c_{11}>c_{12},\quad c_{11}^2>c_{12}^2,\quad c_{44}>0,\\
&c_{33}(c_{11}+c_{12})>c_{13}^2.
\end{aligned}\tag{4}$$

To solve Eqs. (1), we take[30]

$$u_r=\frac{1}{r}\psi_{,\theta}-\phi_{,r},\quad u_\theta=-\frac{1}{r}\phi_{,\theta}-\psi_{,r},\quad u_z=\chi_{,z}.\tag{5}$$

Upon using Eqs. (5) in Eqs. (1), we find that ϕ, χ, ψ, and T satisfy the equations

$$\left(c_{11}\nabla_1^2+c_{44}\frac{\partial^2}{\partial z^2}-\rho\frac{\partial^2}{\partial t^2}\right)\phi-(c_{13}+c_{44})\frac{\partial^2\chi}{\partial z^2}=\beta_1T,\tag{6a}$$

$$-(c_{13}+c_{44})\nabla_1^2\phi+\left(c_{44}\nabla_1^2+c_{33}\frac{\partial^2}{\partial z^2}-\rho\frac{\partial^2}{\partial t^2}\right)\chi=\beta_3T,\tag{6b}$$

$$\left(c_{66}\nabla_1^2+c_{44}\frac{\partial^2}{\partial z^2}-\rho\frac{\partial^2}{\partial t^2}\right)\psi=0,\tag{6c}$$

$$K_1\nabla_1^2T+K_3\frac{\partial^2T}{\partial z^2}-\rho C_e\frac{\partial T}{\partial t}=T_0\frac{\partial}{\partial t}\left(\beta_1\nabla_1^2\phi+\beta_3\frac{\partial^2\chi}{\partial z^2}\right),\tag{6d}$$

where $c_{66}=(c_{11}-c_{12})/2$, $\nabla_1^2=\partial^2/\partial r^2+(1/r)(\partial/\partial r)+(1/r^2)(\partial^2/\partial\theta^2)$.

Equation (6c) in the above system in ψ gives a purely transverse wave, which is not affected by the temperature. This wave is polarized in planes perpendicular to the z axis and may be referred to as an SH wave. We assume that the disturbance is time harmonic through the factor $e^{i\omega t}$ and hence the system of equations (6) becomes

$$\left(\nabla_1^2+c_2\frac{\partial^2}{\partial z^2}+\frac{\rho\omega^2}{c_{11}}\right)\phi-c_3\frac{\partial^2\chi}{\partial z^2}=\frac{\beta_1}{c_{11}}T,$$

$$-c_3\nabla_1^2\phi+\left(c_2\nabla_1^2+c_1\frac{\partial^2}{\partial z^2}+\frac{\rho\omega^2}{c_{11}}\right)\chi=\frac{\beta_1\bar{\beta}T}{c_{11}},$$

$$\nabla_1^2T+\bar{K}\frac{\partial^2T}{\partial z^2}-\frac{\rho C_e i\omega T}{K_1}=\frac{\beta_1T_0i\omega}{K_1}\left(\nabla_1^2\phi-\bar{\beta}\frac{\partial^2\chi}{\partial z^2}\right), \tag{7}$$

$$\left(c_4\nabla_1^2+c_2\frac{\partial^2}{\partial z^2}+\frac{\rho\omega^2}{c_{11}}\right)\psi=0,$$

where

$$c_1=\frac{c_{33}}{c_{11}},\quad c_2=\frac{c_{44}}{c_{11}},\quad c_3=\frac{c_{13}+c_{44}}{c_{11}},\quad c_4=\frac{c_{66}}{c_{11}},$$

$$\bar{\beta}=\frac{\beta_3}{\beta_1},\quad \bar{K}=\frac{K_3}{K_1}. \tag{8}$$

We consider the free vibration of a cylindrical panel (see Fig. 1) subjected to the simply supported thermally insulated edge boundary conditions at all of its four edges, we can write three displacement functions and the temperature change as

$$\psi(r,\theta,z,t)=\bar{\psi}(r)\sin(m\pi z)\cos(n\pi\theta/\alpha)e^{i\omega t},$$

$$\phi(r,\theta,z,t)=\bar{\phi}(r)\sin(m\pi z)\sin(n\pi\theta/\alpha)e^{i\omega t},$$

$$\chi(r,\theta,z,t)=\bar{\chi}(r)\sin(m\pi z)\sin(n\pi\theta/\alpha)e^{i\omega t}, \tag{9}$$

$$T(r,\theta,z,t)=\bar{T}(r)\sin(m\pi z)\sin(n\pi\theta/\alpha)e^{i\omega t},$$

where $r'=r/R$, $z'=z/L$, $T'=T/T_0$, $R=(a+b)/2$. The solutions are applicable to both closed hollow cylinders and open ones (panels), depending upon whether $n\pi/\alpha$ is an integer or not.

Here dashes on the right-hand side have been suppressed for convenience. The substitution of Eqs. (9) into Eqs. (7) gives

$$(\nabla_2^2+k_1^2)\bar{\psi}=0, \tag{10a}$$

$$(\nabla_2^2+g_1)\bar{\phi}+g_2\bar{\chi}=\frac{\beta_1T_0R^2\bar{T}}{c_{11}}, \tag{10b}$$

$$-c_3\nabla_2^2\bar{\phi}+c_2(\nabla_2^2+g_3)\bar{\chi}=\frac{\beta_1\bar{\beta}T_0R^2}{c_{11}}\bar{T}, \tag{10c}$$

$$(\nabla_2^2+g_4)\bar{T}=\frac{i\omega\beta_1}{K_1}(\nabla_2^2\bar{\phi}-\bar{\beta}t_L^2\bar{\chi}), \tag{10d}$$

where $\nabla_2^2=(1/r)(d/dr)[r(d/dr)]-\beta^2/r^2$ and

$$g_1=c_2(\Omega^2-t_L^2),\quad g_2=c_3t_L^2,\quad g_3=\Omega^2-\frac{c_1}{c_2}t_L^2,$$

$$g_4=\frac{-ic_2}{\chi^*}\left(\Omega^2-i\frac{\bar{K}t_L^2\chi^*}{c_2}\right),\quad \beta=n\pi/\alpha,$$

$$k_1^2=2c_{44}(\Omega^2-t_L^2)/(c_{11}-c_{12}),\quad \Omega=\frac{\omega}{\omega_s},\quad \omega_s=v_2/R, \tag{11}$$

$$t_L=\frac{m\pi R}{L},\quad v_2=\sqrt{c_{44}/\rho},\quad \omega^*=\frac{C_ec_{11}}{K_1},\quad \chi^*=\omega/\omega^*.$$

Here v_2 is the velocity of elastic wave in the cylindrical panel and ω^* is the characteristic frequency of the panel.

Equation (10a) is a Bessel equation with its possible solution as

$$\bar{\psi}=\begin{cases}A_4J_\beta(k_1r)+B_4Y_\beta(k_1r), & k_1^2>0,\\ A_4r^\beta+B_4r^{-\beta}, & k_1^2=0,\\ A_4I_\beta(k_1'r)+B_4K_\beta(k_1'r), & k_1^2<0,\end{cases} \tag{12}$$

where $k_1'^2=-k_1^2$, J_β and Y_β are, respectively, Bessel functions of the first and second kinds, while I_β and K_β are modified Bessel functions of the first and second kinds, respectively, and A_4, and B_4 are two arbitrary constants. Generally, $k_1^2\neq0$ so that the specific situation $k_1^2=0$ will not be discussed in the following analysis. We go on with our derivation by taking the form of $\bar{\psi}$ for $k_1^2<0$; the derivation for $k_1^2>0$ is obviously similar.

From the rest of equations (10), one can obtain

$$(\nabla_2^2-m_1^2)(\nabla_2^2-m_2^2)(\nabla_2^2-m_3^2)\bar{\phi}=0, \tag{13}$$

where m_i^2, $i=1,2,3$ are the roots of the equation [$\mathrm{Re}(m_i)\geqslant0$]

$$m^6+Am^4+Bm^2+C=0, \tag{14}$$

in which the coefficients A, B, and C are given by

$$A=g_1+g_3+g_4+\frac{c_3g_2}{c_2}-i\epsilon c_2\Omega^2/\chi^*,$$

$$B=g_1g_3+g_1g_4+g_3g_4+\frac{c_3}{c_2}g_2g_4+\frac{i\epsilon_1\Omega^2\bar{\beta}}{\chi^*}\times\left(g_2-\frac{c_2g_3}{\beta}+(c_3+\bar{\beta})t_L^2\right), \tag{15}$$

$$C=g_1g_3g_4+i\epsilon_1\Omega^2t_L^2\bar{\beta}^2g_1/\chi^*,$$

$$\epsilon_1=\beta_1^2T_0/\rho C_ec_{11}.$$

For uncoupled thermoelasticity ($\epsilon_1=0$), Eq. (14) reduces to

$$m^4+\bar{B}m^2+\bar{C}=0, \tag{16}$$

$$m^2+g_4=0, \tag{17}$$

where $\bar{B}=g_1+g_3+c_3/c_2g_2$, $\bar{C}=g_1g_3$.

Equation (16) is similar to one as obtained and discussed by Chen *et al.*[22] in case of elastokinetics and Eq. (17) corresponds to the thermal mode (T mode). The functions $\bar{\phi}$, $\bar{\chi}$, and $\bar{T}$ are obtained as

$$\overline{\phi}(r)=\sum_{i=0}^{3}\left[A_iI_\beta(m_ir)+B_iK_\beta(m_ir)\right],$$

$$\overline{\chi}(r)=\sum_{i=0}^{3}a_i\left[A_iI_\beta(m_ir)+B_iK_\beta(m_ir)\right], \tag{18}$$

$$\overline{T}(r)=\sum_{i=0}^{3}b_i\left[A_iI_\beta(m_ir)+B_iK_\beta(m_ir)\right],$$

where

$$a_i=\frac{(c_3+\overline{\beta})m_i^2+\overline{\beta}g_1}{c_2m_i^2+c_2g_3-\overline{\beta}g_2},\quad i=1,2,3,$$

$$b_i=\frac{c_{11}}{\beta_1T_0R^2}\left(\frac{c_2m_i^4+[c_2(g_1+g_3)+c_3g_2]m_i^2+c_2g_1g_3}{c_2m_i^2+c_2g_3-\beta g_2}\right), \tag{19}$$

$$i=1,2,3.$$

Here we noticed that Eq. (14) generally possesses complex roots and hence modified Bessel functions constitute the solutions. Also, we take

$$\overline{\psi}(r)=A_4I_\beta(k_1'r)+B_4K_\beta(k_1'r).$$

III. FREQUENCY EQUATION

The displacements, temperature, and stresses are obtained as

$$\begin{aligned}
u_r&=(-\overline{\phi}'-\beta\overline{\psi}/r)\sin(m\pi z)\sin(\beta\theta)e^{i\omega t},\\
u_\theta&=(-\overline{\psi}'-\beta\overline{\phi}/r)\sin(m\pi z)\cos(\beta\theta)e^{i\omega t},\\
u_z&=\overline{\chi}t_L\cos(m\pi z)\sin(\beta\theta)e^{i\omega t},\\
T&=\overline{T}\sin(m\pi z)\sin(\beta\theta)e^{i\omega t},\\
\overline{\sigma}_{rr}&=\Bigg[\left[-\overline{\phi}''-(1-2c_4)\overline{\phi}'r^{-1}+\beta^2(1-2c_4)\overline{\phi}r^{-2}\right]\\
&\quad-2c_4\beta\left(\frac{\overline{\psi}'}{r}-\frac{\overline{\psi}}{r^2}\right)-(c_3-c_2)\overline{\chi}t_L^2-\frac{\beta_1T_0R}{c_{11}}\overline{T}\Bigg]\\
&\quad\times\sin(m\pi z)\sin(\beta\theta)e^{i\omega t},\\
\overline{\sigma}_{rz}&=\frac{c_2t_L}{2}\left(-\overline{\phi}-\frac{\beta\overline{\psi}}{r}+\overline{\chi}'\right)\cos(m\pi z)\sin(\beta\theta)e^{i\omega t},\\
\overline{\sigma}_{r\theta}&=c_4\left(\frac{-2\beta\overline{\phi}'}{r}+\frac{2\beta\overline{\phi}}{r^2}-\overline{\psi}''+\frac{\overline{\psi}'}{r^2}-\beta^2\overline{\psi}/r^2\right)\\
&\quad\times\sin(m\pi z)\cos(\beta\theta)e^{i\omega t},
\end{aligned} \tag{20, 21}$$

where a prime denotes differentiation with respect to r, $\overline{u}_i=u_i/R$, $i=r$, θ, z are three nondimensional displacements, and $\overline{\sigma}_r=\sigma_r/c_{11}$, $\overline{\sigma}_{rz}=\sigma_{rz}/c_{11}$, $\overline{\sigma}_{r\theta}=\overline{\sigma}_{r\theta}/c_{11}$ are three nondimensional stresses. Considering the traction-free insulated boundary conditions at the lower and upper surfaces $r=a,b$ and making use of Eqs. (21), one can get the free vibration equation as follows:

$$|E_{ij}|=0\quad(i,j=1,2,...,8), \tag{22}$$

where

$$\begin{aligned}
E_{11}&=-2c_4\beta[k_1'I_\beta(k_1't_1)/t_1-I_\beta(k_1't_1)/t_1^2],\\
E_{12}&=-2c_4\beta[k_1'K_\beta'(k_1't_1)/t_1-K_\beta(k_1't_1)/t_1^2],\\
E_{13}&=-m_1^2I_\beta''(m_1t_1)-\frac{1-2c_4}{t_1}m_1I_\beta'(m_1t_1)\\
&\quad+\left((c_3-c_2)t_L^2a_1+\frac{\beta_1T_0R}{c_{11}}b_1\right)I_\beta(m_1t_1),\\
E_{15}&=-m_2^2I_\beta''(m_2t_1)-\frac{1-2c_4}{t_1}m_2I_\beta'(m_2t_1)\\
&\quad+\left((c_3-c_2)t_L^2a_2+\frac{\beta_1T_0R}{c_{11}}b_2\right)I_\beta(m_2t_1),\\
E_{17}&=-m_3^2I_\beta''(m_3t_1)-\frac{1-2c_4}{t_1}m_3I_\beta'(m_3t_1)\\
&\quad+\left[(c_3-c_2)t_L^2a_3+\frac{\beta_1T_0R}{c_{11}}b_3\right]I_\beta(m_3t_1)\\
E_{21}&=\frac{\beta}{t_1}I_\beta(k_1't_1),\quad E_{23}=I_\beta(m_1t_1)-a_1m_1I_\beta'(m_1t_1),\\
E_{25}&=I_\beta(m_2t_1)-a_2m_2I_\beta'(m_2t_1),\\
E_{27}&=I_\beta(m_3t_1)-a_3m_3I_\beta'(m_3t_1),\\
E_{31}&=-k_1'^2I_\beta''(k_1't_1)+\frac{k_1'}{t_1}I_\beta'(k_1^1t_1)-\frac{\beta^2}{t_1^2}I_\beta(k_1't_1),\\
E_{33}&=-\frac{2\beta m_1}{t_1}I_\beta'(m_1t_1)-\frac{2\beta}{t_1^2}I_\beta(m_1t_1),\\
E_{35}&=\frac{-2\beta}{t_1}m_2I_\beta'(m_2t_1)+\frac{2\beta}{t_1^2}I_\beta(m_2t_1),\\
E_{37}&=\frac{-2\beta}{t_1}m_3I_\beta'(m_3t_1)+\frac{2\beta}{t_1^2}I_\beta(m_3t_1),\quad E_{41}=0,\\
E_{43}&=b_1m_1I_\beta(m_1t_1),\quad E_{45}=b_2m_2I_\beta(m_2t_1),\\
E_{47}&=b_3m_3I_\beta(m_3t_1).
\end{aligned} \tag{23}$$

Here E_{ij} $(j=2,4,6,8)$ can be obtained by just replacing a modified Bessel function of first kind in E_{ij} $(i=1,3,5,7)$ with the ones of the second kind, respectively, while E_{ij} $(j=5,6,7,8)$ can be obtained by just replacing t_1 in E_{ij} $(j=1,2,3,4)$ with t_2, respectively, where $t_1=a/R=1-t^*/2$ and $t_2=b/R=1-t^*/2$ and $t^*=(b-a)/R$ is the thickness to mean radius ratio of the panel.

For uncoupled thermoelasticity $(\epsilon_1=0)$, the above analysis reduces to one as obtained and discussed by Chen *et al.*[22] in the case that elastokinetics and thermal effects get decoupled.

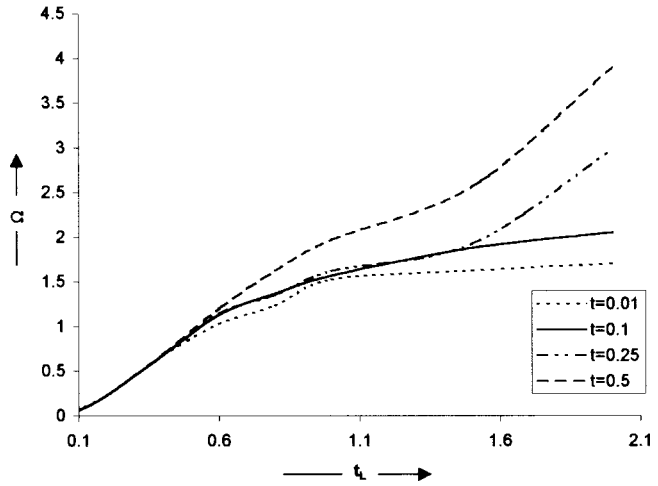


FIG. 2. The variation of the lowest frequency of a simply supported traction-free thermally insulated zinc cylindrical shell with $t_L = mR/L$ for $l = 1$.

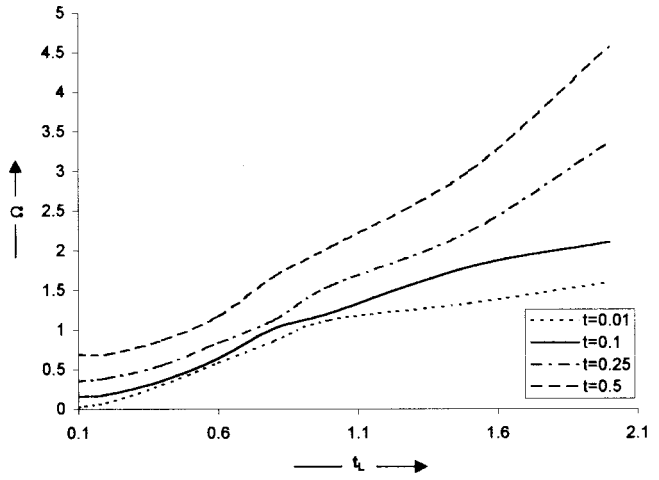


FIG. 3. The variation of the lowest frequency of a simply supported traction-free thermally insulated zinc cylindrical shell with $t_L = mR/L$ for $l = 2$.

A. Free vibration of an isotropic cylindrical panel

For isotropic materials, we have

$$c_1 = 1, \quad c_2 = \frac{\mu}{\lambda + 2\mu} = c_4, \quad c_3 = 1 - c_2, \quad \overline{\beta} = 1, \quad \overline{K} = 1.$$

where λ, μ are the Lame parameters.

Using these values in various relevant relations and equations, the above analysis will reduce to represent free vibrations of an isotropic cylindrical panel.

IV. NUMERICAL RESULTS AND DISCUSSIONS

For the purpose of numerical illustrations we consider the case of free vibrations of a closed circular cylindrical shell. For closed cylindrical shells, the central angle $\alpha = 2\pi$ and the integer n must be even since the shell vibrates in circumferencial full waves. Thus the frequency equation for a closed cylindrical shell can be obtained by setting $\beta = l$ ($l = 1,2,3,...$), where l is the circumferencial wave number in Eqs. (23). We consider zinc material for the purpose of numerical calculations, the physical data[27] for which is given below:

$$\rho = 7.14 \times 10^3 \ \text{kg m}^{-3}, \quad T_0 = 296 \ \text{K},$$

$$c_{11} = 1.628 \times 10^{11} \ \text{Nm}^{-2},$$

$$c_{12} = 0.362 \times 10^{11} \ \text{Nm}^{-2}, \quad c_{13} = 0.508 \times 10^{11} \ \text{N m}^{-2},$$

$$c_{33} = 0.627 \times 10^{11} \ \text{Nm}^{-2}, \quad c_{44} = 0.385 \times 10^{11} \ \text{Nm}^{-2},$$

$$\beta_1 = 5.75 \times 10^6 \ \text{Nm}^{-2}\,\text{deg}^{-1},$$

$$\beta_3 = 5.17 \times 10^6 \ \text{Nm}^{-2}\,\text{deg}^{-1},$$

$$C_e = 3.9 \times 10^2 \ \text{J kg}^{-1}\,\text{deg}^{-1},$$

$$K_1 = 1.24 \times 10^2 \ \text{wm}^{-1}\,\text{deg}^{-1},$$

$$K_3 = 1.24 \times 10^2 \ \text{Wm}^{-1}\,\text{deg}^{-1},$$

$$\epsilon_1 = 2.21 \times 10^{-2}, \quad \omega_1^* = 5.01 \times 10^{11} \ \text{s}^{-1}.$$

The variations of lowest frequencies $\Omega = \omega R(\rho/c_{44})^{1/2}$ of a simply supported cylindrical shell of zinc have been shown in Figs. 2 and 3, respectively, for two values ($l = 1$ and 2) of the circumferential wave number with respect to $t_L = mR/L$ for different values ($t^* = 0.01$, 0.1, 0.25 and 0.5) of the thickness to mean radius of the shell.

From Figs. 2 and 3 it is observed that the lowest frequencies of the shell first increases rapidly to become steady for $t \geqslant 0.8$ in case of $t^* = 0.01$ and 0.1 for $l = 1$ and 2. The increase in lowest frequencies for $t^* = 0.25$ and 0.5 is noticed to be sharp and fast with $t_L = mR/L$ for $l = 1$ and 2. The comparison of both Figs. 2 and 3 reveals that the variations of lowest frequencies in case of $l = 1$ remain close for $0.1 < t_L < 0.5$ and quite dispersive from each other for $t_L \geqslant 0.5$ for all values of t^* whereas the variations in case of $l = 2$ are distinct and dispersive for all values of t_L and t^*. In case of uncoupled thermoelasticity ($\epsilon_1 = 0$) the lowest frequencies are found to be completely in agreement to those obtained by Chen *et al.*[22] and a significant difference in values is noticed for coupled thermoelasticity and thus the frequencies get modified by thermoelastic coupling effects here.

V. CONCLUSIONS

In this paper we directly use modified Bessel functions with complex arguments to study vibration problems of homogeneous, transversely isotropic, cylindrical panels based on three-dimensional thermoelasticity. Three displacement potential functions are introduced so that the equations of motion and heat conduction are uncoupled and simplified. It is noticed that a purely transverse mode is independent of temperature change and the rest of the motion.

A comparison of numerical results with those of related publications proves the feasibility and effectiveness of the present method. The effects of various thermal and mechanical parameters on the natural frequencies of a closed circular cylindrical shell are investigated and the results are presented graphically in Figs. 2 and 3. The results for uncoupled theory of thermoelasticity are also deduced at appropriate stages in this work.

Because neither assumption on the deformation of the panel nor simplified technology is introduced in the analysis, the present method and solutions are strictly exact. The work reported here is more general, with an application to circular cylindrical ''panels'' of arbitrary thickness, from thin shells to extremely thick ones. This can be used in applications involving aerospace, offshore, and/or submarine constructions, pressure vessels, etc. It can also be used to check the applicability of various kinds of two-dimensional simplified shell theory in elastokinetics and numerical methods such as FEM and BEM. The solutions obtained are applicable to both closed hollow cylinders and open ones (panels), depending upon whether $n\pi/\alpha$ is an integer or not.

ACKNOWLEDGMENTS

The author is thankful to the referees for their useful suggestions and REC, Hamirpur and MHRD, GOI, New Delhi for providing financial assistance to complete this work through a FIP under CoE scheme.

[1] E. Y. W. Tsui, ''Natural vibrations of cylindrical panels,'' J. Eng. Mech. **94**, 1435–1444 (1968).
[2] N. R. Maddox, H. E. Plumblee, and W. W. King, ''Frequency analysis of cylindrically curved panel with clamped and elastic boundaries,'' J. Sound Vib. **12**, 225–249 (1970).
[3] J. J. Webster, ''Free vibrations of rectangular curved panels,'' Int. J. Mech. Sci. **10**, 571–582 (1971).
[4] M. Petyt, ''Vibration of curved panels,'' J. Sound Vib. **15**, 381–395 (1971).
[5] T. Mizusawa, ''Application of spline strip method to analyse vibrations of open cylindrical shells,'' Int. J. Numer. Methods Eng. **26**, 663–676 (1988).
[6] M. Ohga, H. Takao, and T. Shigematsu, ''Natural frequencies and modes of open cylindrical shells with a circumferential thickness taper,'' J. Sound Vib. **183**, 143–156 (1995).
[7] T. Mizusawa and T. Kato, ''Applications of the spline prism method to analyze vibration of thick circular cylindrical panels,'' Int. J. Solids Struct. **33**, 967–976 (1996).
[8] C. W. Bert, J. L. Baker, and D. M. Egle, ''Free vibrations of multilayer anisotropic cylindrical shells,'' J. Compos. Mater. **3**, 1173–1186 (1969).
[9] R. M. Jones and H. Margan, ''Buckling and vibration of crossply laminated circular cylindrical shells,'' Am. Inst. Aeron. Astron. **13**, 664–671 (1975).
[10] W. Soedel, ''Simplified equations and solutions for the vibration of orthotropic cylindrical shells,'' J. Sound Vib. **87**, 555–566 (1983).
[11] K. P. Soldatos, ''Influence of thickness shear deformation on free vibrations of rectangular plates and cylinders of antisymmetric angle ply construction,'' J. Sound Vib. **119**, 111–137 (1987).
[12] K. H. Ip, W. K. Chan, P. C. Tse, and T. C. Lai, ''Vibration analysis of orthotropic cylindrical shells with free ends by the Rayleigh–Ritz Method,'' J. Sound Vib. **195**, 117–135 (1996).
[13] K. P. Soldatos and V. P. Hadhgeorgian, ''Three dimensional solution of the free vibration problem of homogeneous isotropic cylindrical shells and panels,'' J. Sound Vib. **137**, 369–384 (1990).
[14] J. Y. So and A. W. Leissa, ''Free vibrations of thick hollow circular cylinders from three dimensional analysis,'' ASME J. Vibr. Acoust. **119**, 89–95 (1997).
[15] X. Y. Jiang, ''3-D vibration analysis of fiber reinforced composite laminated cylindrical shells'' ASME J. Vibr. Acoust. **119**, 46–51 (1997).
[16] J. R. Fan and K. W. Ding, ''Analytical solution for thick closed laminated cylindrical shells,'' Int. J. Mech. Sci. **35**, 657–668 (1993).
[17] J. Q. Ye and K. P. Soldatos, ''3-D vibrations of laminated cylinders and cylindrical panels with symmetric and antisymmetric cross-ply layup,'' Composites Eng. **4**, 429–444 (1994).
[18] I. Mirsky, ''Axi-symmetric vibrations of orthotropic cylinders,'' J. Acoust. Soc. Am. **36**, 2106–2112 (1964).
[19] I. Mirsky, ''Wave propagation in transversely isotropic circular cylinders (Part I and II),'' J. Acoust. Soc. Am. **37**, 1016–1026 (1965).
[20] I. Mirsky, ''Vibrations of orthotropic, thick, cylindrical shells,'' J. Acoust. Soc. Am. **36**, 41–51 (1964).
[21] K. T. Chau, ''Vibrations of transversely isotropic circular cylinders,'' ASME J. Appl. Mech. **61**, 964–970 (1994).
[22] W. Chen, J. Cai, G. Ye, and H. Ding, ''Vibration analysis of orthotropic cylindrical shells with free ends by the Rayleigh–Ritz method,'' J. Sound Vib. **195**, 117–135 (1996).
[23] M. J. P. Musgrave, *Crystal Acoustics* (Holden-Day, San Francisco, CA, 1970).
[24] E. Dieulesaint and D. Royer, *Elastic Waves in Solids* (Wiley, New York, 1980).
[25] P. Chadwick and L. T. C. Seet, ''Wave propagation in transversely isotropic heat conducting elastic material,'' Mathematika **17**, 255–274 (1970).
[26] P. Chadwick, ''Basic properties of plane harmonic waves in a prestressed heat conducting elastic material,'' J. Therm. Stresses **2**, 193–214 (1979).
[27] H. Singh and J. N. Sharma, ''Generalized thermoelastic waves in a transversely isotropic media,'' J. Acoust. Soc. Am. **77**, 1046–1053 (1985).
[28] J. N. Sharma and H. Singh, ''Propagation of generalized thermoelastic waves in cubic crystals,'' Arch. Mech. **42**, 19–30 (1990).
[29] J. N. Sharma and R. S. Sidhu, ''On the propagation of plane harmonic waves in anisotropic generalized thermoelasticity,'' Int. J. Eng. Sci. **24**, 1511–1516 (1986).
[30] V. P. Buchwald, ''Rayleigh waves in transversely isotropic media,'' Aust. Q. J. Mech. Appl. Math. **14**, 193–304 (1961).

Optimum sensor–actuator distance for decentralized acoustic control

Arthur P. Berkhoff[a)]
TNO Institute of Applied Physics, P.O. Box 155, 2600 AD Delft, The Netherlands

(Received 9 May 2000; accepted for publication 30 March 2001)

This paper presents simulation results of a decentralized system for the active minimization of noise transmitted through a plate. The systems are analyzed for harmonic disturbances with respect to stability, convergence, reduction of transmitted sound power, the distance between actuators and sensors, and sensitivity for reverberating environments. Assuming a particular stabilization scheme, it is shown that an optimum exists for the actuator–sensor distance in a broadband sense. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1378352]

PACS numbers: 43.40.Vn, 43.40.At, 43.40.Rj [PJR]

I. INTRODUCTION

In most implementations for active noise and vibration control it is assumed that the cost function to be minimized by the controller is a time-averaged, summed square value of all error signals, such as microphone signals. This cost function is minimized by a simultaneous optimization of the strengths of all actuator signals.[1] Although such a global, or centralized, configuration has desirable properties such as guaranteed convergence and stability, there are also some serious drawbacks. For large numbers of actuators and sensors the computational load for the processor can be very large, especially for the case of random primary noise. In a decentralized, or local, control configuration, multiple independent controllers are used. The cost function for a particular controller is obtained from a subset of the sensor signals, which is minimized by driving a subset of the secondary sources. Mostly, the sensors and actuators are local to the controller. In this paper, therefore, the terms decentralized and local have a similar meaning. The attractiveness of a local or decentralized configuration is mainly due to the need for reduced processor requirements, improved convergence, and a reduced number of connections between the sensors, actuators, and controller.

An analysis of decentralized control configurations for active noise control has been presented by Elliott and Boucher.[2] In the latter article, a noise reduction system is described having a panel of actuators arranged in a first plane and a plurality of error sensors in a second plane. The first and second planes are parallel to one another. Elliott and Boucher present an analysis of a decentralized adaptive feedforward control system. They also present results of some physical examples of two-actuator/two-sensor systems and three-actuator/three-sensor systems. In these examples, the mutual distances between the error sensors and the actuators are introduced as important parameters to derive conditions as to when such a system is stable. In some cases the decentralized system is inherently unstable; stability can then be enforced by introduction of an effort weighting term. This term has a detrimental effect on system performance, however. The paper by Elliott and Boucher also gives a stability analysis using the Gershgorin theorem. A stability analysis based on a generalization of the Gershgorin theorem is presented in Ref. 3.

Other works on decentralized acoustic control are known,[4] but the literature on the subject seems relatively sparse, especially the combined analysis of the physics and parameters of the control algorithms. In this paper, optimum sensor positioning for decentralized configurations is addressed. For single secondary sources optimum sensor positioning has been studied by Qiu, Hansen, and Li.[5] For multi-input multi-output systems, using regular centralized control, this has been studied by Guo, Pan, and Bao.[6] Using wave-number analysis, Johnson[7] showed that a simple gain would be effective in controlling the radiation of all the supersonic wave-number components. In the present paper, it is shown that the stabilization of the control system in combination with the physics of the problem leads to an optimum in the sensor–actuator distance.

II. ANALYSIS

A derivation is given for the analysis of multiple local control systems, where interaction is present between the control systems through the physics of the problem. The derivation is based on combining the different implicit formulations for the control signals corresponding to the minimization of each of the local cost functions.

Computational nodes, indicated by index $i=1,...,I$, are defined that operate on a subset of the sensor signals and drive a subset of the actuators. It is assumed that each of the computational nodes, which act as local controllers, tries to minimize a cost function based on sensor signals local to that node by driving actuators which are also local to that node. The scalar cost functions J_i for the I nodes i are written as

$$J_i = \mathbf{p}_i^H \mathbf{p}_i + \mathbf{u}_i^H \beta_i \mathbf{u}_i, \quad i=1,...,I, \tag{1}$$

with

$$\mathbf{p}_i = \mathbf{W}_i \mathbf{p}, \tag{2}$$

in which $\mathbf{p}$ is an $M\times 1$ vector of sensor signals, $\mathbf{W}_i$ is a weighting matrix of dimensions $P\times M$, which provides a selection and weighting of P out of a total of M sensor sig-

a)Electronic mail: berkhoff@tpd.tno.nl

 0001-4966/2001/110(1)/260/7/$18.00

nals used as error inputs for node i; $\mathbf{u}_i$ is a $K\times$one-dimensional control signal for node i and β_i is a $K\times K$ dimensional effort weighting matrix. The sensor signals $\mathbf{p}$ result from the superposition of primary field contributions $\mathbf{p}_p$ and the secondary field contributions $\mathbf{p}_s$ due to N actuators. The latter contributions are given by $\mathbf{Gu}$, where $\mathbf{u}$ is an $N\times 1$ vector denoting the control signals that drive the actuators and $\mathbf{G}$ is an $M\times N$ matrix of transfer functions between control signals and sensor signals. Hence,

$$\mathbf{p}=\mathbf{p}_p+\mathbf{Gu}. \tag{3}$$

Each node drives K actuators, so $N=IK$. The contribution of the self-term $\mathbf{G}_i\mathbf{u}_i$ of $\mathbf{Gu}$ for node i will be separated from the other terms as follows:

$$\mathbf{Gu}=\mathbf{G}_i\mathbf{u}_i+\mathbf{G}_i'\mathbf{u}_i', \quad i=1,...,I, \tag{4}$$

where the $M\times K$-dimensional matrix $\mathbf{G}_i$ denotes the columns of $\mathbf{G}$ corresponding to controller i. The complement $\mathbf{u}_i'$ of $\mathbf{u}_i$ has dimensions $(N-K)\times 1$; the complement $\mathbf{G}_i'$ of $\mathbf{G}_i$ has dimensions $M\times(N-K)$. As a result, the local cost function becomes

$$\begin{aligned}J_i=&\mathbf{p}_p^H\mathbf{F}_i\mathbf{p}_p+\mathbf{p}_p^H\mathbf{F}_i\mathbf{G}_i\mathbf{u}_i+\mathbf{p}_p^H\mathbf{F}_i\mathbf{G}_i'\mathbf{u}_i'+\mathbf{u}_i^H\beta_i\mathbf{u}_i\\&+\mathbf{u}_i^H\mathbf{G}_i^H\mathbf{F}_i\mathbf{p}_p+\mathbf{u}_i^H\mathbf{G}_i^H\mathbf{F}_i\mathbf{G}_i\mathbf{u}_i+\mathbf{u}_i^H\mathbf{G}_i^H\mathbf{F}_i\mathbf{G}_i'\mathbf{u}_i'\\&+\mathbf{u}_i'^H\mathbf{G}_i'^H\mathbf{F}_i\mathbf{p}_p+\mathbf{u}_i'^H\mathbf{G}_i'^H\mathbf{F}_i\mathbf{G}_i\mathbf{u}_i+\mathbf{u}_i'^H\mathbf{G}_i'^H\mathbf{F}_i\mathbf{G}_i'\mathbf{u}_i',\end{aligned} \tag{5}$$

with $\mathbf{F}_i=\mathbf{W}_i^H\mathbf{W}_i$. This cost function is recognized as the standard Hermitian form

$$J_i=\mathbf{u}_i^H\mathbf{A}_i\mathbf{u}_i+\mathbf{u}_i^H\mathbf{b}_i+\mathbf{b}_i^H\mathbf{u}_i+c_i, \tag{6}$$

where

$$\mathbf{A}_i=\mathbf{G}_i^H\mathbf{F}_i\mathbf{G}_i+\beta_i, \tag{7}$$

$$\mathbf{b}_i=\mathbf{G}_i^H\mathbf{F}_i\mathbf{p}_p+\mathbf{G}_i^H\mathbf{F}_i\mathbf{G}_i'\mathbf{u}_i', \tag{8}$$

$$\begin{aligned}c_i=&\mathbf{p}_p^H\mathbf{F}_i\mathbf{p}_p+\mathbf{p}_p^H\mathbf{F}_i\mathbf{G}_i'\mathbf{u}_i'+\mathbf{u}_i'^H\mathbf{G}_i'^H\mathbf{F}_i\mathbf{p}_p\\&+\mathbf{u}_i'^H\mathbf{G}_i'^H\mathbf{F}_i\mathbf{G}_i'\mathbf{u}_i'.\end{aligned} \tag{9}$$

The conjugate derivatives of J_i with respect to the corresponding control signals $\mathbf{u}_i$ are[8]

$$\frac{\partial J_i}{\partial \mathbf{u}_i^*}=2(\mathbf{A}_i\mathbf{u}_i+\mathbf{b}_i), \quad i=1,...,I. \tag{10}$$

In view of Eq. (10), the minimization of J_i with respect to $\mathbf{u}_i$ by setting $\partial J_i/\partial\mathbf{u}_i^*=0$ leads to a set of implicit relations for the different $\mathbf{u}_i$. This can be formulated as

$$\min_{u_i} J_i\rightarrow\mathbf{A}_i\mathbf{u}_i=-b_i, \quad i=1,...,I. \tag{11}$$

Substitution of Eqs. (7) and (8) yields

$$(\mathbf{G}_i^H\mathbf{F}_i\mathbf{G}_i+\beta_i)\mathbf{u}_i=-\mathbf{G}_i^H\mathbf{F}_i\mathbf{p}_p-\mathbf{G}_i^H\mathbf{F}_i\mathbf{G}_i'\mathbf{u}_i', \quad i=1,...,I. \tag{12}$$

A simplification results from the use of Eq. (4):

$$\mathbf{G}_i^H\mathbf{F}_i\mathbf{Gu}+\beta_i\mathbf{u}_i=-\mathbf{G}_i^H\mathbf{F}_i\mathbf{p}_p, \quad i=1,...,I, \tag{13}$$

which are I sets of K linear equations in N unknown components of $\mathbf{u}$. Introducing the $M\times N$ matrix

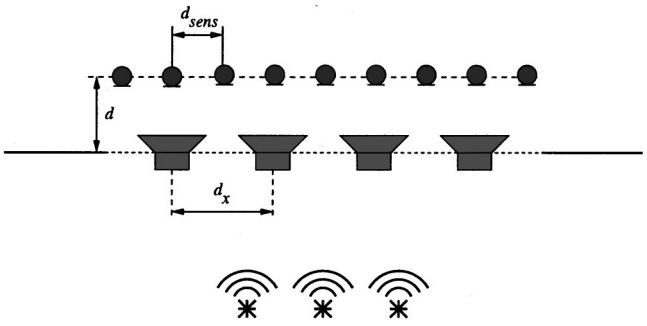


FIG. 1. Active noise control configuration.

$$\hat{\mathbf{G}}=[\mathbf{F}_1\mathbf{G}_1,\mathbf{F}_2\mathbf{G}_2,...,\mathbf{F}_I\mathbf{G}_I], \tag{14}$$

and the $N\times N$ block-diagonal matrix

$$\beta=\begin{bmatrix}\beta_1 & 0 & \cdots & 0\\ 0 & \beta_2 & \cdots & 0\\ \cdots & & & \cdots\\ 0 & 0 & \cdots & \beta_I\end{bmatrix}, \tag{15}$$

the linear system of Eq. (13) can be written more compactly as

$$(\hat{\mathbf{G}}^H\mathbf{G}+\beta)\mathbf{u}=-\hat{\mathbf{G}}^H\mathbf{p}_p. \tag{16}$$

The present result explicitly includes the weighting factors for the error sensors. To arrive at the solution for $\mathbf{u}$ an iterative procedure is implemented in the system, such as the procedure described by Elliott, Boucher, and Nelson.[9] For interpretation of the behavior of the decentralized system the reader is referred to Ref. 2.

III. SIMULATIONS

Simulation results are given for an active control system for the reduction of sound transmitted through a plate, such as the configuration in Fig. 1. A possible configuration for

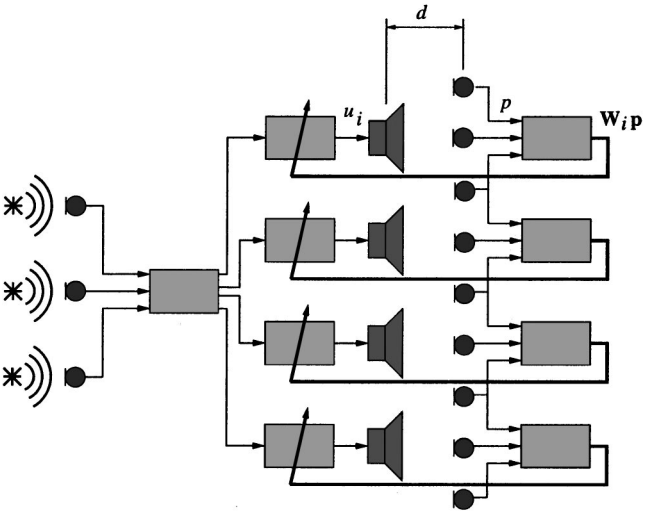


FIG. 2. Decentralized control configuration using four local controllers, one loudspeaker, and three microphones per controller ($M=9$, $N=4$, $P=3$, and $K=1$), and three sensors for the feedforward link.

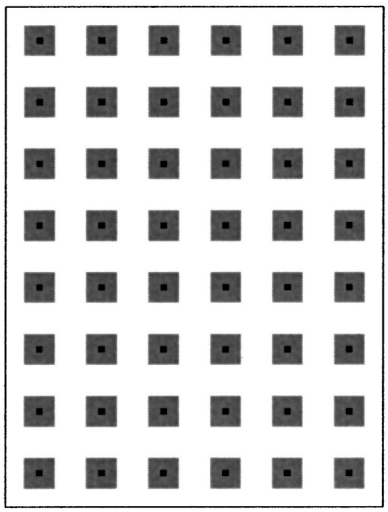


FIG. 3. Configuration using a grid of $6\times8=48$ sensors (small squares) and $6\times8=48$ sources (large squares).

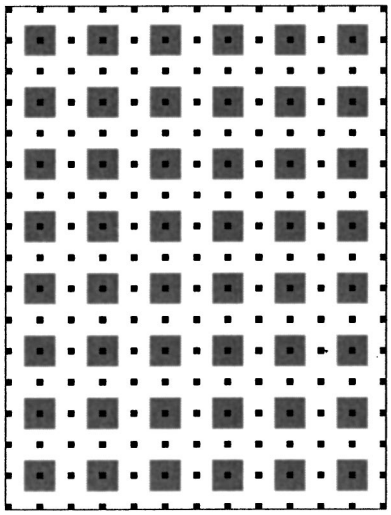


FIG. 4. Configuration using a grid of $13\times17=221$ sensors (small squares) and $6\times8=48$ sources (large squares).

the controller is given in Fig. 2, which also includes a feedforward link for the reference signals. The reference signals can be obtained from a signal that is independent of the actuator signals, such as a tacho signal. The resulting operation of the control system is feedforward. Alternatively, the reference signals could be chosen to be identical to the error signals, which results in a feedback system. If the influence of the secondary path on the reference sensor is subtracted in the controller then the system can be regarded as feedforward. The assumption in the present paper is that for each controller a reference signal is available that has good correlation with the primary signal on the error sensor. The sensors are pressure sensors that are placed at various distances from the plate. The actuators are loudspeakers which operate as constant velocity sources, which are positioned directly on the plate. The plate is assumed to be a 1-mm-thick aluminum plate of $60\,\mathrm{cm}\times80\,\mathrm{cm}$, having a modulus $7\times10^{10}\,\mathrm{Pa}$, Poisson ratio of 0.3, hysteretic damping $\eta=0.02$, and a density of $2.6\times10^{3}\,\mathrm{kg\,m^{-3}}$. The plate is simply supported and the incident field is a plane wave arriving at a direction of 60° to the plate normal. Two configurations are discussed. The first configuration consists of $6\times8=48$ actuators and $6\times8=48$ sensors, as shown in Fig. 3. The second configuration consists of $6\times8=48$ actuators and $13\times17=221$ sensors, as shown in Fig. 4. For both cases the mutual actuator distance equals $d_x=0.1\,\mathrm{m}$ in both in-plane directions.

As opposed to active control systems which minimize a global quadratic error criterion, stability is not guaranteed in multiple local systems. Assuming an iterative procedure to solve Eq. (16), the system is stable if the real parts of the eigenvalues λ_i, $i=1,...,N$ of the matrix $\hat{\mathbf{G}}^H\mathbf{G}+\beta$ are positive.[2] The effort weighting matrix is taken to be the diagonal matrix $\beta=\beta\mathbf{I}$. If the system is unstable for $\beta=0$, the value of β will be set equal to $-\min_i \mathrm{Re}\,\lambda_i$, which makes the system just stable. Increasing the value of β further would enhance the stability margin and improve the speed of convergence of the iterative procedure, but also increase the residual radiated power. The convergence of some iterative procedures is governed by the ratio of the largest singular value κ_1 to the smallest singular value κ_N,[9] i.e., the condition number of the matrix $\hat{\mathbf{G}}^H\mathbf{G}+\beta$.[10]

A. Methods

The models describing the vibration of the plate can be found in Ref. 11. The pressures $\mathbf{p}_p$ and $\mathbf{p}_s$ were computed with a weak form of a Fourier-type extrapolation technique in which singularities were evaluated by analytical integration.[12] In principle, an integral equation method or a boundary element method as described in Ref. 13 can also be used, but the latter methods are less efficient for geometries of this size and larger. Formulas for zero-extrapolation distance can be found in Ref. 14. The sound power was obtained from the sum of primary velocity and velocity of the secondary sources for each point on the plate. The resulting total velocity was transformed to the wave-number domain. Subsequently, the wave-number domain pressure was computed[14] and the spatial pressure distribution was obtained after the inverse transformation. The radiated power was obtained from the real part of the product of conjugated pressure and velocity integrated over the plate area in the spatial domain.

B. Results

The sound powers without control and with control for various configurations are shown in Fig. 5. All sound powers are given in arbitrary units, expressed in dB. The dimensionality of the decentralized systems is indicated by $M\times N$, $P\times K$. The cost functions for the 221×48, 9×1 system were obtained by specifying ones for the nine corresponding entries in the weighting matrix $\mathbf{W}_i$ and zeros at the other $221-9=212$ positions. For each i the corresponding weighting matrix $\mathbf{W}_i$ was constructed in such a way. The cost functions for the 48×48, 1×1 system were obtained by specifying a 1 for the single corresponding entry in the weighting matrix $\mathbf{W}_i$ and zeros at the other 47 positions.

It can be seen that, for $d=0$, the global 48×48, 1×1 system lead to identical results. This agrees with the results

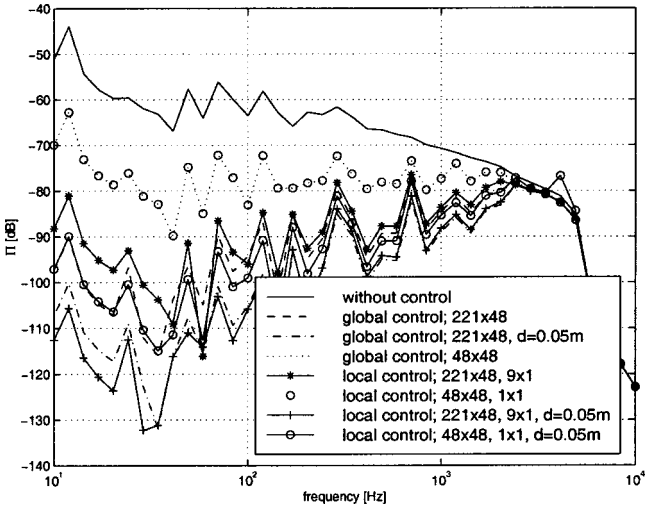


FIG. 5. Sound power II radiated from plate without control, with global control and with local control.

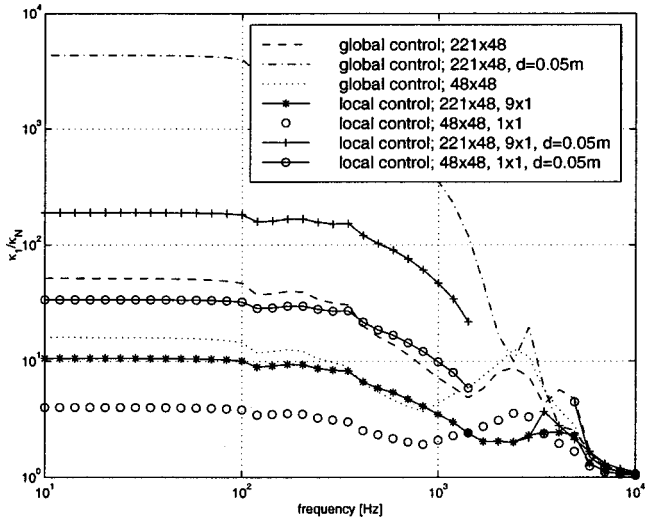


FIG. 6. Condition numbers κ_1/κ_N corresponding to Fig. 5.

of Elliott and Boucher,[2] who showed that the steady-state actuator driving signals for these two square systems are identical if the decentralized system is stable. In Fig. 5, it can also be seen that the results for the nonsquare global 221 $\times 48$ system and the local 221×48, 9×1 system are not identical.

It was found that reductions could be obtained for frequencies for which the mutual distance d_x between the actuators were smaller than half of a wavelength. Moreover, the distance d between the sensors and the plate turns out to be an important parameter. Larger reductions are obtained if the pressure sensors are moved away from the plate. This distance d cannot be made arbitrarily large because of stability issues. The point of instability is reached at, depending on the configuration, 0.25–0.4 of a wavelength from the plate if the ratio d/d_x is larger than a certain minimum value. If this ratio is smaller than this value, then the system is stable for all frequencies. Elliott and Boucher[2] show that the decentralized square systems are stable for all distances d if monopole sources are used. In the present paper, extended (distributed) sources are used which lead to potentially unstable systems for large d. This was also seen if, instead of using a single extended source for each local control system, multiple monopole sources were used which were driven in parallel, each of which were identical to that of Elliott and Boucher.[2]

The corresponding condition numbers are shown in Fig. 6. If a positive value of β was used to make the system stable, then the condition number is not shown.

1. Influence of d on the reduction

From the previous results it was found that the distance d between the actuator plane and the sensor plane has a considerable influence on the achievable reduction of radiated sound power. On the one hand, a larger distance d leads to improved estimation of the radiated sound power and corresponding larger reductions. On the other hand, the distance d determines the frequency above which the system has to be stabilized by increasing the value of β, which leads to smaller reductions.

Clearly, two contradicting requirements for d have to be satisfied for broadband reductions. This is illustrated in Fig. 7, which shows the radiated sound power for various distances d. It can be seen that, for small d, reductions are increased by increasing d, particularly at low frequencies. However, the system has to be stabilized above the frequency where d becomes 0.25–0.4 of a wavelength. This stabilization leads to smaller reductions at high frequencies.

Hence, for broadband applications there might be an optimum value for d. In order to determine this distance, it is assumed that all frequencies are taken into account for which half of the wavelength is larger than the actuator spacing d_x. For the present configuration, this corresponds to all frequencies smaller than $f=c/(2d_x)=1715\,\mathrm{Hz}$, with c the speed of sound. The latter frequency is indicated by a dashed line in Fig. 7. This frequency corresponds to the upper frequency for which an active control system using a global error cri-

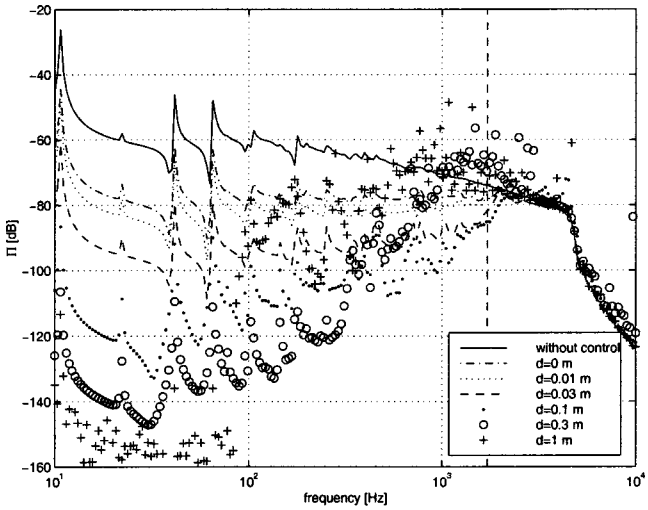


FIG. 7. Sound power II radiated from plate without control and with local control using a 48×48, 1×1 system, i.e., using a total of 48 sensors and 48 actuators, 1 sensor and 1 actuator for each independent controller with $d_x=0.1$ m and the distance d between the actuator plane and the sensor plane as parameter. The frequency $f=c/2d_x$, for which half of the wavelength equals the actuator distance $d_x=0.1$ m, is indicated by a dashed line.

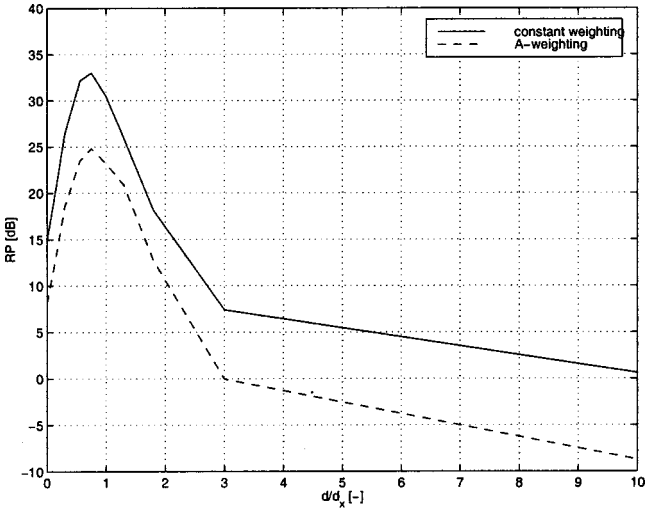


FIG. 8. Broadband reduction of sound power corresponding to Fig. 7, taking into account all frequencies $f<c/2d_x$, with c the speed of sound and d_x the actuator spacing.

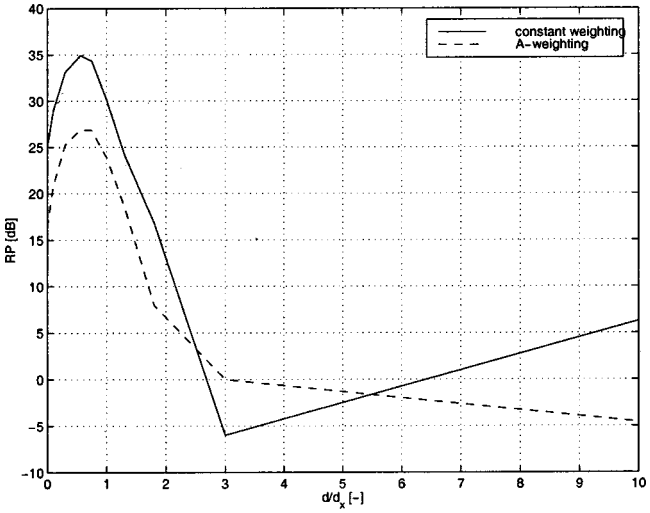


FIG. 10. Broadband reduction of sound power corresponding to Fig. 9, taking into account all frequencies $f<c/2d_x$, with c the speed of sound and d_x the actuator spacing.

terion, or using a local error criterion with small d, leads to significant reductions of radiated sound power (see Fig. 5). For the present 48×48, 1×1 system, the sensor spacing is identical to the actuator spacing. The broadband reductions RP for various values of d normalized to actuator spacing d_x are shown in Fig. 8. Indeed, it can be seen that there is a maximum in the reduction of broadband radiated sound power, both for constant weighting and for A weighting. The maximum reduction is obtained for $d_x/2\leqslant d\leqslant d_x$.

It is noted that the present sensor configuration is not designed to directly estimate the radiated power. Nevertheless, the location of the sensor does have a substantial effect on the amount of the reduction of radiated power and the sensor configurations as discussed in the present paper, and Ref. 2 can be quite useful for the reduction of radiated sound power.

The results for a 221×48, 9×1 system, having half the distance between the sensors, are shown in Figs. 9 and 10. Figure 9 shows the sound power radiated from a plate in such a system, whereas Fig. 10 shows the broadband reduction, again for all frequencies $f\leqslant c/(2d_x)$. It can be seen that the maximum reduction which can be obtained is similar. The optimum value for d, as obtained from Fig. 10, is also within the range $d_x/2\leqslant d\leqslant d_x$, although the peak in the reduction is wider than in Fig. 8. In practice, therefore, the value of d for the 9×1 system will often be chosen somewhat smaller than for the 1×1 system. The results for a global control system are shown in Figs. 11 and 12. The differences with the preceding local control systems are mainly in the high-frequency range. This leads to larger values for d as well as less-pronounced maxima. In Fig. 13 a comparison is given of decentralized (local) control and global control. It can be seen that at high frequencies the reductions with decentralized control are less than with global control because of the required stabilization.

From the results it can also be seen that at high frequencies the performance of the global control system is better

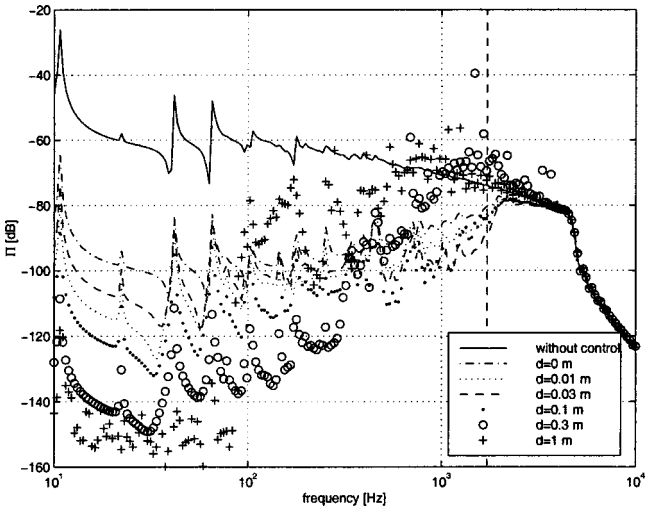


FIG. 9. As Fig. 7, for a 221×48, 9×1 system, i.e., using a total of 221 sensors and 48 actuators, 9 sensors and 1 actuator for each independent controller.

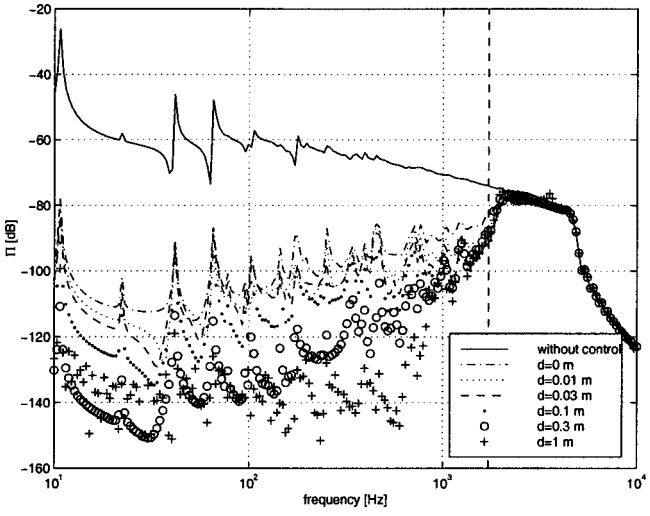


FIG. 11. As Fig. 7, for global control using a total of 221 sensors and 48 actuators.

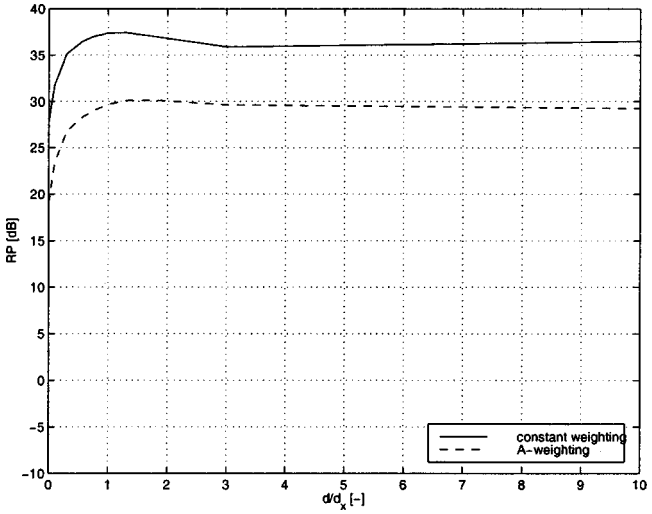


FIG. 12. Broadband reduction of sound power corresponding to Fig. 11, taking into account all frequencies $f<c/2d_x$, with c the speed of sound and d_x the actuator spacing.

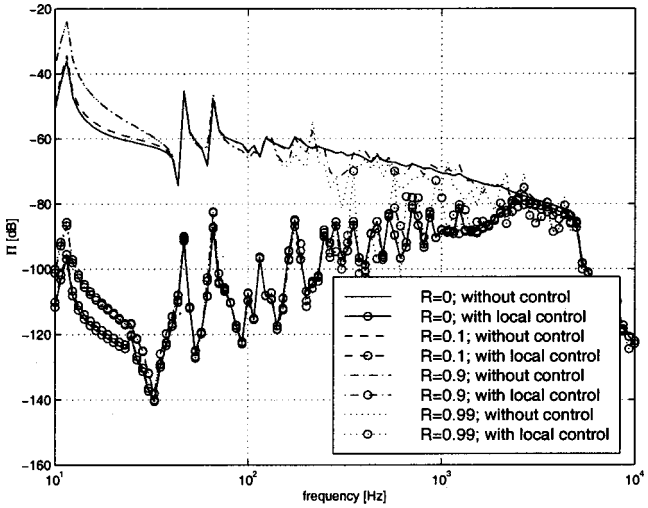


FIG. 14. Sound power II radiated from a plate without control and with local control for a reflecting wall at 1 m distance having a reflection coefficient R.

than that of the local control systems. This seems a contradiction with the results of Elliott and Boucher,[2] who derive that the actuator driving signals are identical for local and global control for stable square systems. An important difference is that for the present configuration extended sources are used, which lead to potentially unstable systems at high frequencies. Hence, because the stability constraint can be violated for the present configuration at high frequencies, the derivation leading to the equalness of the control signals for local control and global control in Ref. 2 is inappropriate for the present configuration at these frequencies. Furthermore, at such frequencies, the local systems actually have to be stabilized and/or the control action limited, which leads to performance differences between local control and global control.

2. Performance in reverberating environment

The performance of the local control system was also investigated for the case of a reflecting parallel plane. The distance of this plane to the actuators was taken to be 1 m. The reduction which can be obtained with this configuration is shown in Fig. 14. It can be seen that for reflection coefficients smaller than or equal to 0.9 the control system remains stable and leads to reasonable reductions. For a reflection coefficient of 0.99 the possible reduction above approximately 500 Hz becomes less than for lower reflection coefficients.

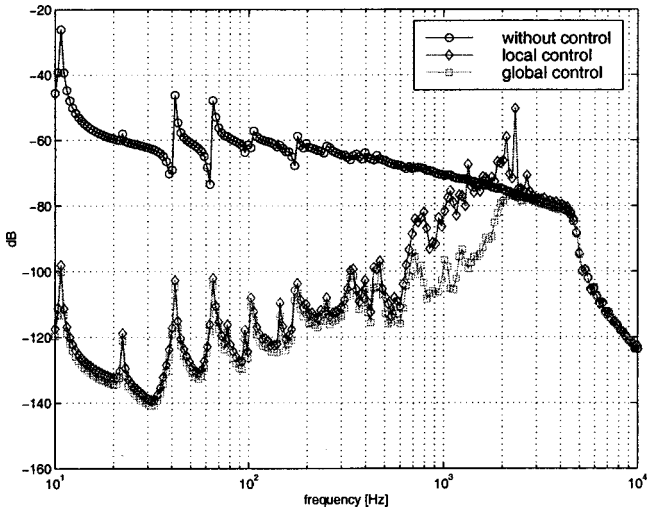


FIG. 13. Sound power transmitted through the plate without control, with global control and local control, both for $d=0.3$ m.

C. Discussion

The spectral characteristics of the primary noise have some influence on the optimum for d. In the following, an explanation is given for why this influence is rather small. For the active control problem at hand, the residual at high frequencies is larger than the residual at low frequencies, simply because it is easier to reduce low-frequency components. Therefore, the high-frequency range will dominate the optimum for d. If the primary noise does not contain significant spectral components at high frequencies, then the optimum for d will be at higher values of d/d_x. However, in the latter case, the mutual distance between the actuators can be considered to be smaller than strictly necessary and the total number of actuators unnecessarily large. Hence, in the case of essentially broadband primary disturbances up to frequency $f_{\max}$ and an economic actuator configuration designed for $f \leqslant f_{\max}$, there will be a clear optimum value for d, being relatively independent of the spectral characteristics of the primary noise, if the primary spectrum is reasonably flat in the neighborhood of $f_{\max}$.

Additional factors might influence the optimum value for d. For smaller d, the condition number κ_1/κ_N of the system is lower, and often, therefore, the convergence of adaptive schemes better. Furthermore, the present stabilization procedure leads to the maximum obtainable reductions at high frequencies. For improved convergence, larger values of β may be required, leading to smaller reductions at high frequencies. Experiments were performed to investigate the amount of performance degradation by increasing β. It was

found that using a value of β which was twice the value required to make the system stable did not lead to significant changes of the optimum value for d. The differences of the reductions which were obtained by the doubling of β were less than 3 dB. A performance advantage of a larger value of β is obtained at frequencies higher than $f=c/(2d_x)$, where the possible increases of the sound power become smaller.

These considerations can lead to an optimum for d, which is somewhat smaller than given by Figs. 8 and 10. If the speakers are moved away from the plate while keeping the sensors on the same position, the stability margin increases. However, the reduction of radiated sound power will not be as good because the system now tends to minimize the sound field in the immediate vicinity of the sensors, while giving worse results at other positions with corresponding smaller reductions.

The active techniques as described in the present paper could be used in conjunction with passive techniques, which often are simpler and more cost effective at high frequencies. If passive means are used at high frequencies, then the number of microphones and loudspeakers could be reduced because of the reduction of the maximum operating frequency of the active system.

The acoustic transmission through a plate could also be obtained by using actuators that directly act on the plate. In that case, there is a strong and more-complicated interaction between the actuators and the sensors. If reductions for all frequencies below 1715 Hz are to be obtained for a plate of 60 cm$\times$80 cm using structural actuators and structural sensors, a control system of at least 45 input channels and 45 output channels[15] is needed, which, if fully coupled, is a very large system by current standards, especially for systems suitable for broadband signals. A decentralized control system using collocated and matched structural actuators and sensors can be used to increase the damping of the plate.[16] Similar results can be obtained with a decentralized configuration in which collocated piezoelectric patch actuators and velocity sensors are used.[17] The present decentralized configuration allows the use of compact, independently operating systems that do not interact with the structure and that eventually could be produced at low cost.

IV. CONCLUSIONS

A decentralized control configuration for active reduction of sound transmission has been studied. Assuming a particular stabilization method, it has been shown that for broadband harmonic disturbances there is an optimum distance d between the loudspeakers and the microphones that was expressed as a constant times the mutual distance between the loudspeakers.

ACKNOWLEDGMENT

The author would like to thank Dr. J. J. Hutter for his suggestions.

[1] P. A. Nelson and S. J. Elliott, *Active Control of Sound* (Academic, New York, 1992).

[2] S. J. Elliott and C. C. Boucher, ''Interaction between multiple feedforward active control systems,'' IEEE Trans. Speech Audio Process. **2**, 521–530 (1994).

[3] Y. Elgrichi and E. Zeheb, ''Stability of multichannel sound control systems,'' IEE Proc. Vision Image Signal Process. **144**, 1–7 (1997).

[4] I. Stothers, ''Adaptive control system having multiple inputs and multiple outputs,'' Patent No. GB 2310512 A, Application No. 9703686.7, assigned to Lotus Cars Ltd.

[5] X. Qiu, C. H. Hansen, and X. Li, ''A comparison of near-field acoustic error sensing strategies for the active control of harmonic-free field sound radiation,'' J. Sound Vib. **215**, 81–103 (1998).

[6] J. Guo, J. Pan, and C. Bao, ''Actively created quiet zones by multiple control sources in free space,'' J. Acoust. Soc. Am. **101**, 1492–1501 (1997).

[7] M. E. Johnson, ''Analysis of active control systems for controlling sound radiation from infinite structures,'' in Proc. ACTIVE 1999, The Institute of Noise Control Engineering of the USA (1999), edited by S. Douglas, pp. 435–446 (unpublished).

[8] S. S. Haykin, *Adaptive Filter Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1986).

[9] S. J. Elliott, C. C. Boucher, and P. A. Nelson, ''The behavior of a multiple channel active control system,'' IEEE Trans. Signal Process. **40**, 1041–1052 (1992).

[10] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. (Johns Hopkins, Baltimore, MD, 1989).

[11] C. R. Fuller, S. J. Elliott, and P. A. Nelson, *Active Control of Vibration* (Academic, London, 1996).

[12] A. P. Berkhoff, J. M. Thijssen, and R. J. F. Homan, ''Simulation of ultrasonic imaging with linear arrays in causal absorptive media,'' Ultrasound Med. Biol. **22**, 245–259 (1996).

[13] R. D. Ciskowski and C. A. Brebbia, *Boundary Element Methods in Acoustics* (Elsevier, London, 1991).

[14] E. G. Williams and J. D. Maynard, ''Numerical evaluation of the Rayleigh integral for planar radiators using the FFT,'' J. Acoust. Soc. Am. **72**, 2020–2030 (1982).

[15] A. P. Berkhoff, ''Sensor scheme design for active structural acoustic control,'' J. Acoust. Soc. Am. **108**, 1037–1045 (2000).

[16] A. Preumont, *Vibration Control of Active Structures* (Kluwer, Dordrecht, 1997).

[17] S. J. Elliott, P. Gardonio, T. C. Sors, and M. J. Brennan, ''Active vibroacoustic control with multiple local feedback loops,'' Proc. SPIE, Smart Materials and Integrated Systems, ed. by L. P. Davis, Vol. **4327-83** (2001).

Active attenuation of the wave transmission through an L-plate junction

Nicole J. Kessissoglou[a)]
Mechanical Engineering, James Cook University, Townsville, Queensland 4811, Australia

(Received 1 March 2000; accepted for publication 30 March 2001)

Active control is applied to an L-shaped plate in order to attenuate the flexural energy transmission from one plate to the other. The coupled plates are simply supported along two parallel sides, and free at the other two ends. Point forces are used to generate the primary and secondary excitation of the plates. The flexural wave coefficients are determined from the boundary conditions, continuity equations at the driving force locations, and continuity equations at the corner junction of the two plates. Bending, shearing, and longitudinal effects are taken into consideration at the corner junction. Under broadband frequency control at a discrete location in plate 2, both the control shaker and the error sensor are optimally located to achieve the best control performance. Results show that when the control force and error sensor are arbitrarily located, the control performance is dependent on the excitation frequency. When both the control force and error sensor are optimally located with respect to the primary shaker location in a symmetrical arrangement, the control performance is both maximized and independent of the excitation frequency. Using single-frequency control to attenuate the total vibrational response of the coupled plates, the error sensor location is strongly mode dependent. It is shown that using a single, properly located control force and a single, properly located error sensor, global attenuation of the L-shaped plate can be achieved. [DOI: 10.1121/1.1378354]

PACS numbers: 43.40.Vn [PJR]

I. INTRODUCTION

Wave propagation through coupled plate structures in an L, T, or cross configuration has been given considerable attention in the past few decades. In one of the earliest works, Cremer, Heckl, and Unger[1] examined the vibrations of two infinite plates coupled together at a right angle in terms of elastic waves that propagate through the structure. Incidence waves propagating both normally and obliquely generate both reflected and transmitted waves at the structural discontinuity. The wave coefficients are determined from continuity equations at the plate junction. Guyader and co-workers[2,3] derived the global modes of thin rectangular plates coupled in an L shape and a cross junction. Using global modes, the time-averaged kinetic energy was determined for each structure for the case of random broadband excitation. Cuschieri[4] investigated the power flow through the junction of a finite L-shaped plate using a mobility power flow approach. Shen and Gibbs[5] used an approximate method to calculate the bending vibrations of coupled plates, by means of expressing the bending displacement amplitude function as a linear combination of coordinate functions. The frequency response curves for a series of T and L combinations of rectangular plates are given for variations in material damping and receiver position. Although much attention has been given to the dynamic response of L-shaped plates, to the author's knowledge there has been no work conducted to actively attenuate the wave transmission in such structures. Previous works on active vibration control include active control of homogeneous plates[6] and ribbed plates.[7,8]

In this paper, active control is applied to a point force excited L-shaped plate, in order to attenuate the flexural energy transmission from one plate to the other. A single control force is used in close proximity to the primary shaker, and the error signal is located in the connecting plate. Under broadband frequency control, results show that when both the control shaker and error sensor are optimally located in a symmetrical arrangement around the primary shaker, the control performance is maximized and independent of the excitation frequency. Under single-frequency control, the control performance is strongly dependent on the error sensor position. Location of the error sensor at an antinodal point for a mode will result in global attenuation of the L-shaped plate.

II. ANALYTICAL DEVELOPMENT OF THE L-SHAPED PLATE MODEL

The L-shaped plate structure consists of two finite elastic plates coupled together at a right angle as shown in Fig. 1. Both plates are simply supported at $y=0$ and $y=L_y$, and free at the other two ends corresponding to $x_1=L_{x1}$ and $x_2=L_{x2}$. The junction of the two plates corresponds to $x_1=0$ and $x_2=0$. Point force excitation at a location of (x_0,y_0) in plate 1 is used to drive the coupled structure, and is described by the following Dirac delta function:

$$F(x_0,y_0)=F_0\delta(x-x_0)\,\delta(y-y_0). \tag{1}$$

Due to the boundary conditions, the plate flexural displacement is described by both a modal solution in the y direction, and a traveling-wave solution along the x_1 and x_2 directions. General solutions for the primary response can be obtained

a)Electronic mail: nicole.kessissoglou@jcu.edu.au

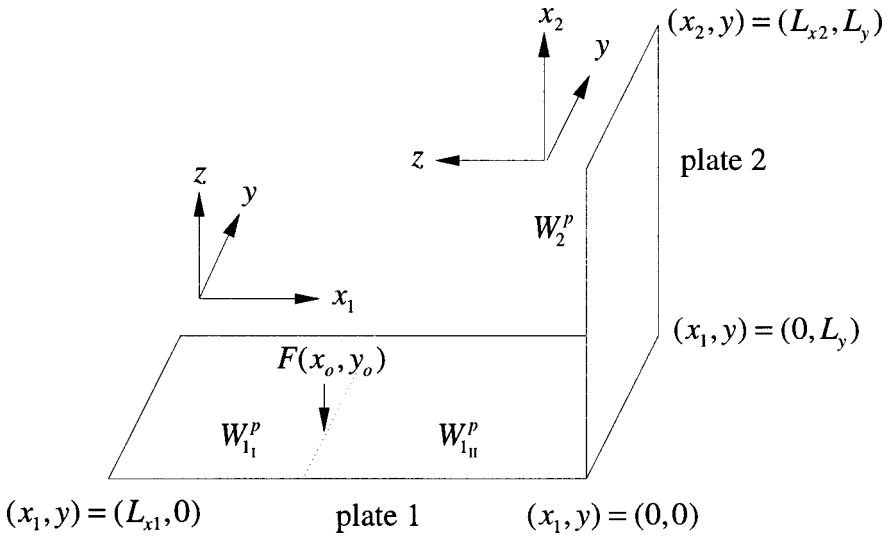


FIG. 1. Finite L-shaped plate showing the three regions of the coupled structure under point force excitation.

for three regions of the plate corresponding to[6]

$$W_{1_I}^p(x_1,y)=\sum_{m=1}^{\infty}(A_1^p e^{-jk_{x1}x_1}+B_1^p e^{jk_{x1}x_1}+C_1^p e^{-k_{n1}x_1}+D_1^p e^{k_{n1}x_1})\sin(k_y y),\quad L_{x1}\leqslant x_1\leqslant x_0, \tag{2}$$

$$W_{1_{II}}^p(x_1,y)=\sum_{m=1}^{\infty}(A_2^p e^{-jk_{x1}x_1}+B_2^p e^{jk_{x1}x_1}+C_2^p e^{-k_{n1}x_1}+D_2^p e^{k_{n1}x_1})\sin(k_y y),\quad x_0\leqslant x_1\leqslant 0, \tag{3}$$

$$W_2^p(x_2,y)=\sum_{m=1}^{\infty}(A_3^p e^{-jk_{x2}x_2}+B_3^p e^{jk_{x2}x_2}+C_3^p e^{-k_{n2}x_2}+D_3^p e^{k_{n2}x_2})\sin(k_y y),\quad 0\leqslant x_2\leqslant L_{x2}, \tag{4}$$

where A_i^p and B_i^p are the coefficients of the propagating waves, and C_i^p, D_i^p are the coefficients of the near-field decay waves. $k_y=m\pi/L_y$ is the modal wave number along the y direction, where m is the mode number. $k_p=(\omega^2\rho h/D)^{1/4}$ is the *in vacuo* plate flexural wave number, where $D=Eh^3/12(1-\nu^2)$ is the flexural rigidity of the plate, and ρ, h, E, and ν are, respectively, the density, thickness, Young's modulus, and Poisson's ratio of the plate. $k_{x1}=\sqrt{k_p^2-k_y^2}$ and $k_{n1}=\sqrt{k_p^2+k_y^2}$ are the wave numbers along the x direction in plate 1 of the propagating and near-field decay waves, respectively, and are derived from the plate classical equation of motion.[6] If the two plates are of the same material and thickness, the wave numbers in the x direction in plate 2 are the same as those in plate 1, that is, $k_{x2}=k_{x1}$ and $k_{n2}=k_{n1}$.

The 12 unknown coefficients A_i^p, B_i^p, C_i^p, and D_i^p for $i=1-3$ can be determined from (i) structural boundary conditions at $x_1=L_{x1}$ and $x_2=L_{x2}$, (ii) continuity equations at the driving force location (x_0,y_0), and (iii) continuity equations at the structural junction. At the free ends corresponding to $x_1=L_{x1}$ and $x_2=L_{x2}$, the bending moment and net vertical shear force are zero, and are, respectively, described by the following classical relations:[9]

$$\frac{\partial W}{\partial x^2}+\nu\frac{\partial W}{\partial y^2}=0, \tag{5}$$

$$\frac{\partial^3 W}{\partial x^3}+(2-\nu)\frac{\partial^3 W}{\partial x\,\partial y^2}=0. \tag{6}$$

At the driving force location, there are four coupling equations to describe the continuity of the plate response, corresponding to continuity of the plate displacement, slope, moment, and shear force:

$$W_{1_I}(x_0,y)=W_{1_{II}}(x_0,y), \tag{7}$$

$$\frac{\partial W_{1_I}(x_0,y)}{\partial x}=\frac{\partial W_{1_{II}}(x_0,y)}{\partial x}, \tag{8}$$

$$\frac{\partial^2 W_{1_I}(x_0,y)}{\partial x^2}=\frac{\partial^2 W_{1_{II}}(x_0,y)}{\partial x^2}, \tag{9}$$

$$\frac{\partial^3 W_{1_I}(x_0,y)}{\partial x^3}-\frac{\partial^3 W_{1_{II}}(x_0,y)}{\partial x^3}=\frac{2F_0}{L_y D}\sin(k_y y_0)\sin(k_y y). \tag{10}$$

The four continuity equations at the junction of the two plates corresponding to $x_1=0$ and $x_2=0$ have been previously established by Cremer *et al.*[1] These continuity equations are presented in this paper for both completeness of the analytical model and in accordance with the sign convention shown in Fig. 1. The first boundary condition corresponds to continuity of the bending moment from plate 1 to 2 resulting in

$$\frac{\partial^2 W^p_{1_{II}}}{\partial x^2} = \frac{\partial^2 W^p_2}{\partial x^2}. \tag{11}$$

The second boundary condition corresponds to continuity of slope, which yields

$$\frac{\partial W^p_{1_{II}}}{\partial x} = \frac{\partial W^p_2}{\partial x}. \tag{12}$$

The shear force in plate 1 induces a longitudinal force in plate 2. Similarly, the shear force in plate 2 generates a longitudinal force in plate 1. Also, at the junction, the amplitude of the flexural displacement in plate 1 is equal to the amplitude of the longitudinal waves in plate 2, and similarly, the amplitude of the flexural displacement in plate 2 is equal to the amplitude of the longitudinal waves in plate 1. Hence, the third and fourth boundary conditions can be given by[1]

$$D\frac{\partial^3 W^p_{1_{II}}}{\partial x^3} = j\omega\rho h c_L W^p_{1_{II}}, \tag{13}$$

$$D\frac{\partial^3 W^p_2}{\partial x^3} = -j\omega\rho h c_L W^p_2, \tag{14}$$

where $c_L = \sqrt{E/\rho(1-\nu^2)}$ is the longitudinal wave speed and $k_L = \omega/c_L$.

Using the equations for the boundary and continuity conditions, the 12 unknown coefficients can be obtained from the matrix expression $[\alpha_p][\mathbf{A_p}] = [\mathbf{F_p}]$, where $[\alpha_p]$ is a 12×12 matrix, and is given in Appendix A. The coefficient matrix $[\mathbf{A_p}]$ and the force matrix $[\mathbf{F_p}]$ are given by

$$[\mathbf{A_p}] = [A^p_1 \quad B^p_1 \quad C^p_1 \quad D^p_1 \quad A^p_2 \quad B^p_2 \quad C^p_2 \quad D^p_2 \quad A^p_3 \quad B^p_3 \quad C^p_3 \quad D^p_3]^T, \tag{15}$$

$$[\mathbf{F_p}] = \left[0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad \frac{2F_0}{L_y D}\sin k_y y_0 \quad 0 \quad 0 \quad 0 \quad 0\right]^T. \tag{16}$$

Solutions for the unknown coefficients of matrix $[\mathbf{A_p}]$ are determined by $[\mathbf{A_p}] = [\alpha_p]^{-1}[\mathbf{F_p}]$.

III. ACTIVE CONTROL

A single point control force of amplitude F_s is used to generate secondary excitation at a location of (x_s, y_s) in plate 1. Expressions for the secondary flexural displacement in three regions of the L-shaped structure can be obtained in the same way as for the primary displacement. The secondary displacements $W^s_{1_I}$, $W^s_{1_{II}}$, and W^s_2 are functions of wave coefficients A^s_i, B^s_i, C^s_i, and D^s_i (for $i=1-3$) and mode number m', for the three regions $L_{x1} \leqslant x_1 \leqslant x_s$, $x_s \leqslant x_1 \leqslant 0$, and $0 \leqslant x_2 \leqslant L_{x2}$, respectively. Solutions for the unknown coefficients can also be determined by $[\mathbf{A_s}] = [\alpha_s]^{-1}[\mathbf{F_s}]$. $[\alpha_s]$ is the same as matrix $[\alpha_p]$ but is instead a function of m', and with x_0 replaced by x_s. Similarly, $[\mathbf{F_s}]$ is the same as $[\mathbf{F_p}]$ in Eq. (16), but F_0 and y_0 are replaced by F_s and y_s.

A cost function is developed based on the feedforward adaptive lms algorithm to minimize the far-field flexural energy transmission, by means of minimizing the total squared plate displacement at the error sensor. In order to perform the control processing, the cost function needs to be expressed as a quadratic function of the control force. Adopting the approach described in Ref. 6, the primary and secondary plate flexural displacements can be described by $W^p_j = [\mathbf{A_p}] \times [E^p_j]\sin(k_y y)$ and $W^s_j = [\mathbf{A_s}][E^s_j]\sin(k_y y)$, for $j=1,2$, where

$$[E^p_{1_I}] = [e^{-jk_{x1}x_1} \quad e^{jk_{x1}x_1} \quad e^{-k_{n1}x_1} \quad e^{k_{n1}x_1} \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]^T, \quad L_{x1} \leqslant x_1 \leqslant x_0, \tag{17}$$

$$[E^p_{1_{II}}] = [0 \quad 0 \quad 0 \quad 0 \quad e^{-jk_{x1}x_1} \quad e^{jk_{x1}x_1} \quad e^{-k_{n1}x_1} \quad e^{k_{n1}x_1} \quad 0 \quad 0 \quad 0 \quad 0]^T, \quad x_0 \leqslant x_1 \leqslant 0, \tag{18}$$

$$[E^p_2] = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad e^{-jk_{x2}x_2} \quad e^{jk_{x2}x_2} \quad e^{-k_{n2}x_2} \quad e^{k_{n2}x_2}]^T, \quad 0 \leqslant x_2 \leqslant L_{x2}, \tag{19}$$

and $[E^s_{1_I}]$, $[E^s_{1_{II}}]$, and $[E^s_2]$ are similar to Eqs. (17)–(19), but are functions of m' instead. Since the primary and secondary excitation matrices ($[\mathbf{F_p}]$ and $[\mathbf{F_s}]$) have a nonzero element in one row only, the coefficients matrices $[\mathbf{A_p}]$ and $[\mathbf{A_s}]$ can be written in terms of a single column of the inverse matrices $[\alpha_p]^{-1}$ and $[\alpha_s]^{-1}$, respectively, resulting in $[\mathbf{A_p}] = [\alpha_p]^{-1}_{n,8}[\mathbf{F_p}]_8$ and $[\mathbf{A_s}] = [\alpha_s]^{-1}_{n,8}[\mathbf{F_s}]_8$, where $n=1$ to 12. $[\alpha]^{-1}_{n,8}$ corresponds to the eighth column of the inverse matrix $[\alpha]^{-1}$. Hence, the coefficient matrices can be written as

$$[\mathbf{A_p}]=[\alpha_p]^{-1}_{n,8}\frac{2F_0}{L_yD}\sin(k_y y_0), \tag{20}$$

$$[\mathbf{A_s}]=[\alpha_s]^{-1}_{n,8}\frac{2F_s}{L_yD}\sin(k_y y_s). \tag{21}$$

It is now possible to describe the flexural displacement of the L-shaped plate in terms of the primary and control force amplitudes by

$$W^p_j(x,y)=F_0G_p,$$

where

$$G_p=\sum_{m=1}^{\infty}\frac{2}{L_yD}\sin(k_y y_0)[\alpha_p]^{-1}_{n,8}[E^p_j]\sin(k_y y), \tag{22}$$

$$W^s_j(x,y)=F_sG_s,$$

where

$$G_s=\sum_{m'=1}^{\infty}\frac{2}{L_yD}\sin(k_y y_s)[\alpha_s]^{-1}_{n,8}[E^s_j]\sin(k_y y). \tag{23}$$

The flexural energy transmission is to be minimized at an error sensor location in the far field of plate 2. The total flexural displacement in plate 2 is the superposition of the primary transmitted flexural waves and the secondary flexural waves generated by the point control force, that is, $W^{\text{total}}_2(x_2,y)=W^p_2(x_2,y)+W^s_2(x_2,y)$. The total squared plate displacement at an error sensor location in plate 2 of (x_e,y_e) can now be expressed as a quadratic function of the control force by

$$W^{\text{total}}_2(W^{\text{total}}_2)^*=F_sG_sG^*_sF^*_s+F^*_sG^*_sG_pF_0+F_sG_sG^*_pF^*_0+F_0G_pG^*_pF^*_0. \tag{24}$$

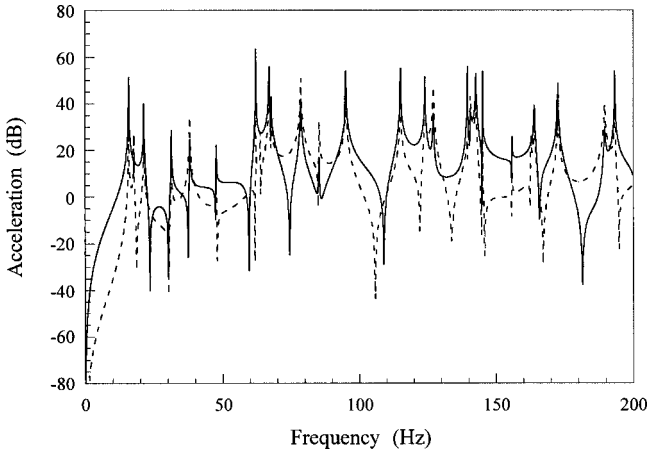


FIG. 2. Acceleration distributions at positions I (——) and V (- - -).

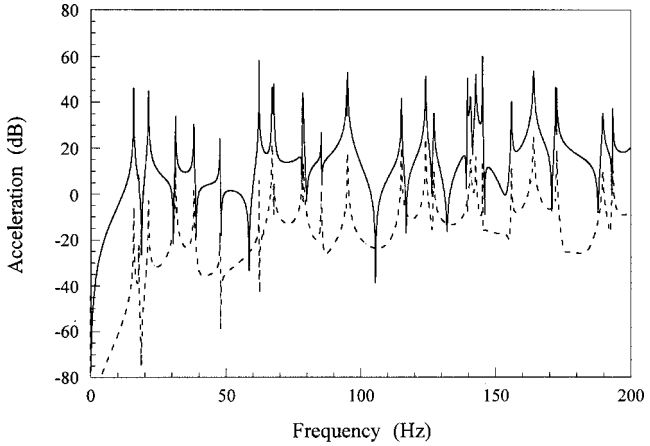


FIG. 3. Acceleration distributions at positions II (——) and III (- - -).

By standard means, the optimal control force resulting in the minimum averaged squared plate displacement is obtained as[6]

$$F_s|_{\text{opt}}=-F_0\frac{G_pG^*_s}{G_sG^*_s}. \tag{25}$$

It is important to note that due to orthogonality of mode relationships, the products $G_pG^*_p$, $G_pG^*_s$, and $G^*_pG_s$ are functions of mode number m, whereas the product $G_sG^*_s$ is a function of mode number m'.

IV. RESULTS AND DISCUSSION

The properties of aluminum were chosen for the material parameters, where $\rho=2700\,\text{kg/m}^3$, $E=6.9\times10^{10}\,\text{N/m}^2$, and $\upsilon=0.33$. In the analytical modeling, the internal distributed damping in the structure was included in the complex Young's modulus by $\tilde{E}=E(1+j\eta)$, where $\eta=0.001$ is the structural loss factor. The two plates have the same thickness of $h=0.0016\,\text{m}$ and width $L_y=0.5\,\text{m}$. The length of plate 1 is $L_{x1}=-1.2\,\text{m}$ (the negative sign is due to the sign convention as shown in Fig. 1) and the length of plate 2 is $L_{x2}=0.6\,\text{m}$. The primary shaker was located at a position of $(x_0,y_0)=(0.618L_{x1},0.618L_y)=(-0.74\,\text{m},0.31\,\text{m})$, which

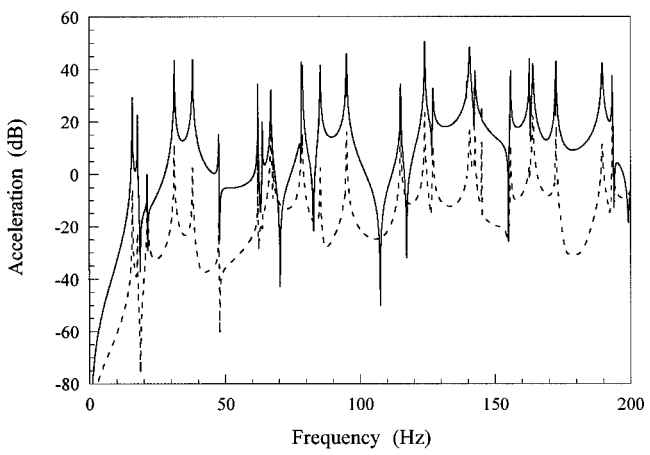


FIG. 4. Acceleration distributions at positions IV (——) and VI (- - -).

TABLE I. Positions of the accelerometer (in meters) in the three sections of the L plate to measure the primary frequency responses.

Position I	(−1.0, 0.31)	Section I, plate 1
Position II	(−0.54, 0.31)	Section II, plate 1
Position III	(−0.002, 0.31)	Section II, plate 1
Position IV	(0.002, 0.31)	Plate 2
Position V	(0.4, 0.31)	Plate 2
Position VI	(0.25, 0.2)	Plate 2

results in simultaneous excitation of all structural modes. The amplitude of the primary shaker is unity for all frequencies.

Feedforward adaptive control was used in order to attenuate the plate vibration at the error sensor location. As evident in the analytical model, simultaneous control of all the structural modes that contribute to the response at the frequency of interest was performed. Both single- and broadband frequency control were conducted, varying both the control force and error sensor locations. All control and error sensor locations are given in meters.

A. Primary frequency response

Figures 2–4 correspond to the frequency responses up to 200 Hz at the six accelerometer locations defined in Table I. Figure 2 compares the frequency responses obtained in section I of plate 1 (position I) and in plate 2 at position V. Figure 2 shows that the frequency responses at any accelerometer location in the L-shaped structure are similar, where the peaks correspond to structural resonances. For some frequencies, the amplitudes of the resonance peaks in plate 2 are greater than those in plate 1 at the same frequencies. Figure 3 compares the frequency responses obtained in the near field (position III) and far field (position II) of the corner junction in section II of plate 1. Figure 3 shows that the response curves at locations which are close to the corner junction are generally lower than elsewhere. Finally, Fig. 4 compares the frequency responses obtained at two different locations in plate 2 corresponding to positions IV and VI. Figure 4 shows that close to the corner junction in plate 2, the response curves are lower. These primary results are consistent with those in Ref. 5.

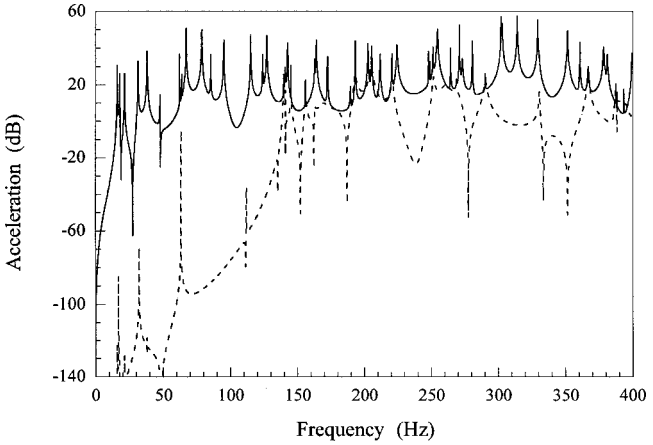


FIG. 6. Primary (——) and controlled (-----) acceleration distributions at primary force, control force, and error sensor locations of (−0.74, 0.31), (−0.54, 0.19), and (0.37, 0.25), respectively.

B. Effect of error sensor location under broadband frequency control

Broadband frequency control up to 400 Hz was performed at a single discrete location in plate 2. Initial investigation of the effect of the error sensor location in plate 2 on the control performance revealed that the y_e coordinate of the error sensor affected the control performance significantly greater than the x_e coordinate, provided that the x_e location is in the far field of the corner junction. A fixed control force location in plate 1 was chosen to be symmetrical with the primary force location in the y direction. For the fixed control force location of $(x_s, y_s) = (-0.54, 0.19)$, three error sensor positions at a fixed x_e coordinate of $0.618 L_{x2}$ were chosen in order to investigate the effect of the error sensor location on the control performance. The error sensor locations in plate 2 correspond to $(x_{e1}, y_{e1}) = (0.37, 0.31)$, which is in line with the primary force along the x direction; $(x_{e2}, y_{e2}) = (0.37, 0.19)$, which is in line with the control force in the x direction; and $(x_{e3}, y_{e3}) = (0.37, 0.25)$, which is

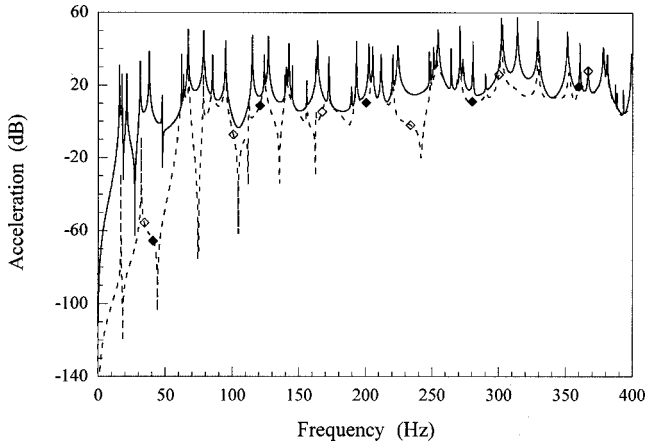


FIG. 5. Primary (——) and controlled acceleration distributions at a fixed control force location in plate 1 of (−0.54, 0.19), and error sensor locations in plate 2 of (0.37, 0.31) (--◆--) and (0.37, 0.19) (--◇--).

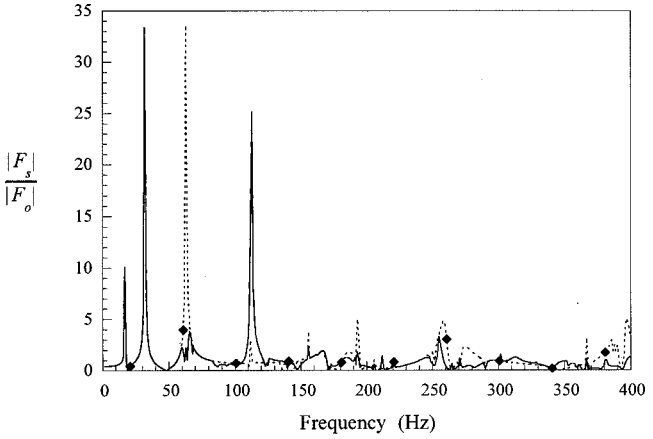


FIG. 7. Optimal control force amplitudes at a fixed control force location of (−0.54, 0.19), and error sensor locations of (0.37, 0.31) (——), (0.37, 0.19) (- - -), and (0.37, 0.25) (--◆--).

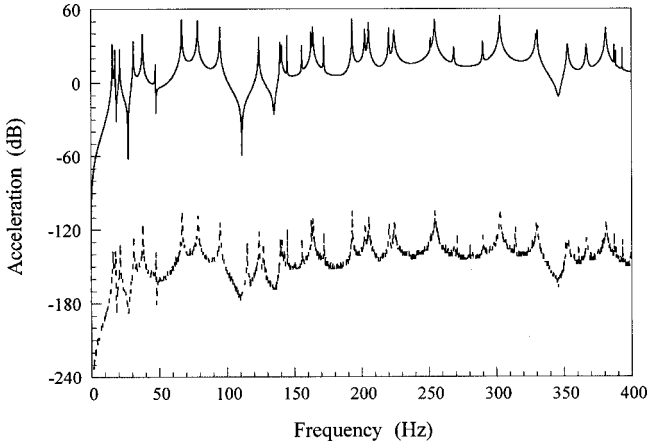


FIG. 8. Primary (——) and controlled (-----) acceleration distributions at a primary force location of (−0.74, 0.31), and the optimized control force and error sensor locations of (−0.74, 0.19) and (0.37, 0.25), respectively.

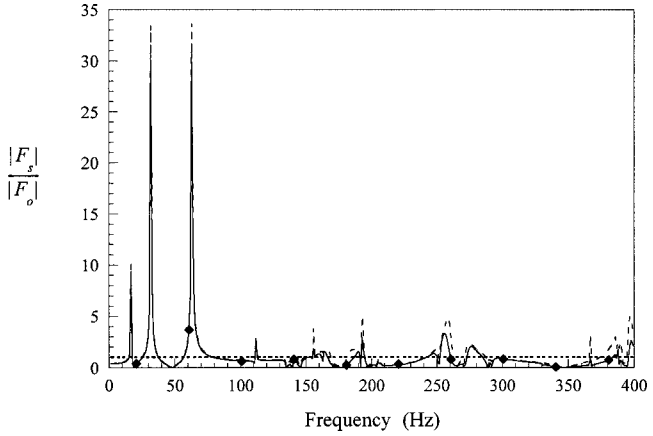


FIG. 10. Optimal control force amplitudes at a fixed error sensor location of (0.37, 0.25), and the control force locations of (−0.74, 0.19) (· · · ·), (−0.54, 0.31) (- - -), and (−0.54, 0.25) (—◆—).

midway between the primary and control forces. Figures 5 and 6 show the uncontrolled and controlled acceleration distributions for the three error sensor locations. In Fig. 5, it is shown that there is no difference in the level of attenuation achieved when the error sensor is in line with either the primary or control forces in the x direction. Figure 6 shows that when the error sensor is located midway between the primary and control shakers, attenuation of most of the resonance peaks is achieved. The corresponding optimal control force amplitudes are shown in Fig. 7, which shows that at discrete resonant frequencies, an extremely large control force is required.

C. Effect of control force location under broadband frequency control

For a fixed error sensor location at $(x_e, y_e) = (0.37, 0.25)$ in plate 2, three control force locations in plate 1 were chosen in order to investigate the effect of the control force location on the control performance. The control force positions correspond to $(x_{s1}, y_{s1}) = (-0.74, 0.19)$, which is both in line with the primary force and symmetrical about the error sensor location in the y direction; $(x_{s2}, y_{s2}) = (-0.54, 0.31)$, which is in line with the primary force along the x direction; and $(x_{s3}, y_{s3}) = (-0.54, 0.25)$, which is in line with error sensor location along the x direction. Figure 8 shows that the optimal location for the control force corresponds to the first position, that is, when the control force is both in line with the primary force and symmetrical about the error sensor location in the y direction. Further simulations reveal that this optimal location is not dependent on the x_e coordinate, provided that the x_e location is not very close to the corner junction. Away from the optimal control force location, the control performance is decreased (Fig. 9). The corresponding optimal control force amplitudes are shown in Fig. 10. An interesting observation is that when the control force and error sensor locations are optimized in a symmetrical arrangement with respect to the primary shaker location, the corresponding optimal control force amplitude is always unity.

Figure 9 shows that when the control force and error sensor are positioned in a symmetrical arrangement, with the control force in line and symmetric with the primary force in the y direction (in plate 1), and the error sensor placed midway between the two forces in plate 2, the control performance is both maximized and frequency independent. Under

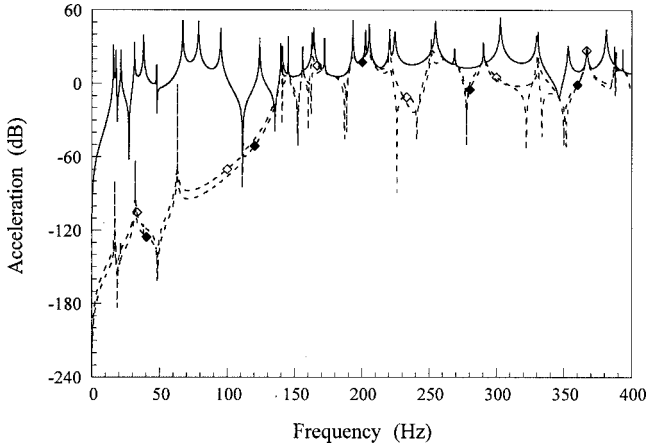


FIG. 9. Primary (——) and controlled acceleration distributions at a fixed error sensor location of (0.37, 0.25), and control force locations of (−0.54, 0.31) (--◆--) and (−0.54, 0.25) (--◇--).

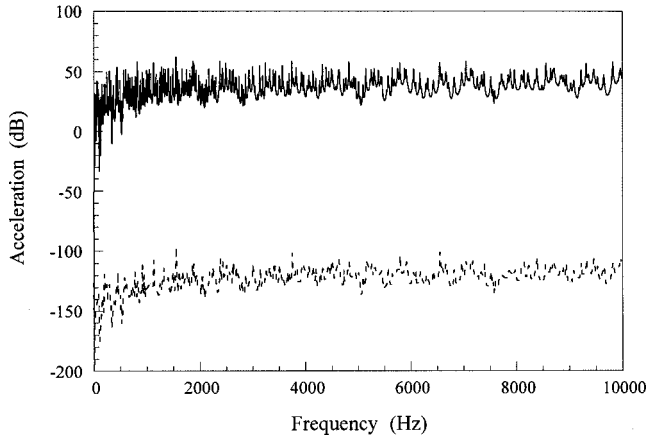


FIG. 11. Primary (——) and controlled (- - -) acceleration distributions at the optimal control force and error sensor arrangement under broadband frequency control.

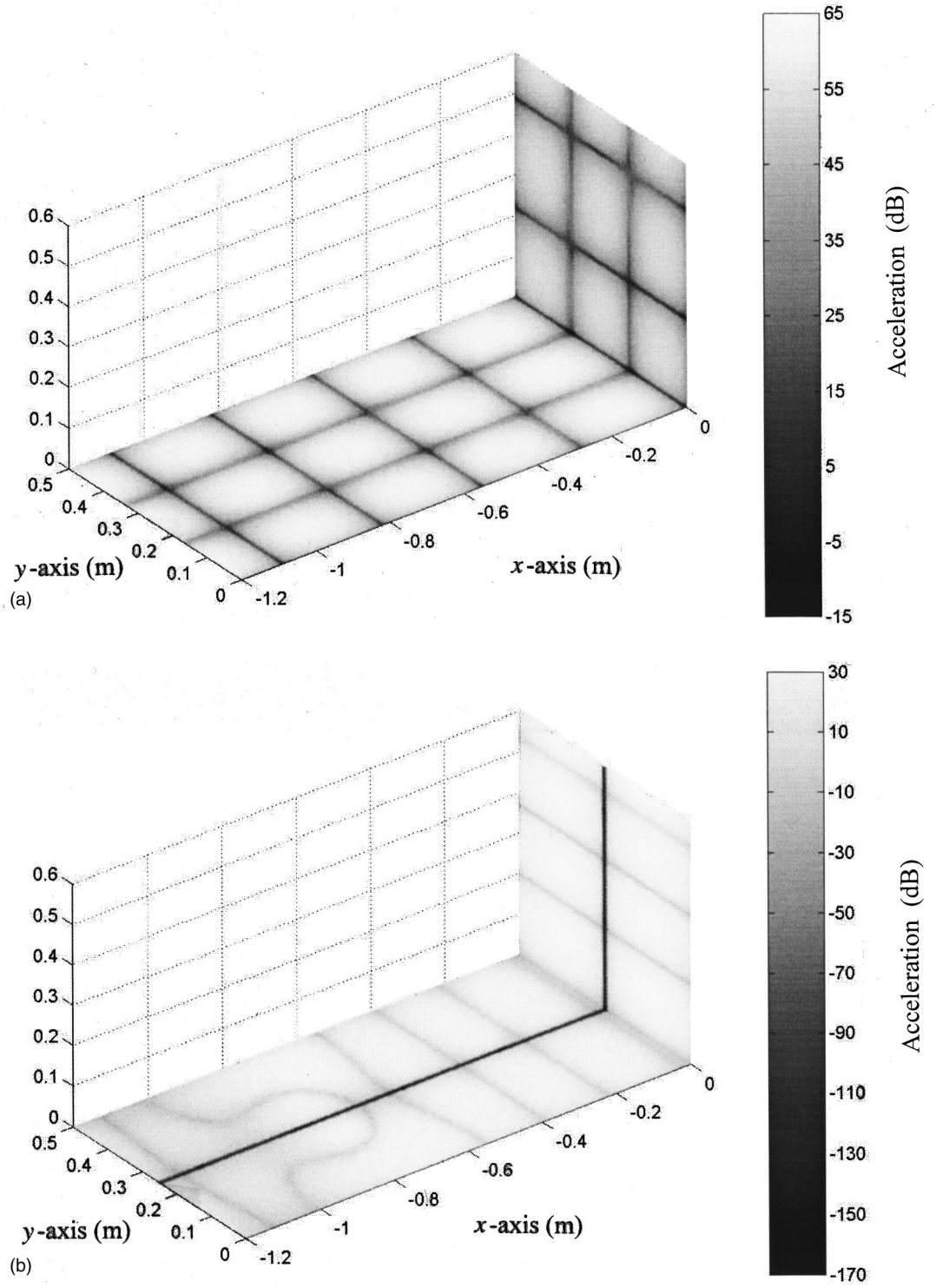


FIG. 12. (a) Contour plot of the mode shape at 193.45 Hz. (b) Contour plot of the controlled response at 193.45 Hz and an error sensor location of (0.37, 0.25) in plate 2.

this optimized application, broadband frequency control was conducted for a frequency range up to 10 kHz. Figure 11 shows that the control performance is still maximized and independent of the excitation frequency.

D. Single-frequency control

Single-frequency control is now investigated to examine the effect of the control force and error sensor locations on the global response of the L-shaped plate. Two modes of vibration were examined. The first mode corresponding to a resonance frequency of 193.45 Hz has an antinodal line midway in the *y* direction, whereas the second mode at a resonance frequency of 115.26 Hz has a nodal line located midway in the *y* direction. Figures 12 and 13, respectively, show contour plots of the uncontrolled and controlled responses for the two resonance frequencies. In Figs. 12(a) and 13(a), the mode shapes at the resonance frequencies are given, which clearly display the nodal lines. Figures 12(b) and

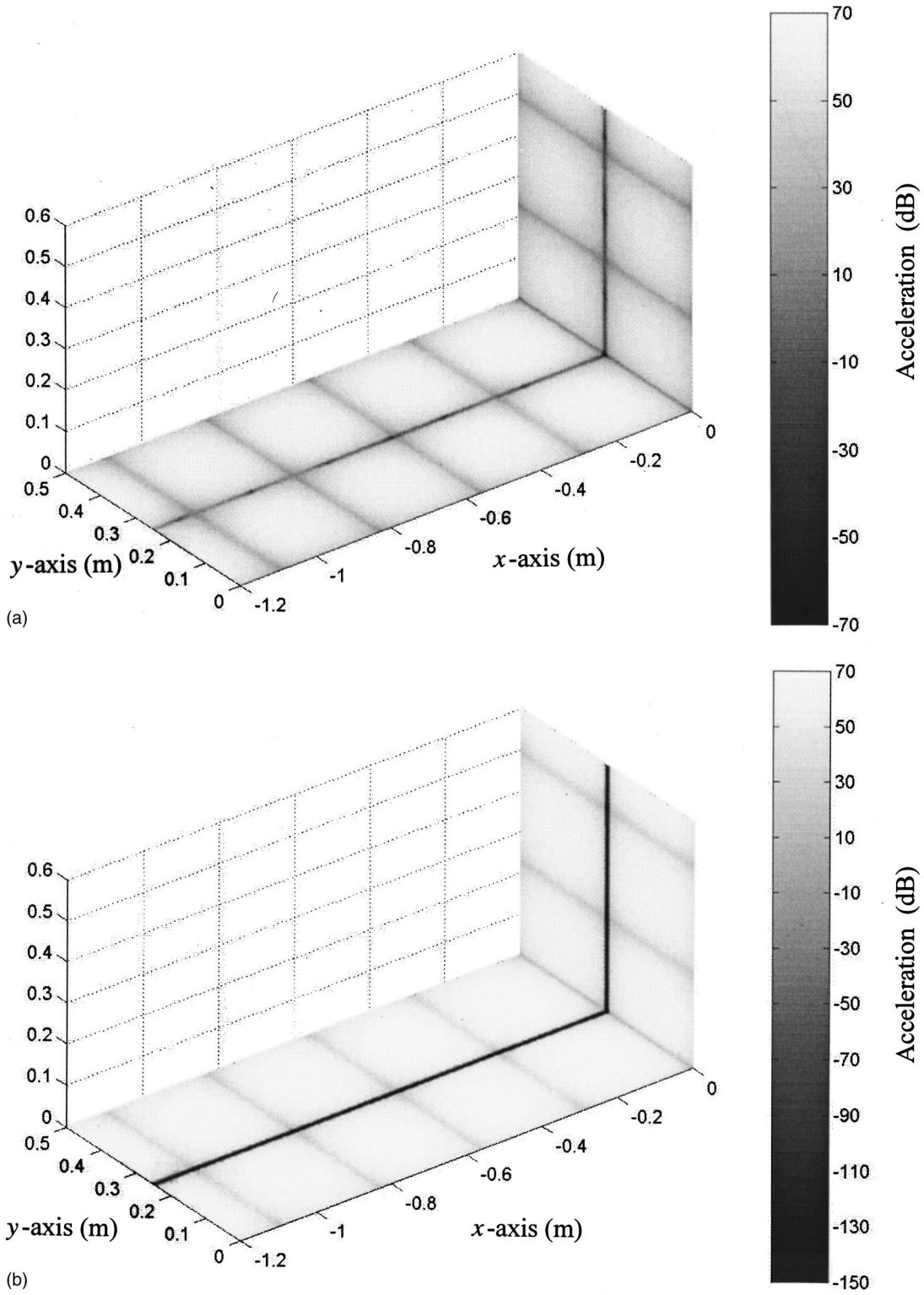


FIG. 13. (a) Contour plot of the mode shape at 115.26 Hz. (b) Contour plot of the controlled response at 115.26 Hz and an error sensor location of (0.37, 0.25) in plate 2. (c) Contour plot of the controlled response at 115.26 Hz and an error sensor location of (0.35, 0.375) in plate 2.

13(b) give the controlled response of the L-shaped plate using the symmetrical control arrangement under broadband frequency control. That is, the primary and control shaker locations in plate 1 are, respectively, $(-0.74, 0.31)$ and $(-0.74, 0.19)$, and the error sensor was located at (0.37, 0.25) in plate 2. Figure 12(b) shows that the vibrational response of the L-shaped plate at 193.45 Hz has been globally attenuated. This is attributed to the fact that the symmetrical control application results in the error sensor located at an antinodal position for this mode. However, at the resonance frequency of 115.26 Hz [Fig. 13(b)], attenuation in the L-shaped plate occurs only at the nodal line located midway in the y direction, which coincides with the error sensor location. In this case, the optimal location for broadband fre-

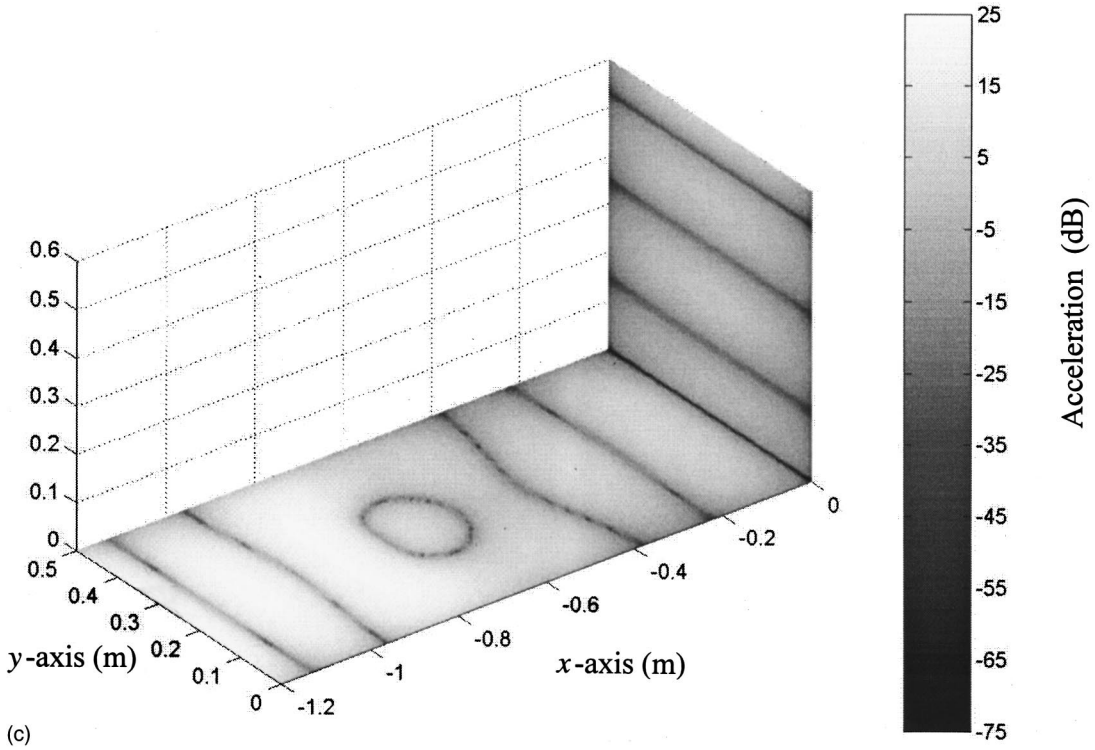


FIG. 13. (*Continued.*)

quency control is not suitable in order to achieve global attenuation of the L-shaped plate at a single resonance frequency, although the results for broadband frequency control are confirmed as there is significant attenuation along the midway line in the y direction. Figure 13(c) shows the controlled response at 115.26 Hz for the error sensor located at an antinodal point of (0.35, 0.375) in plate 2. In this case, global attenuation of the L-shaped plate is achieved.

Under single-frequency control, the error sensor location is strongly mode dependent. If the error sensor is located along a nodal line of a mode, then only attenuation of the nodal line will occur, which is clearly undesirable. However, global attenuation of the L-shaped plate can be achieved using just a single, properly located error sensor at an antinodal point.

V. CONCLUSIONS

Active attenuation of the flexural wave transmission through the junction of an L-shaped plate has been analytically investigated. Under point force excitation, a combination of a traveling wave and modal solution was used to describe the flexural displacement in the finite coupled plates. The effects of the excitation frequency on the control performance at a discrete location for various control force and error sensor locations were shown. When the locations of the control force and/or the error sensor are not optimized, the control performance is strongly dependent on the excitation frequency, that is, greater attenuation is achieved at some frequencies than others. However, as observed in Fig. 11, when both the control force and error sensor locations are optimized, global attenuation is achieved at all frequencies, and the control performance is frequency independent. The optimized arrangement of the control application corresponds to a symmetrical arrangement, with the control force in line with the primary force in the y direction, and the error sensor placed midway between the primary and control forces. The symmetrical arrangement for the optimal control application can be attributed to the simply supported edges at $y=0$ and $y=L_y$. Due to the simply supported edges, the modes are distinctly defined along the y direction. Hence, with the symmetrical control arrangement, the control force, and error sensor will, respectively, excite and observe/attenuate the same set of modes. Under broadband frequency control at a single location in plate 2, the following observations can be made:

- The error sensor is not dependent on the x_e location, provided it is located away from the corner junction.
- Global attenuation at the discrete error sensor location is achieved for all modes regardless of whether the error sensor location coincides with a nodal or antinodal line of the modes.
- Under the optimal control arrangement, the amplitude ratio of the control force to the primary force is unity for all frequencies.

Under single-frequency control, where the objective is to globally attenuate the plate response, the nodal lines of each mode play a significant role in the control performance. Global reduction of the response of the L-shaped plate can be achieved using a single, properly located error sensor at an antinodal point of each mode.

APPENDIX A

$$
[\alpha]=\begin{bmatrix}
(-k_{x1}^2-\upsilon k_y^2)e^{-jk_{x1}L_{x1}} & (-k_{x1}^2-\upsilon k_y^2)e^{jk_{x1}L_{x1}} & (k_{n1}^2-\upsilon k_y^2)e^{-k_{n1}L_{x1}} & (k_{n1}^2-\upsilon k_y^2)e^{k_{n1}L_{x1}} & 0 & 0 \\
jk_{x1}(k_{x1}^2+(2-\upsilon)k_y^2)e^{-jk_{x1}L_{x1}} & -jk_{x1}(k_{x1}^2+(2-\upsilon)k_y^2)e^{jk_{x1}L_{x1}} & -k_{n1}(k_{n1}^2-(2-\upsilon)k_y^2)e^{-k_{n1}L_{x1}} & k_{n1}(k_{n1}^2-(2-\upsilon)k_y^2)e^{k_{n1}L_{x1}} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
e^{jk_{x1}x_0} & e^{-jk_{x1}x_0} & e^{k_{n1}x_0} & e^{-k_{n1}x_0} & -e^{jk_{k1}x_0} & -e^{-jk_{x1}x_0} \\
-jk_{x1}e^{jk_{x1}x_0} & jk_{x1}e^{-jk_{x1}x_0} & -k_{n1}e^{k_{n1}x_0} & k_{n1}e^{-k_{n1}x_0} & jk_{x1}e^{jk_{x1}x_0} & -jk_{x1}e^{-jk_{x1}x_0} \\
-k_{x1}^2e^{jk_{x1}x_0} & -k_{x1}^2e^{-jk_{x1}x_0} & k_{n1}^2e^{k_{n1}x_0} & k_{n1}^2e^{-k_{n1}x_0} & k_{x1}^2e^{jk_{x1}x_0} & k_{x1}^2e^{-jk_{x1}x_0} \\
jk_{x1}^3e^{jk_{x1}x_0} & -jk_{x1}^3e^{-jk_{x1}x_0} & -k_{n1}^3e^{k_{n1}x_0} & k_{n1}^3e^{-k_{n1}x_0} & -jk_{x1}^3e^{jk_{x1}x_0} & jk_{x1}^3e^{-jk_{x1}x_0} \\
0 & 0 & 0 & 0 & -k_{x1}^2 & -k_{x1}^2 \\
0 & 0 & 0 & 0 & -jk_{x1} & jk_{x1} \\
0 & 0 & 0 & 0 & j(k_{x1}^3-k_p^4/k_L) & -j(k_{x1}^3+k_p^4/k_L) \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

$$
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & (-k_{x2}^2-\upsilon k_y^2)e^{-jk_{x2}L_{x2}} & (-k_{x2}^2-\upsilon k_y^2)e^{jk_{x2}L_{x2}} & (k_{n2}^2-\upsilon k_y^2)e^{-k_{n2}L_{x2}} & (k_{n2}^2-\upsilon k_y^2)e^{k_{n2}L_{x2}} \\
0 & 0 & jk_{x2}(k_{x2}^2+(2-\upsilon)k_y^2)e^{-jk_{x2}L_{x2}} & -jk_{x2}(k_{x2}^2+(2-\upsilon)k_y^2)e^{jk_{x2}L_{x2}} & -k_{n2}(k_{n2}^2-(2-\upsilon)k_y^2)e^{-k_{n2}L_{x2}} & k_{n2}(k_{n2}^2-(2-\upsilon)k_y^2)e^{k_{n2}L_{x2}} \\
-e^{k_{n1}x_0} & -e^{-k_{n1}x_0} & 0 & 0 & 0 & 0 \\
k_{n1}e^{k_{n1}x_0} & -k_{n1}e^{-k_{n1}x_0} & 0 & 0 & 0 & 0 \\
-k_{n1}^2e^{k_{n1}x_0} & -k_{n1}^2e^{-k_{n1}x_0} & 0 & 0 & 0 & 0 \\
k_{n1}^3e^{k_{n1}x_0} & -k_{n1}^3e^{-k_{n1}x_0} & 0 & 0 & 0 & 0 \\
k_{n1}^2 & k_{n1}^2 & k_{x2}^2 & k_{x2}^2 & -k_{n2}^2 & -k_{n2}^2 \\
-k_{n1} & k_{n1} & jk_{x2} & -jk_{x2} & k_{n2} & -k_{n2} \\
-(k_{n1}^3+jk_p^4/k_L) & (k_{n1}^3-jk_p^4/k_L) & 0 & 0 & 0 & 0 \\
0 & 0 & j(k_{x2}^3+k_p^4/k_L) & -j(k_{x2}^3-k_p^4/k_L) & -(k_{n2}^3-jk_p^4/k_L) & (k_{n2}^3+jk_p^4/k_L)
\end{bmatrix}.
$$

[1] L. Cremer, M. Heckl, and E. E. Ungar, *Structure-Borne Sound*, 2nd ed. (Springer, Berlin, 1988).
[2] J. L. Guyader, C. Boisson, and C. Lesueur, ''Energy transmission in finite coupled plates, Part I: Theory,'' J. Sound Vib. **81**, 81–92 (1982).
[3] C. Boisson, J. L. Guyader, P. Millot, and C. Lesueur, ''Energy transmission in finite coupled plates, Part II: Application to an L-shaped structure,'' J. Sound Vib. **81**, 93–105 (1982).
[4] J. M. Cuschieri, ''Structural power-flow analysis using a mobility approach of an L-shaped plate,'' J. Acoust. Soc. Am. **87**, 1159–1165 (1990).
[5] Y. Shen and B. M. Gibbs, ''An approximate solution for the bending vibrations of a combination of rectangular thin plates,'' J. Sound Vib. **105**, 73–90 (1986).
[6] X. Pan and C. H. Hansen, ''Active control of vibratory power transmission along a semi-infinite plate,'' J. Sound Vib. **184**, 585–610 (1995).
[7] N. J. Kessissoglou and J. Pan, ''An analytical investigation of the active attenuation of the plate flexural wave transmission through a reinforcing beam,'' J. Acoust. Soc. Am. **102**, 3530–3541 (1997).
[8] A. J. Young and C. H. Hansen, ''Control of flexural vibration in a stiffened plate using piezoceramic actuators and an angle stiffener,'' J. Active Control **1**, 277–301 (1995).
[9] A. W. Leissa, *Vibrations of Plates*, (American Institute of Physics, Woodbury, NY, 1993).

The prediction of façade effects from a point source above an impedance ground

Siu Hong Tang and Kai Ming Li[a)]
Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hong Kong

(Received 6 September 2000; accepted for publication 5 April 2001)

In many environmental noise prediction schemes, an empirical correction factor of 2.5 or 3.0 dB is normally added for the calculation of noise levels 1 meter in front of a reflecting façade. In this paper, theoretical and experimental studies have been conducted to examine the validity and accuracy of such approximations for a point source. The theoretical analysis involves the extension of the classical Weyl–van der Pol formula to include the effect of a reflecting façade. Experimental measurements have been conducted in an anechoic chamber to validate the theoretical predictions. It has been demonstrated that the simple empirical correction factor is adequate in cases for the assessment of the façade noise levels near the ground. However, this correction factor is not sufficient to predict the noise levels high above the ground. Some adjustments of the correction factor are required for a better estimation of sound fields in front of a reflecting façade above an impedance ground. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1377049]

PACS numbers: 43.50.Gf, 43.28.En, 43.50.Rq [MRS]

I. INTRODUCTION

The problem of noise is one of the most important environmental issues in recent years. With increased exposure and the greater sensitivity of much of the population, the number of noise-related complaints has grown at an alarming rate over the past few years. There is current legislation imposing limits of acceptable noise levels in many countries. However, in reflection of popular demand, these limits look set to be tightened further in the future. There has been an increased interest in the development of accurate numerical schemes for the prediction of environmental noise in a complex urban environment.[1–3]

A typical situation that requires the prediction of noise levels is for buildings situated close to a road of high traffic volume. As a general rule of thumb, a correction factor of 3 dB (or 2.5 dB for the guideline developed in the UK[4]) has been widely used for the calculation of noise in front of a reflecting façade. This empirical correction factor is related to sound propagation from an incoherent line source radiating noise from a continuous traffic flow. More recently,[5] a sequence of point sources has been used to model the traffic flow on a road in which each vehicle is treated as an equivalent point source. The total sound field is obtained by summing contributions from these point sources incoherently. However, many outdoor ground surfaces have different acoustical characteristics ranging from hard reflecting ground to grassland of fairly high flow resistivity and snow-covered ground of low flow resistivity.[6] It is well-known that the effect of ground impedance has greatly influenced the overall sound-pressure level received at a modest distance from the noise source. Therefore, an addition of a single empirical correction factor in front of the façade due to this reflection of sound may not be adequate to represent the whole situation.

In a series of studies,[7] we wish to investigate the interference effects of a point source and its images in a complex urban environment. In the present study, we examine a somewhat less complicated urban situation: the façade effect on the propagation of sound due to a point source above an impedance ground. Theoretical and experimental studies will be conducted to investigate the effect of the reflecting plane on the spectrum received at some distance away from the noise source. Although meteorological and topographical conditions have significant effects on sound propagation outdoors, they are ignored in the present study because our principal aim is to explore the effect of a reflecting façade above an absorbing ground.

Theoretical prediction of sound fields near a reflecting wall involves an extension of the classical Weyl–van der Pol theory to include the effect of a reflecting façade. The theory behind the prediction model is outlined in Sec. II. Indoor experiments are conducted to validate the theoretical model and they are presented in Sec. III. In Sec. IV, we use the theoretical model to calculate the sound fields above three different types of ground surfaces in front of a reflecting façade. Finally, some concluding remarks are offered in Sec. V.

II. THEORY

There is extensive literature for the sound field above an absorbing ground.[8–10] However, the asymptotic solution for a harmonic source above an absorbing ground in the presence of a reflecting façade is not readily available in the literature. Hence, it is derived and included here for future reference, although many authors would have written down the solution empirically without going through detailed asymptotic evaluation. The basic formulation for the sound field above an impedance ground in vicinity of a façade is as follows. Without loss of generality, the impedance ground is

[a)]Author to whom correspondence should be addressed. Electronic mail: mmkmli@polyu.edu.hk

 0001-4966/2001/110(1)/278/11/$18.00

located at the plane of $z=0$, the reflecting façade located at the plane of $x=0$, and a harmonic source located at $(x_s,0,z_s)$ where $x_s, z_s>0$. We wish to determine the sound field for the quarter in which $x>0$ and $z>0$; see Fig. 1 for the schematic diagram of the problem. The time-dependent factor, $e^{-i\omega t}$, is understood and suppressed for brevity in our subsequent analysis. The propagation of sound, as represented by the velocity potential, is governed by the inhomogeneous Helmholtz equation

$$\nabla^2\phi+k^2\phi=-\delta(x-x_s)\,\delta(y)\,\delta(z-z_s), \tag{1}$$

where k is the wave number given by ω/c with ω the angular frequency of the source and c the speed of sound in air. Specifying the plane of $z=0$ to be an absorbing ground of effective admittance β_z and the plane of $x=0$ to be a reflecting façade of effective admittance β_x, the velocity potential must satisfy the boundary conditions

$$\frac{\partial\phi}{\partial z}+ik\beta_z\phi=0 \quad \text{at} \quad z=0, \tag{2}$$

and

$$\frac{\partial\phi}{\partial x}+ik\beta_x\phi=0 \quad \text{at} \quad x=0. \tag{3}$$

There is an inherent assumption of the above boundary conditions: the angle between the flat ground and façade is $\pi/2$, i.e., the two surfaces are perpendicular to each other. This assumption is not too restrictive and is generally applicable for the case where the prediction of sound fields exterior to a tall building is required.

The posed problem can be solved readily by representing the solution in terms of Fourier integrals as follows. According to the theory of geometrical acoustics (see, for example, Ref. 11, Sec. 9.8), the source and its three images can be identified immediately as shown in Fig. 1, but careful considerations are necessary. We have to take into account the diffracted waves as a result of reflections from the impedance ground and façade. The positions of all ''sources'' can be identified straightforwardly as $(x_s,0,z_s)$, $(x_s,0,-z_s)$, $(-x_s,0,z_s)$, and $(-x_s,0,-z_s)$. We also note that according to the geometrical-acoustic principles,[11] there will be no wave diffraction term due to the inside edge of the wedge because its angle is $\pi/2$. Also, we ignore the diffraction wave contribution due to the impedance discontinuity (i.e., the ground and façade have different impedance) at the edge. These two assumptions can be supported by the good agreement of the theoretical prediction and experimental results for the case of ground and/or façade having different impedance; see Sec. III below.

Assuming the sound field is composed of four terms due to the source and its images, we can write the solution in the form of

$$\begin{aligned}
\phi=&-\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\exp[ik_x|x-x_s|+ik_y y+ik_z|z-z_s|]}{k^2-k_x^2-k_y^2-k_z^2}\,dk_x\,dk_y\,dk_z\\
&-\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}V_z\,\frac{\exp[ik_x|x-x_s|+ik_y y+ik_z(z-z_s)]}{k^2-k_x^2-k_y^2-k_z^2}\,dk_x\,dk_y\,dk_z\\
&-\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}V_x\,\frac{\exp[ik_x(x-x_s)+ik_y y+ik_z|z-z_s|]}{k^2-k_x^2-k_y^2-k_z^2}\,dk_x\,dk_y\,dk_z\\
&-\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}V_{xz}\,\frac{\exp[ik_x(x+x_s)+ik_y y+ik_z(z+z_s)]}{k^2-k_x^2-k_y^2-k_z^2}\,dk_x\,dk_y\,dk_z,
\end{aligned} \tag{4}$$

where V_z, V_x, and V_{xz} are the plane-wave reflection coefficients due to the reflections from the impedance ground and façade. They can be determined by applying the boundary conditions, Eqs. (2) and (3), on Eq. (4). With some tedious algebraic manipulations, we can show that

$$V_z=\frac{k_z-k_0\beta_z}{k_z+k_0\beta_z}, \tag{5a}$$

$$V_x=\frac{k_x-k_0\beta_x}{k_x+k_0\beta_x}, \tag{5b}$$

and

$$V_{xz}=V_x\times V_z=\frac{k_x-k_0\beta_x}{k_x+k_0\beta_x}\times\frac{k_z-k_0\beta_z}{k_z+k_0\beta_z}. \tag{5c}$$

Note also that negative signs are present in all integrals in Eq. (4) as a result of the specification of the problem where a source of strength -1 is used in Eq. (1).

In Eq. (4), we can identify that the first term corresponds to the direct wave term. The second and third terms correspond to waves, which have a single reflection from the impedance ground and the façade, respectively. The last term of Eq. (4) corresponds to the wave emanating from the source that hits the ground and façade before reaching the receiver. Although it is possible to evaluate these integrals numerically by means of the standard Gauss–Laguerre technique, as detailed in Ref. 9, we prefer to use the technique of contour integration for the derivation of a closed-form analytical formula which results in a better understanding of the problem. It is because the analytical approach invariably leads to a

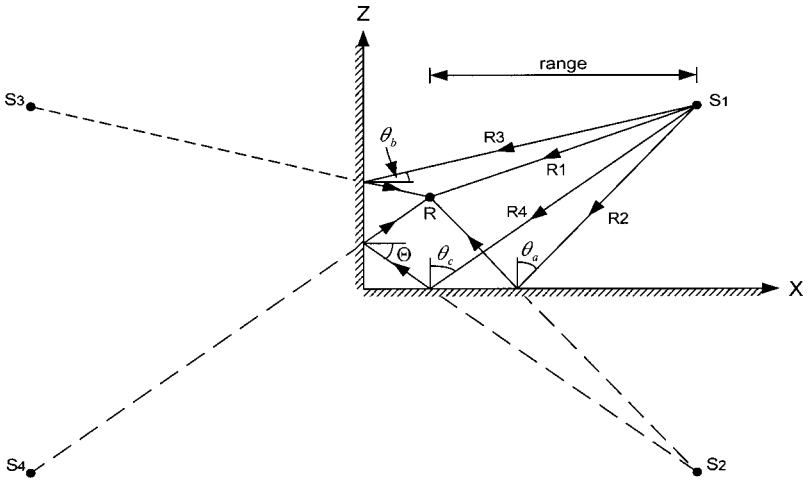


FIG. 1. The source/receiver geometry: three image sources and the sound paths due to the direct wave are shown.

physically interpretable solution in which contributions due to the reflected waves from each of the three image sources can be identified immediately.

Each of the Fourier integrals in Eq. (4) can be estimated asymptotically and their details are relegated to the Appendix. As expected, the solution can be cast in a form of the classical Weyl–van der Pol formula

$$\phi = \frac{e^{ikR_1}}{4\pi R_1} + Q(R_2,\theta_a,\beta_z)\frac{e^{ikR_2}}{4\pi R_2} + Q(R_3,\theta_b,\beta_x)\frac{e^{ikR_3}}{4\pi R_3} + Q(R_4,\theta_c,\beta_z)Q(R_4,\Theta,\beta_x)\frac{e^{ikR_4}}{4\pi R_4}, \tag{6}$$

where Q is the corresponding spherical wave reflection coefficient for the wave reflected from the ground and the façade. The path lengths R_1, R_2, R_3, and R_4 are the distances from the source and its three images to the receiver, respectively. The variables θ_a, θ_b, and θ_c are angles of incidence of reflected waves measured from the normal to the reflecting plane; see Fig. 1. Also, θ_c and Θ are related according to $\cos\Theta = \cos\psi \sin\theta_c$, where ψ is the azimuthal angle of the vertical plane that contains both source and receiver.

Knowing the distance between the image source and receiver (R), the angle of incidence of the reflected wave (θ), and the specific admittance of the reflecting surface (β), the spherical wave reflection coefficient can be determined according to

$$Q(R,\theta,\beta) = R_p + (1-R_p)F(w), \tag{7}$$

where

$$R_p = \frac{\cos\theta - \beta}{\cos\theta + \beta}, \tag{8}$$

$$F(w) = 1 + i\sqrt{\pi} w e^{-w^2}\,\mathrm{erfc}(-iw), \tag{9}$$

and

$$w = +\sqrt{\tfrac{1}{2}ikR}(\cos\theta + \beta). \tag{10}$$

We can see from Eq. (6) that there are two extra terms in the Weyl–van der Pol formula as a result of the presence of reflecting façade above an impedance ground. In fact, the surface wave term is present implicitly in the boundary loss factor $F(w)$, cf. Eq. (9).[12] Hence, all surface wave components are inherently included in the derived formula, Eq. (6). The validity of the derived formula will be confirmed by comparing with experimental measurements indoors.

We also wish to point out that there are many other more accurate asymptotic formulations (see, for example Refs. 13 and 14), and other computationally intensive numerical methodologies (see, for example, Ref. 15) for the computation of the sound fields above an impedance ground. On the other hand, a set of relatively stringent requirements ($kR \gg 1$, $|\beta_e| \ll 1$, and $\theta \approx \pi/2$) is used in deriving the spherical reflection coefficient.[8] However, numerical comparisons of various computational schemes do not reveal any significant discrepancies from that predicted by the Weyl–van der Pol formula for practical geometries and typical outdoor surfaces.[16–18] As the Weyl–van der Pol formula is now widely accepted for predicting outdoor sound,[10,12] we shall use it and its analogous form in our following analysis.

III. EXPERIMENTAL VALIDATIONS

The investigation of the predicted sound field near a building façade involves a simple mathematical modeling as shown in Sec. II. To validate the theoretical model, experiments are conducted in an anechoic chamber of size 6×6×4-m (high). The experimental model is composed of two 2.4×1.8×0.02-m (thick) hardwood boards making contact at one of their edges so that they are aligned at a right angle with each other. The sound field within the region in front of the vertical board consists of a main noise source and three images. The hardwood boards were varnished to provide smooth hard surfaces for parts of the experiments. Two different types of materials, a 1-cm-thick carpet and 4-cm-thick fiberglass, were also used to cover the hardwood board, providing surfaces of finite effective impedance. A combination of the ground and façade with different covering materials was used for measurements.

For the carpet-covered surface, it can be modeled as a locally reacting ground where a two-parameter model[19] can be used to predict the effective admittance as follows:

$$\beta_e = \frac{1}{0.436(1+i)(\sigma_e/f)^{0.5} + 19.48 i\alpha_e/f}. \tag{11}$$

In the above equation, σ_e is the effective flow resistivity and α_e is the effective rate of change of porosity with depth of the material. On the other hand, an extended reaction model should be used to estimate the impedance of the surface covered with a layer of fiberglass. For the fiberglass used in our experiments, we find that a one-parameter Delaney and Bazley model[20] can be used conveniently to calculate the effective admittance of the fiberglass-covered ground. The real and imaginary parts of the normalized impedance of the ground ($Z_1 = Z_R + iZ_X$), which are dependent on the flow resistivity of ground, are determined by

$$Z_R = 1 + 9.08(\sigma_e/f)^{0.75}, \tag{12}$$

$$Z_X = 11.9(\sigma_e/f)^{0.73}. \tag{13}$$

For the extended reaction ground, the effective admittance β_e is determined by[21]

$$\beta_e = m\sqrt{n^2 - \sin^2\theta}, \tag{14}$$

where $m(\equiv k_1/k = c/c_1)$ is the sound-speed ratio n $(\equiv \rho/\rho_1)$ is the density ratio. In these two ratios, ρ is the density of air and the subscript 1 denotes the corresponding properties of the ground. The real and imaginary parts of the wave number of the ground $k_1(k_1 = k_R + ik_x)$ are given by

$$k_R/k = 1 + 10.8(\sigma_e/f)^{0.70}, \tag{15}$$

$$k_X/k = 10.3(\sigma_e/f)^{0.59}. \tag{16}$$

Use of Eqs. (12)–(16) and the relationship of $1/Z = mn$ permit the determination of the effective admittance of the fiberglass-covered ground. In their paper, Li *et al.*[21] have demonstrated that use of Eq. (14) in the Weyl–van der Pol formula leads to accurate predictions of sound fields even for materials with flow resistivity as low as 1000 Pa s m^{-2}. Since the outdoor ground surfaces normally have much higher flow resistivity, we shall use this ''effective admittance approach'' in our subsequent numerical simulations.

A Tannoy driver with a tube of 3 cm internal diameter and 1 m long was used as a point source in all indoor experiments. The sound source was connected to a maximum length sequence system analyzer (MLSSA) with an MLS card installed in a PC.[22] The analyzer was connected to a B&K 2713 amplifier. The MLSSA system was used both as the signal generator for the source and as the signal-processing analyzer. A BSWA TECH MK224 $\frac{1}{2}$-in. condenser microphone and a BSWA TECH MA201 preamplifier were used together as the receiver. Both source and receiver were placed at a fixed position by means of a stand; clamps and the position of receiver were adjusted for different sets of measurements.

To characterize the acoustical properties of the indoor ground surfaces, prior measurements were conducted with the carpet and fiberglass, in turn, secured on the hardwood board. Theoretical predictions were compared with experimental measurements for the propagation of sound over an impedance ground, enabling the determination of the ground parameters. A graphical representation of the setup for the ground characterization is shown in Fig. 2. Measurements above the carpet and fiberglass for different source and receiver geometry were obtained. These measurements were

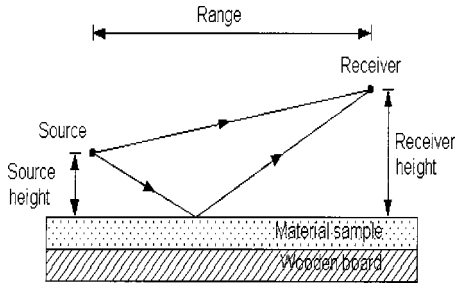


FIG. 2. The source/receiver geometry: characterization of the indoor ground surfaces.

subsequently fitted by means of a two-parameter model;[19] [see Eq. (12)], for the carpet and by means of a one-parameter model[20] for the fiberglass. Best-fit parameters were found to be $\sigma_e = 10\,000$ Pa s m^{-2} and $\alpha_e = 80$ m^{-1} for the carpet and an effective flow resistivity of 35 000 Pa s m^{-2} for fiberglass. Figures 3 and 4 display typical experimental measurements and theoretical predictions of the relative sound-pressure level with the reference free-field level measured at 1 m. These ground parameters and measurement geometries have been used in our subsequent prediction of the propagation of sound above an impedance ground in the presence of a reflecting façade.

To validate the theoretical model developed in Sec. II, three sets of indoor experiments were conducted in which the spectra of relative sound-pressure levels were obtained with different horizontal ranges between the source and receiver. Due to the physical limitations of the anechoic chamber, the range is chosen to vary from 0.6 to 1.5 m at 0.1-m intervals for each measurement. It is worth pointing out that the principal aim of our indoor measurements is to provide useful experimental data for the validation of Eq. (6). They are not intended to be proper scale model experiments, as conducted by others.[23] Hence, we made no attempt to select the most appropriate materials to model the porous ground and absorbing façade. Rather, we chose a hardwood board to model

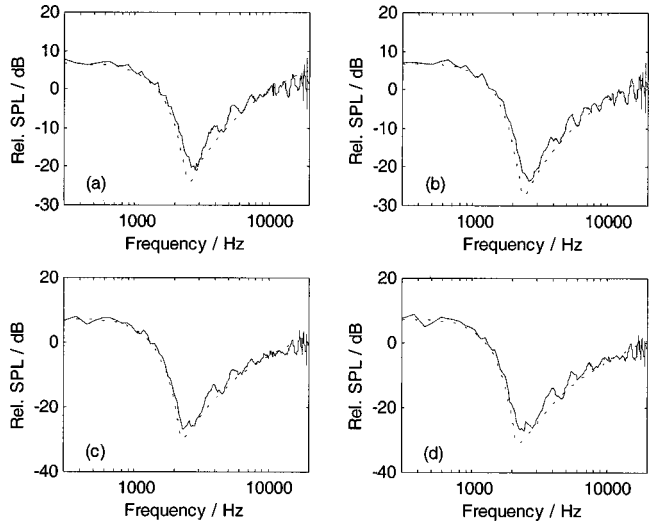


FIG. 3. Characterization of a hard ground covered with carpet. Source height = 0.05 m, receiver height = 0.04 m. Range: (a) 0.6 m; (b) 0.8 m; (c) 1.0 m; (d) 1.2 m. (Dotted line: theoretical prediction; Solid line: experimental result.)

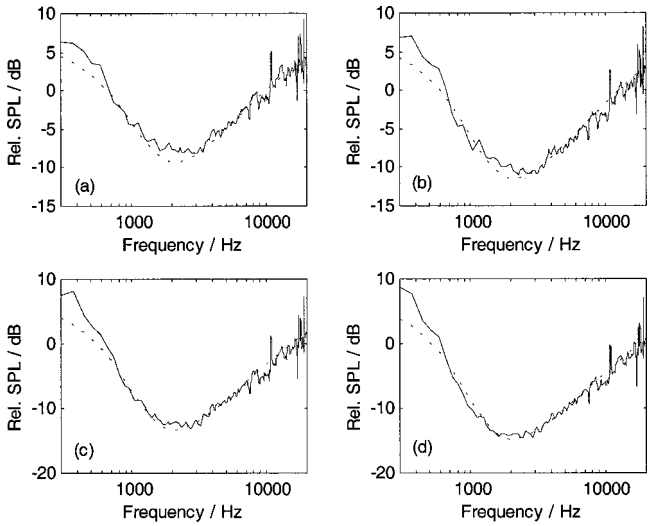


FIG. 4. Characterization of a hard ground covered with fiberglass. Source height = 0.05 m, receiver height = 0.04 m. Range: (a) 0.6 m; (b) 0.8 m; (c) 1.0 m; (d) 1.2 m. (Dotted line: theoretical prediction; Solid line: experimental result.)

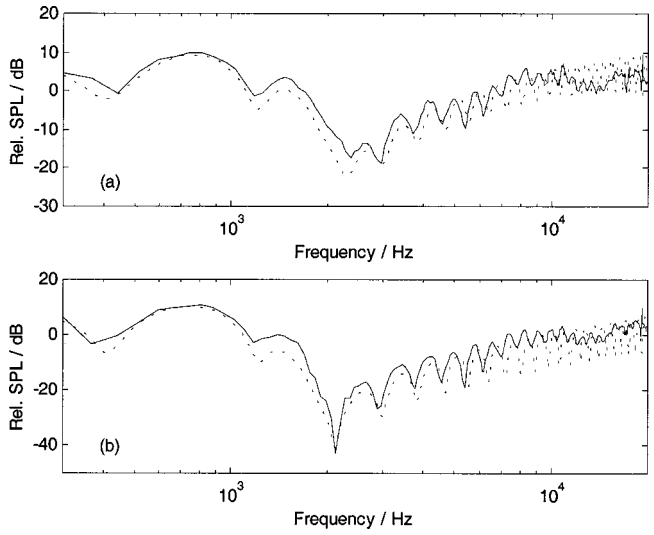


FIG. 6. Experimental data and theoretical predictions for sound propagation over a carpet-covered floor and a hard vertical wall. Source height=0.06 m, receiver height=0.05 m at 0.2 m from façade. Range: (a) 0.6 m; (b) 1.4 m. (Dotted line: theoretical prediction; Solid line: experimental result.)

a hard ground, a carpet-laid surface to model a locally reacting ground, and a surface covered with fiberglass to model an extended reaction ground. But, nevertheless, comparisons of the indoor experimental results and theoretical predictions should shed light on the validity of Eq. (6) in modeling the façade effects numerically.

In the first experimental measurements, hardwood boards were used as the floor and the vertical wall. In this case, both β_z and β_x vanish, and hence, all spherical wave reflection coefficients equal to 1. Equation (6) is used to compute the sound fields that are used to compare with the experimental measurements. Figure 5 shows measured and predicted sound fields as a function of frequency. The receiver was placed 5 cm above the floor and 20 cm in front of the vertical wall. The source was located 6 cm above the floor and at the same principal plane as the receiver, which is

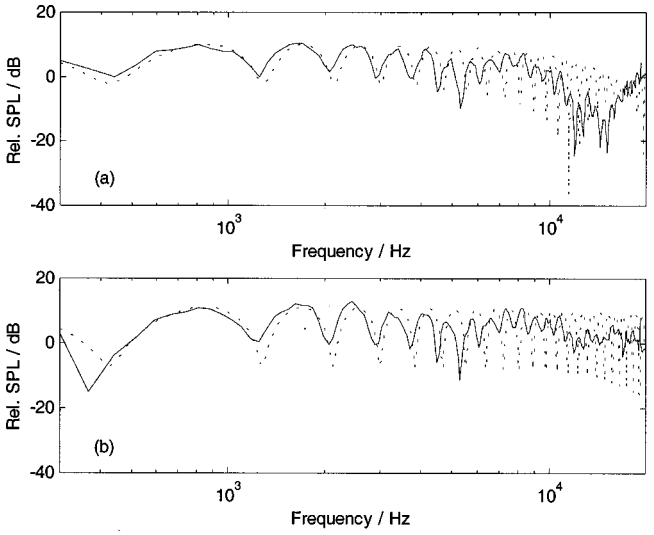


FIG. 5. Experimental data and theoretical predictions for sound propagation over a hard floor in front of a hard vertical wall. Source height=0.06 m, receiver height=0.05 m at 0.2 m from façade. Range: (a) 0.6 m; (b) 1.4 m. (Dotted line: theoretical prediction; Solid line: experimental result.)

mutually perpendicular to the ground and façade. The theoretical predictions are in accord with the experimental measurements. As expected, in the presence of a reflecting façade, the interference patterns in the frequency spectrum are much more intricate as a result of the existence of two extra terms in the Weyl–van der Pol formula.

In the second set of measurements, the carpet was secured on the surface of the horizontal board and the other hardwood board was used as the vertical wall. The orientation of the source and receiver was the same as in the first set of measurements. The source and receiver heights were now measured from the surface of the carpet instead of the floor. Figure 6 displays typical sets of measured data. Good agreement is obtained with the theoretical predictions according to Eq. (6).

In the last case, a layer of 4-cm-thick fiberglass was attached to the surface of the vertical wall, with no change of the source and receiver orientation. The experimental results for the last set of measurements are shown in Fig. 7. Again, the agreement between theoretical predictions and experimental data is generally good. We note that although there are considerable differences in the magnitude of the relative sound-pressure levels, especially at high frequencies in Fig. 5, the positions of interference ''dips'' can be predicted closely by the theoretical model. The discrepancies in the magnitude could be attributed to the fact that the wooden board (vertical wall and ground surfaces) has a nonzero effective acoustic admittance.

IV. PREDICTION OF FAÇADE EFFECTS IN THE PRESENCE OF AN IMPEDANCE GROUND

In Sec. III, we have presented experimental results for the sound fields above an impedance ground in front of a vertical wall. These experimental results agree reasonably well with the theoretical model developed in Sec. II. This gives us the confidence of using the theory to simulate a slightly more complex outdoor environment when the re-

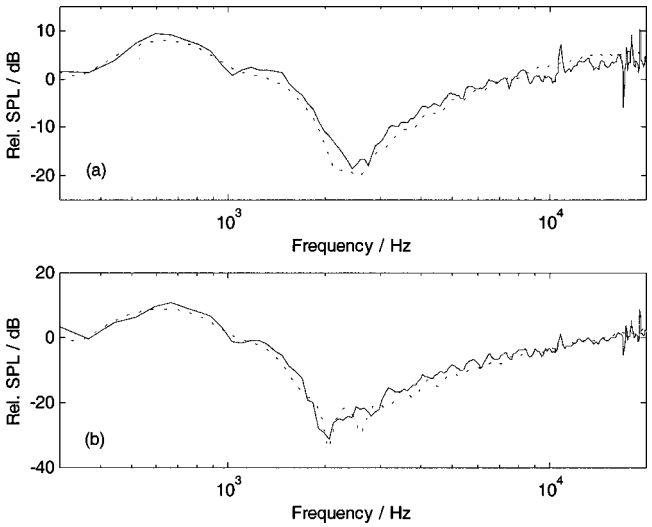


FIG. 7. Experimental data and theoretical predictions for sound propagation over a carpet-covered floor and a hard vertical wall covered with fiberglass. Source height=0.06 m, receiver height = 0.05 m at 0.2 m from façade. Range: (a) 0.6 m; (b) 1.4 m. (Dotted line: theoretical prediction; Solid line: experimental result.)

ceiver is situated above an impedance ground, and is in close proximity to a building façade. As a general rule in many environmental noise prediction schemes, a correction factor of 3 dB (2.5 dB for the guideline developed in the UK[4]) is added to account for the façade effect, as the vertical wall is assumed to be an acoustically hard surface. Here, we shall investigate the validity of such approximation as follows.

In the present study, the distribution of an A-weighted sound-pressure level (SPL) in front of a building façade is simulated. Two types of sources, noise from tire–road interactions and from engines, are taken into consideration because they represent typical sources in the context of environmental noise. The source heights used in the analysis for tire and engine exhaust are 0.01 and 0.3 m, respectively. Although it is possible for us to study the case in which the vertical wall has finite impedance, we assume that it has a perfectly reflecting surface because a building façade is normally a fairly hard surface with materials of high flow resistivity. Three different ground surfaces outside a building are considered. The first one is a typical hard ground, such as a concrete road surface, with negligible effective admittance. The second case to be considered is a simulation of grassland. The two-parameter model with $\sigma_e = 250$ kPa s m^{-2} and $\alpha_e = 100$ m^{-1} is used to calculate the impedance of the grassland. The last case is a snow-covered ground where a simple one-parameter hard-back layer model[21] is used with $\sigma_e = 20$ kPa s m^{-2} and the snow depth of 0.1 m.

In all calculations (i) both source and receiver are situated at the ''principal'' plane which is mutually perpendicular to the ground and façade, and (ii) the receiver is located 1 m in front of the façade. For the ease of interpretation, the numerical simulation is plotted for the difference between the A-weighted sound-pressure level with ($\mathrm{SPL_W}$) and without ($\mathrm{SPL_{W/O}}$) the presence of the reflecting façade

$$\Delta(z) = \mathrm{SPL_W} - \mathrm{SPL_{W/O}}.$$

The case being investigated involves the noise source situated up to 20 m from the reflecting façade and the receiver placed up to 20 m from the ground. Figures 8(a)–(c) show contour plots of the distribution $\Delta(z)$ with the source located 0.01 m above three different types of ground.

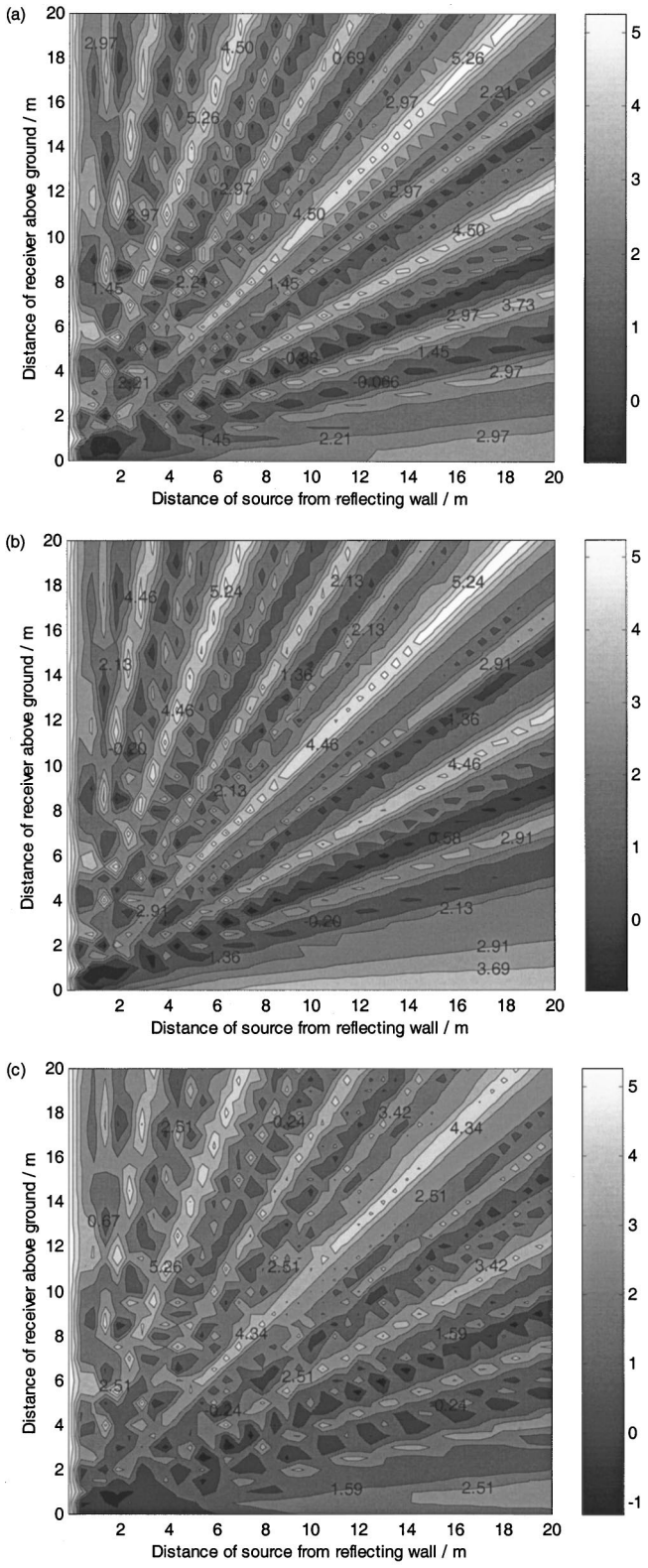


FIG. 8. Predicted difference in A-weighted sound-pressure levels with and without a reflecting façade for tire noise located at 0.01 m above the ground. (a) Hard ground, hard façade; (b) grassland, hard façade; (c) snow-covered ground, hard façade.

TABLE I. Mean values and standard deviations of A-weighted sound-pressure level difference of tire noise.

	Hard ground		Grassland		Snow-covered ground	
Distance from façade (m)	Mean (dB)	Standard deviation (dB)	Mean (dB)	Standard deviation (dB)	Mean (dB)	Standard deviation (dB)
4	2.06	1.32	2.04	1.30	2.07	1.26
8	2.23	1.25	2.19	1.27	2.05	1.15
12	2.37	1.34	2.39	1.36	2.14	1.34
16	2.47	1.30	2.51	1.35	2.24	1.23
20	2.29	1.15	2.30	1.20	1.98	1.07

In these predictions, the source has a broadband spectrum of equal energy in all octave bands, i.e., a white-noise source. The abscissa and ordinate of Fig. 8 represent, respectively, the horizontal distance of the source from the façade and the vertical distance of the receiver from the ground. In the contour plots, those regions that show a level of 3.0 dB (or 2.5 dB for the British guideline) are considered to be satisfactory in applying the empirical factor to correct for the façade effect. Predictions at other planes are not shown as they have a rather similar trend to those shown in the principal plane.

A close examination of Fig. 8 reveals that the façade correction of 2.5 or 3 dB is generally acceptable if the source is near the ground. However, the spread of Δ (i.e., the difference between the maximum and minimum values) for all three impedance grounds is about 6 dB. To allow for the assessment of a simple façade correction at all heights, it is useful to consider the mean value of Δ where the receiver is located at a given distance from the façade. The mean value, Δ, and its standard deviation, s.d., can be determined by

$$\bar{\Delta}=\frac{\int_0^{h_{\max}}\Delta(z)dz}{h_{\max}} \quad \text{and} \quad \text{s.d.}=\frac{\int_0^{h_{\max}}\Delta^2(z)dz}{h_{\max}},$$

where $h_{\max}$ is the maximum height of the receiver used in this statistical analysis. They can be computed easily by a simple numerical integration method.

Table I shows the mean values and the corresponding standard deviations of Δ for the case of tire noise with $h_{\max}$ of 20 m. The mean values and the corresponding standard deviations are close for the case of hard ground and grassland. The snow-covered ground has slightly different values from the other two grounds. This indicates that different impedance grounds do not have significant contributions towards the change in Δ because the source is very near to the ground. In general, Δ is close to the empirical correction factor of 2.5 dB. The confidence level of the predicted SPL difference in front of the façade for different source locations would be expected to fall mainly within the first deviation from the mean values (i.e., mean$\pm$standard deviation). Therefore, it may seem reasonable to assume that the addition of the empirical correction factor is an acceptable guideline in this case. However, it may not be suitable in applying this guideline for an elevated source as we investigate the case for an engine noise source located 0.3 m above the impedance ground.

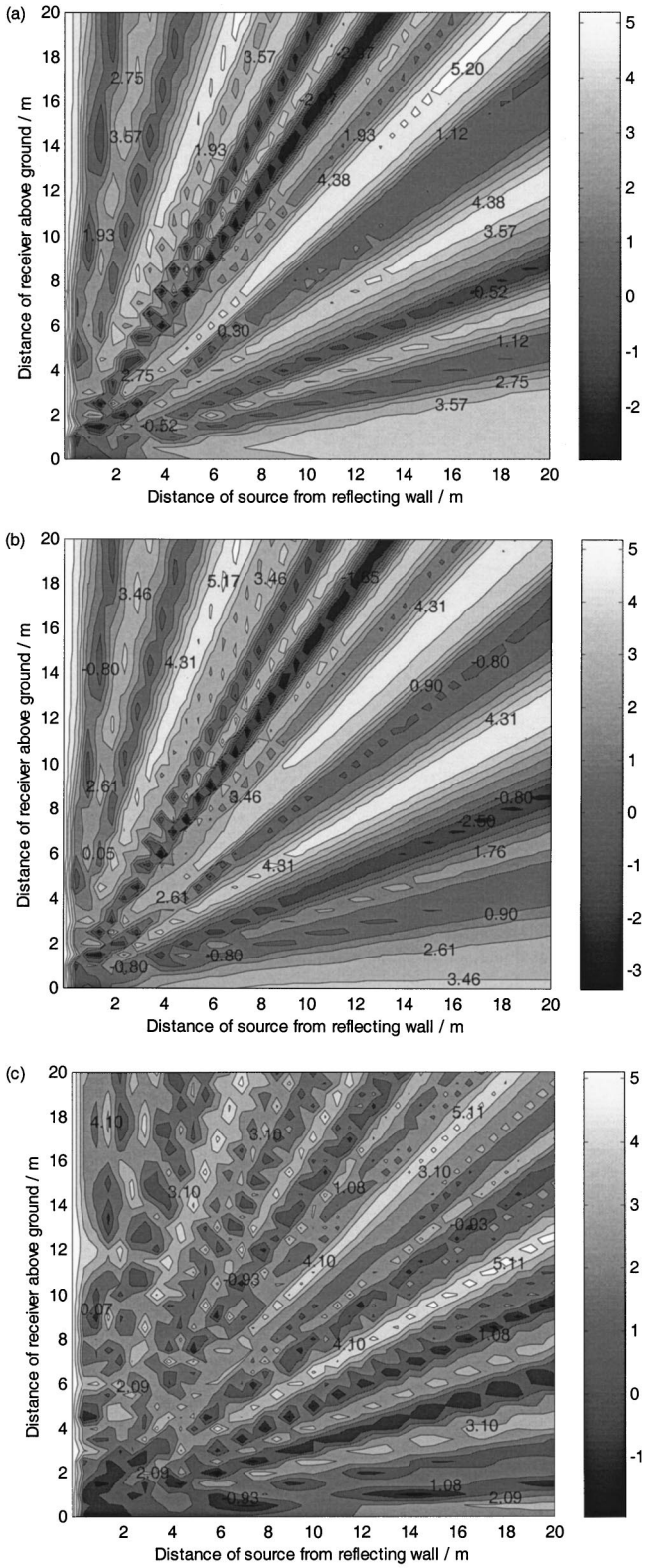


FIG. 9. Predicted difference in A-weighted sound-pressure levels with and without a reflecting façade for engine noise located at 0.3 m above the ground. (a) Hard ground, hard façade; (b) grassland, hard façade; (c) snow-covered ground, hard façade.

Similar contour plots are shown in Figs. 9(a)–(c) for the source located 0.3 m above the ground of different impedance. In this example, a measured noise spectrum for automobiles operating at 30 mph is used[24] instead of the white-

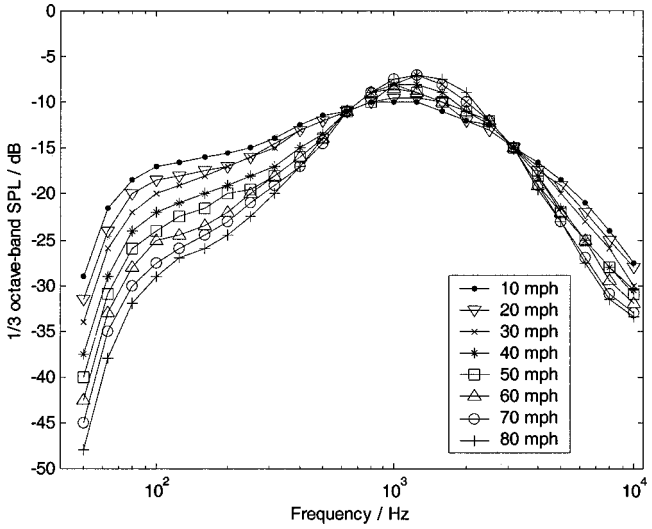


FIG. 10. Emission spectra for automobiles of different speed on average pavement one-third-octave band spectra, relative to A-weighted sound-level emissions.

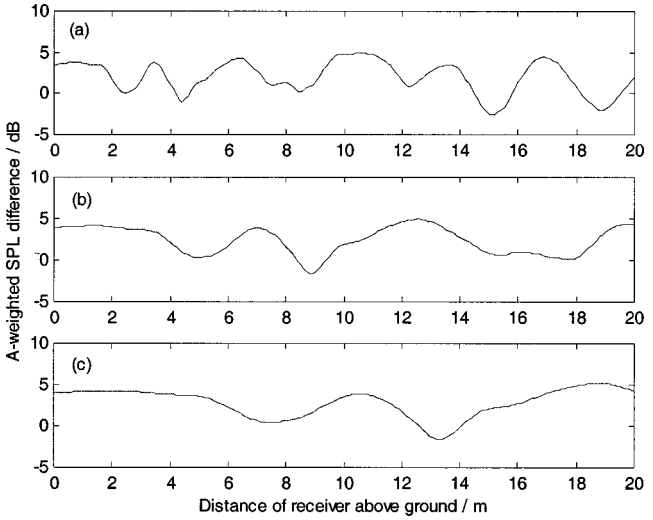


FIG. 11. The predicted A-weighted SPL difference along the height of a building when noise source (engine) is located above a hard ground at (a) 10 m; (b) 20 m; and (c) 30 m from the reflecting façade.

noise spectrum used in the case of tire noise. Figure 10 displays the noise spectra for automobiles operating at different speeds from 10 to 80 mph. It shows that their main differences are at the low frequency contents (below 300 Hz), which are less important in the A-weighted assessment. Hence, we expect that the theoretical predictions for the automobile operating at other speeds do not change significantly and, therefore, for brevity they will not be shown here.

For the contour plots of engine noise, the variation trends in Δ are rather similar to that of the tire noise contour plots. However, the spread of Δ is 7- and 8-dB for snow-covered and hard ground, respectively, and a slightly larger spread of about 8.5 dB is found for grassland. Table II shows the mean values and the corresponding standard deviations of Δ for the case of engine noise.

The standard deviation tends to increase slightly before it drops down with respect to the distance of source away from the reflecting wall. This is because when the source is about 10 m away from the façade, more changes in sound levels are detected in the prediction. Therefore, a higher degree of fluctuation of sound will be experienced. Also, it would be expected that the standard deviation decreases gradually as the source is moved further away from the façade since less fluctuation of sound level is predicted. Figures 11(a)–(c) show the relationship, with engine noise over a hard ground, between the predicted SPL difference in front of the façade due to the variation of source distance at 10, 20, and 30 m from the façade surface. When the noise source is near to a façade (in this case 10 m), a receiver in front of the wall will experience a very random change in sound level as it goes up from the ground. However, if the source moves further away from the façade to distances such as 20 or 30 m, the change in sound level that it experiences will become less random. Therefore, one would expect that at distances even further away from the façade, the SPL difference with respect to the variation of height is to be fairly stable. The phenomenon is similar for both tire and engine exhaust noise above different impedance surfaces.

TABLE II. Mean values and standard deviations of A-weighted sound-pressure level difference of engine noise.

	Hard ground		Grassland		Snow-covered ground	
Distance from façade (m)	Mean (dB)	Standard deviation (dB)	Mean (dB)	Standard deviation (dB)	Mean (dB)	Standard deviation (dB)
4	2.12	1.65	1.96	1.73	1.87	1.37
8	2.23	1.82	1.98	1.84	2.13	1.47
12	2.30	1.87	1.93	1.95	2.12	1.42
16	2.61	1.71	2.13	1.88	2.23	1.51
20	2.42	1.67	1.85	1.75	2.06	1.45

In the case of grassland, the standard deviations are slightly higher than that of the hard and snow-covered ground, hence leading to a larger range of sound-level difference for grassland observed earlier.

Better agreement with the correction factor is found in the region near the ground for both cases of noise source. This agreement is enhanced if the source is located further away from the façade. In general, the agreement works well up to 4 m above the ground if the source is located 20 m away from the façade. However, the accuracy is in doubt for the SPL difference values higher up the ground because of a rather large amount of fluctuation. For engine exhaust noise, a maximum of about 3-dB addition and a minimum of about 6-dB reduction relative to the correction factor are found in some regions in front of the façade as the source is moved away from it. A slightly lower reduction of about 4 dB is found for the case of tire noise. Therefore, the addition of a correction factor may not seem to be totally effective in regions high above the ground and some adjustments may be required.

V. CONCLUDING REMARKS

The asymptotic formula has been derived for the sound field due to a point source above a ground surface in front of a reflecting façade. The formula is validated by comparison

with precise indoor measurements conducted in an anechoic chamber. As expected, the analytical solution is composed of four terms (a source and three image sources) as a result of the reflections from the absorbing ground and façade.

We also examine the guideline of using an additional empirical correction factor of 2.5 or 3.0 dB at 1 m in front of a reflecting façade in many noise prediction schemes. It has been found that an addition of the correction factor for the A-weighted SPL is generally acceptable when the point source is close to the ground. However, the prediction is less satisfactory for an evaluated source above a ground of finite impedance. Some modifications and a stricter condition of using the empirical correction factor are required if a more accurate prediction is needed.

ACKNOWLEDGMENTS

This research was supported in part by the Research Grants Council of the Hong Kong Special Administrative Region (Project No.: PolyU 5151/99E) and the Hong Kong Polytechnic University. S.H.T. gratefully acknowledges the financial support of the Research Committee of the Hong Kong Polytechnic University for the Research Studentship.

APPENDIX: ASYMPTOTIC EVALUATION OF FOURIER INTEGRALS

Let the four separate integrals in Eq. (4) be I_1, I_2, I_3, and I_4. Using the method of contour integration (see for example, Ref. 25), the outer integral of I_1 with respect to k_z can be evaluated exactly to give

$$I_1=\frac{-i}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\exp\{ik_x|x-x_s|+ik_y y+ik_z^*|z-z_s|\}}{2k_z^*}\times dk_x\,dk_y, \tag{A1}$$

where

$$k_z^*=\sqrt{k^2-k_x^2-k_y^2}. \tag{A2}$$

The root taken for k_z^* should be positive real and negative imaginary such that finite amplitude in I_1 can be ensured. We can identify Eq. (A1) as the Sommerfeld integral,[26] which can be evaluated exactly to give

$$I_1=\frac{\exp(ikR_1)}{4\pi R_1}. \tag{A3}$$

We can see that I_1 may be regarded as the direct wave term where R_1 is the distance from the source to receiver; see Fig. 1 for the source–receiver configuration. Alternatively, we can evaluate the k_x integral by the method of contour integration to give

$$I_1=\frac{-i}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\exp\{ik_x^*|x-x_s|+ik_y y+ik_z|z-z_s|\}}{2k_x^*}\times dk_y\,dk_z, \tag{A4}$$

where

$$k_x^*=\sqrt{k^2-k_y^2-k_z^2}, \tag{A5}$$

and, again, the root for k_x^* is taken to be positive real and negative imaginary. The integral in Eq. (A4) is the Sommerfeld integral and its solution is given in Eq. (A3) as expected.

We can apply the same method in the integrals, I_2 and I_3 with the evaluation of k_z integral first for I_2 and k_x integral first for I_3. The reason for the choice of the order in the integration is obvious because of the presence of an extra term, V_z and V_x, in I_2 and I_3 respectively. We can simplify the integrals as follows:

$$I_2=\frac{-i}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{k_z^*-k_0\beta_z}{k_z^*+k_0\beta_z}\times\frac{\exp\{ik_x|x-x_s|+ik_y y+ik_z^*(z+z_s)\}}{2k_z^*}dk_x\,dk_y, \tag{A6}$$

and

$$I_3=\frac{-i}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{k_x^*-k_0\beta_x}{k_x^*+k_0\beta_x}\times\frac{\exp\{ik_x^*(x+x_s)+ik_y y+ik_z|z-z_s|\}}{2k_x^*}dk_y\,dk_z. \tag{A7}$$

Using the standard method of steepest descents with the aid of the pole subtraction method,[27] we can evaluate I_2 and I_3 straightforwardly to give the asymptotic solutions as

$$I_2=Q(R_2,\theta_a,\beta_z)\frac{e^{ikR_2}}{4\pi R_2}, \tag{A8}$$

and

$$I_3=Q(R_3,\theta_b,\beta_x)\frac{e^{ikR_3}}{4\pi R_3}. \tag{A9}$$

Finally, I_4 can be estimated asymptotically in an analogous manner, but there is a slight complication as the integrand contains the reflection factors, V_x and V_z. There are two interesting cases for consideration in the context of the façade effect on sound propagation outdoors. The first case is when the angle of incidence θ_c (see Fig. 1 for the nomenclature) is close to $\pi/2$. In this case, we can evaluate the k_z integral first to yield

$$I_4=\frac{-i}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{k_z^*-k_0\beta_z}{k_{zz}^*+k_0\beta_x}\frac{k_x-k_0\beta_x}{k_x+k_0\beta_x}\times\frac{\exp\{ik_x(x+x_s)+ik_y y+ik_z^*(z+z_s)\}}{2k_z^*}dk_x\,dk_y. \tag{A10}$$

A convenient way to evaluate the integral is to use a spherical polar coordinate (R,μ,ε) centered at the image source, S_4. The receiver is located at (R_4,θ,ψ). Making use of the substitution: $k_x=k\cos\varepsilon\sin\mu$, $k_y=k\sin\varepsilon\sin\mu$, $k_z^*=k\cos\mu$ and $dk_x\,dk_y/k_z=k\sin\mu\,d\theta\,d\varepsilon$, we can transform Eq. (A10) to

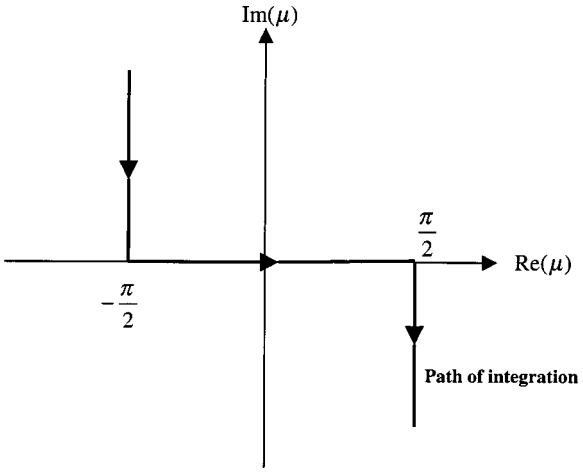


FIG. A1. The path of integration for the integral (A12).

$$I_4=\frac{-i}{2(2\pi)^2}\int_0^{\pi/2-i\infty}\int_0^{2\pi}\frac{\cos\mu-\beta_z}{\cos\mu+\beta_z}\frac{\cos\varepsilon\sin\mu-\beta_x}{\cos\varepsilon\sin\mu+\beta_x}\times e^{ikR_4\cos\theta\cos\mu}e^{ikR_4\sin\theta\sin\mu\cos(\varepsilon-\psi)}\sin\mu\,d\mu\,d\varepsilon. \tag{A11}$$

The integration with respect to ε can be evaluated by the method of stationary phase to give

$$I_4\approx\frac{ik}{4\pi}\int_{-\pi/2+i\infty}^{\pi/2-i\infty}\frac{\cos\mu-\beta_z}{\cos\mu+\beta_z}\frac{\cos\psi\sin\mu-\beta_x}{\cos\psi\sin\mu+\beta_x}\times H_0^{(1)}(kR_4\sin\theta\sin\mu)e^{ikR_4\cos\mu\cos\theta}\sin\mu\,d\mu, \tag{A12}$$

where $H_0^{(1)}(\ )$ is the Hankel function of the first kind. In obtaining Eq. (A12), the following identities[28] have been used:

$$H_0^{(1)}(z)\approx\sqrt{2/(i\pi z)}e^{iz},\quad H_0^{(2)}(z)\approx\sqrt{2i/(\pi z)}e^{-iz},$$

and

$$H_0^{(2)}(-z)=-H_0^{(1)}(z).$$

The path of integration for I_4 is shown in Fig. A1.

In addition to the required condition of θ_c close to $\pi/2$, we further restrict attention to the consideration of the situation where either the source or receiver is close to the façade. In this case, we have $\cos\psi\sin\theta_c\to1$. Hence, the only contribution due to the pole in the integrand of Eq. (A12) comes from the first quotient. We can use the method as detailed in Ref. 8 to evaluate the integral asymptotically. The approximate solution is

$$I_4\approx\frac{1}{4\pi}\frac{\cos\psi\sin\theta_c-\beta_x}{\cos\psi\sin\theta_c+\beta_x}Q(R_4,\theta_c,\beta_z)e^{ikR_4}. \tag{A13}$$

In the second case of interest, we have a small θ_c. We can use the same analysis as before to obtain the following asymptotic solution for I_4:

$$I_4\approx\frac{1}{4\pi}\frac{\cos\theta_c-\beta_z}{\cos\theta_c+\beta_z}Q(R_4,\Theta,\beta_z)e^{ikR_4}, \tag{A14}$$

where $\cos\Theta=\cos\psi\sin\theta_c$. To have a symmetric form for this reflected wave for multiple reflections from the ground and façade, we can include either a ''façade'' wave term in Eq. (A13) or a ground wave term in Eq. (A14). This leads to an interpretable expression

$$I_4\approx\frac{1}{4\pi}Q(R_4,\theta_c,\beta_x)Q(R_4,\Theta,\beta_z)e^{ikR_4}. \tag{A15}$$

[1] K. Heutshi, ''A simple method to evaluate the increase of traffic noise emission level due to buildings, for a long straight street,'' Appl. Acoust. **44**, 259–274 (1995).
[2] E. Wetzel, J. Nicolas, Ph. Andre, and J.-J. Boreux, ''Modeling the propagation pathway of street-traffic noise: Practical comparison of German guidelines and real-world measurements,'' Appl. Acoust. **57**, 97–107 (1999).
[3] E. Walerian, R. Janczur, and M. Czechowicz, ''Sound level forecasting for city centers. I: Sound level due to a road within an urban canyon,'' Appl. Acoust. **62**, 359–380 (2001).
[4] *Calculation of Road Traffic Noise* (H. M. Stationery Office, London, 1988).
[5] E. Walerian and R. Janczur, ''Effect of source model parameters on sound level in built-up area,'' Arch. Acoust. **24**, 145–160 (1999).
[6] J. Nicolas, J. L. Berry, and G. A. Daigle, ''Propagation of sound above a finite layer of snow,'' J. Acoust. Soc. Am. **77**, 67–73 (1985).
[7] K. K. Iu and K. M. Li, ''The propagation of sound in city streets,'' Proceedings of Westprac VII, Vol. 2, 811–814, Kumamoto, Japan (2000).
[8] C. F. Chien and W. W. Soroka, ''A note on the calculation of sound propagation along an impedance plane,'' J. Sound Vib. **69**, 340–343 (1980).
[9] S. N. Chandler-Wilde, ''Ground effects in environmental sound propagation,'' Ph.D. thesis, University of Bradford, 1988.
[10] T. F. W. Embleton, ''Tutorial on sound propagation outdoors,'' J. Acoust. Soc. Am. **100**, 31–48 (1996).
[11] A. D. Pierce, *Acoustics: An Introduction to its Physical Principles and Applications* (Acoustical Society of America, New York, 1989).
[12] K. Attenborough, ''Review of Ground effects on outdoor sound propagation from continuous broadband sources,'' Appl. Acoust. **24**, 289–319 (1988).
[13] T. Kawai, T. Hidaka, and T. Nakajima, ''Sound propagation above an impedance boundary,'' J. Sound Vib. **83**, 125–138 (1982).
[14] S. I. Thomasson, ''Reflection of waves from a point source by an impedance boundary,'' J. Acoust. Soc. Am. **59**, 780–785 (1976).
[15] K. Attenborough, S. Taherzadeh, H. E. Bass, X. Di, R. Raspet, G. R. Becker, A. Güdesen, A. Chrestman, G. A. Daigle, A. L'Espérance, Y. Gabillet, K. E. Gilbert, Y. L. Li, M. J. White, P. Naz, J. M. Noble, and H. A. J. M. van Hoof, ''Benchmark cases for outdoor sound propagation models,'' J. Acoust. Soc. Am. **97**, 173–191 (1995).
[16] L. R Quartararo, ''A theoretical investigation of sound propagation above a half-space of extended reaction,,'' MS thesis, Penn. State University, March, 1993.
[17] K. Attenborough, N. W. Heap, T. L. Richards, and V. V. S. S. Sastry, ''Comments on ground effect analysis: Surface wave and layer potential representation,'' J. Sound Vib. **84**, 289–295 (1982).
[18] C. H. Howorth, ''Sound propagation over rigid porous layers,'' Ph.D. thesis, The Open University, 1991.
[19] K. Attenborough, ''Ground parameter information for propagation modeling,'' J. Acoust. Soc. Am. **92**, 418–427 (1992).
[20] M. E. Delany and E. N. Bazley, ''Acoustical properties of fibrous absorbent materials,'' Appl. Acoust. **3**, 105–116 (1970).
[21] K. M. Li, T. Waters-Fuller, and K. Attenborough, ''Sound propagation from a point source over extended-reaction ground,'' J. Acoust. Soc. Am. **104**, 679–685 (1998).
[22] D. D. Rife, ''Maximum-Length Sequence System Analyzer, Reference Manual,'' Version 10W, 1987–1988.
[23] K. V. Horoshenkov, D. C. Hothersall, and S. E. Mercy, ''Scale modeling of sound propagation in a city street canyon,'' J. Sound Vib. **223**, 795–819 (1999).
[24] US Department of Transportation, *FHWA Traffic Noise Model, Technical Model* (Federal Highway Administration, Office of Environment and Planning, Washington, D.C.).
[25] J. Lighthill, ''Asymptotic behaviour of anisotropic wave systems stimulated by oscillated sources,'' in *Wave Asymptotics*, edited by P. A. Martin and G. R. Wickham (Cambridge University Press, Cambridge, 1992),

Chap. 1. See also J. Lighthill, ''Emendations to a proof in the general three-dimensional theory of oscillating sources of waves,'' Proc. R. Soc. London, Ser. A **427**, 31–42 (1990).
[26] L. M. Brekhovskikh, *Waves in Layered Media* (Academic, New York, 1980), p. 228.
[27] C. F. Chien and W. W. Soroka, ''Sound propagation along an impedance plane,'' J. Sound Vib. **43**, 9–20 (1975).
[28] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover, New York, 1972).

Silencer consisting of two-stage Helmholtz resonator with auto-tuning control

Kosuke Nagaya,[a] Yasudide Hano, and Akihisa Suda
Department of Mechanical Engineering, Gunma University, Kiryu, Gunma 376-8515, Japan

(Received 22 November 1999; revised 20 September 1999; accepted 14 March 2001)

A Helmholtz resonator has advantages over other noise control methods in that it does not require energy to function and it can be applied to high-frequency noise. In this article, a noise reduction method based on a Helmholtz resonator is discussed in reference to application to a blower. However, as the frequency of noise generated by a blower, although high, varies over time, the Helmholtz resonator is ineffectual if applied directly. In order to reduce varying high-frequency noise, a new type of silencer, comprised of a two-stage rotary auto-tuning resonator, is proposed. A fast Fourier transform (FFT) analysis of noise obtained by a noise meter is used for control. The frequencies of the peak intensities obtained by the FFT vary according to the variations in source frequency. Two significant peaks in the frequency domain are considered, the cost functions of which are taken, allowing for frequency variation. The silencer is tuned by minimizing the cost function. The system and control algorithm are presented in this report, along with the results of experimental tests that were conducted on a typical blower in order to verify the effectiveness of the system and method. [DOI: 10.1121/1.1370524]

PACS numbers: 43.50.Ki, 43.50.Gf [MRS]

I. INTRODUCTION

Noise reduction is important for a variety of machines. One popular method is active control, by which sound with opposite phase to the noise is superposed on the noise.[1–5] Such a technique requires a constant supply of energy, and the control system is complicated. When the phase between the control signal and the noise is perfectly opposite (π out of phase), this type of control works well and noise is reduced. However, if the phase between the noise and the control signal differs by any margin from this ideal, the active control signal generates noise. Hence, it is difficult to control high-frequency noise by active control, and so it is only useful in low-frequency applications. In a survey of conventional noise reduction applications, it was found that it is typically difficult to control noise in the kHz frequency range.

Passive control using a resonator does not require energy, and provides noise suppression over a wide frequency range at the kHz level.[5,6] However, passive control is only applicable to a constant frequency range, a limitation that has prevented it from being used in practice. If the resonant frequency of the resonator can be dynamically tuned to the noise frequency, the resonator will work for a wide range of frequencies. Nin *et al.*[7] and Izumi and Narikiyo[8] presented a tuned resonator in which the volume of the resonator chamber is tuned. However, such a resonator is bulky. Izumi[9] discussed the use of a compact tuned Helmholtz resonator, however, the device was limited in that the noise had to be purely sinusoidal and between 40 and 200 Hz, which is too low for practical use. Cheng *et al.*[10] presented a tunable resonator that utilized a linear motor to change the inlet such that the Helmholtz resonator could be switched between single and double. Alternatively, an adaptive algorithm may be used to tune the resonant frequency of the resonator. von Flotow, Beard, and Bailey[11] presented tuning laws for mechanical vibration absorbers, which are the vibrational equivalent to an acoustic Helmholtz resonator. Bernhard, Hall, and Jones[12] and Bedoit *et al.*[13] also discussed adaptive passive noise control. In those studies, the resonance frequency is controlled such that the resonator functions in the principal mode only. In all these studies, it has been difficult to control high-frequency noise using a resonator. A multistage resonator is an option, as it can control not only the principal mode, but also high-frequency components. The frequency response of the human auditory system is not constant, and so it is important to use A-weighted sound pressure levels in the control of resonators. Additionally, it remains difficult to suppress high-frequency noise using a self-tuning multistage resonator, and in this case, the frequency of blower noise varies over time. For these reasons, the resonators presented in these previous papers are not directly applicable to the problem of blower noise.

In the present article, a silencer is presented that has none of these limitations, and which is applicable to general noise with high-frequency components. In particular, the principal frequency is high in comparison to previously proposed methods (in the experiment, we consider 100 Hz–3 kHz), and A-weighted sound pressure levels are considered. The control algorithm presented in this paper, involving a fast Fourier transform FFT, is an important development for the multistage resonator because the previous methods are unable to control a two-stage resonator. The silencer is applicable to blower noise that has high-frequency components with frequency variation during blade rotation. The silencer is required only for the tuning of the passive resonator and does not effect the energy balance of the system under control. This system does not require an upstream reference sensor, as was required for previous methods.

[a]Author to whom correspondence should be addressed. Electronic mail: nagaya@me.gunma-u.ac.jp

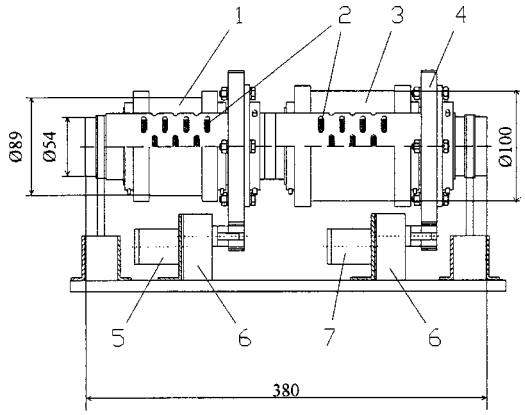


FIG. 1. Geometry of the silencer with a two-stage auto-tuning Helmholtz resonator.

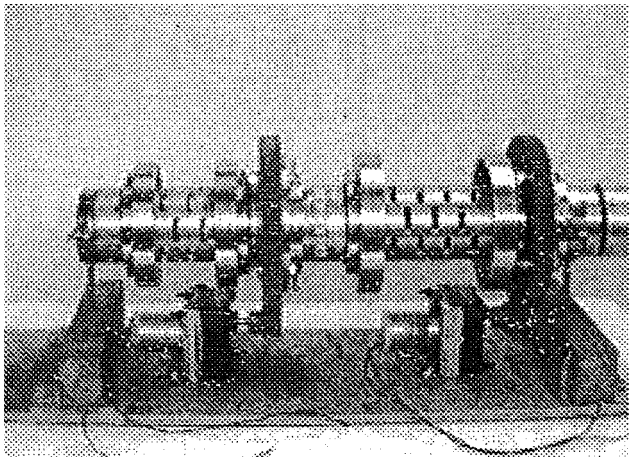


FIG. 3. Photograph of the silencer used in the experiment (outer drum is removed).

II. AUTO-TUNING SILENCER

Porous absorbers provide good broadband attenuation at high frequencies. High-frequency noise is an important consideration because it is relatively loud to the human ear. A schematic diagram of the porous absorber with auto-tuning control is shown in Fig. 1, cross sections of the silencer are shown in Fig. 2, and a photograph of the device is shown in Fig. 3. Blower noise has a number of frequency components; two significant noises in the frequency domain are considered in the present article. Hence, the silencer consists of two stages, one of which attenuates low-frequency noise and the other high-frequency noise. Each silencer consists of two inner cylinders; A and B, and an outer cylinder C, as shown in Fig. 2. There are a number of rectangular holes with rounded comers in cylinders A and B. Cylinders B and C are connected and separated by a cavity, whereas cylinders A and B are not connected and there is no gap. Hence, when cylinder C rotates, cylinder B rotates, and the area of holes between cylinders A and B varies. Cylinder C has a gear that is rotated by a motor. Looking at Helmholtz resonator 1 in Fig. 1, gear 4 is connected to the upper cylinder (cylinder C). The gear is engaged by the small gear of the motor. The upper cylinder is rotated by control motor 7 such that the area of the holes between the two inner cylinders A and B varies according to the angle of rotation. Sound is introduced into cylinder A on the right-hand end of the silencer. Thus, the sound propagates through the cavity between cylinders B and C via the holes between cylinders A and B. Damping occurs at the resonant frequency of the resonator because of the constricting effect of the holes (called the Helmholtz resonator). The resonant frequency varies according to the area of the holes. Since the area of the holes is variable in this device, the resonant frequency of the resonator can be varied by motor control. The silencer has two resonators with different resonant frequency ranges. In this experiment, a hair dryer was used as the blower. The important frequencies of the noise are in the range 100–3000 Hz. Hence, the designed resonance frequencies of the first stage are in the range 100–1000 Hz, and those of the second stage are in 1000–3000 Hz. Although energy is used to create the resonance, no energy is required to control the noise. This is the advantage of the present silencer.

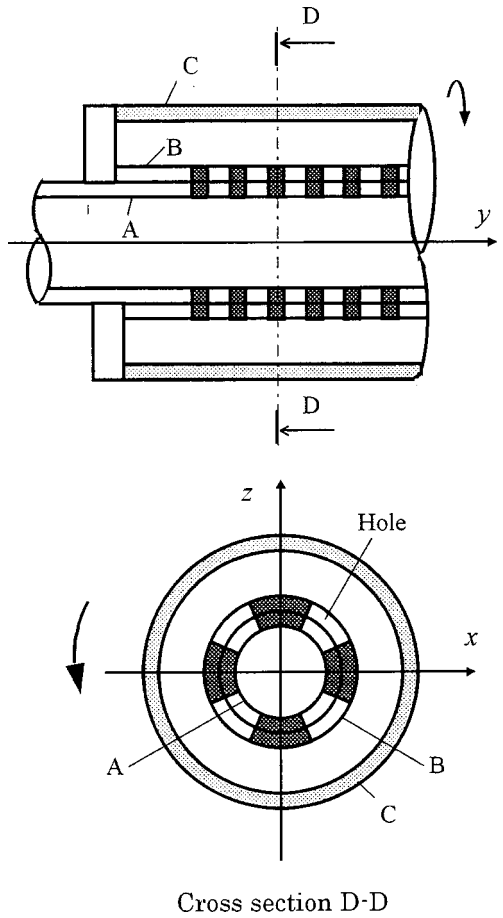


FIG. 2. Cross section of the resonator.

III. SILENCER CONTROL METHOD

In the silencers proposed in previous studies,[7–13] the resonant frequency was directly controlled, and hence, it is difficult to control noise having high-frequency components with frequency variations. With blower noise, there is a constant pressure due to the airflow, which makes control difficult for conventional control methods. Hence, previously described methods are not suitable for this problem (results of conventional methods are shown in the Appendix). In the

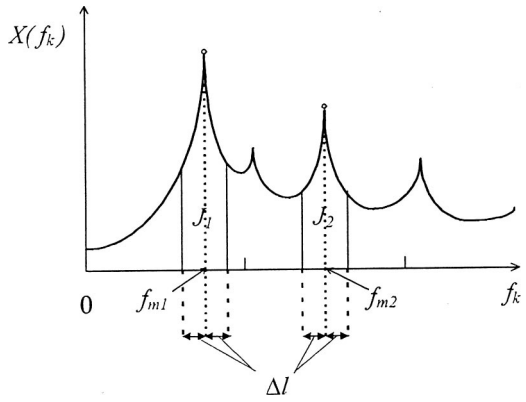


FIG. 4. Frequency variation bands for making the cost function.

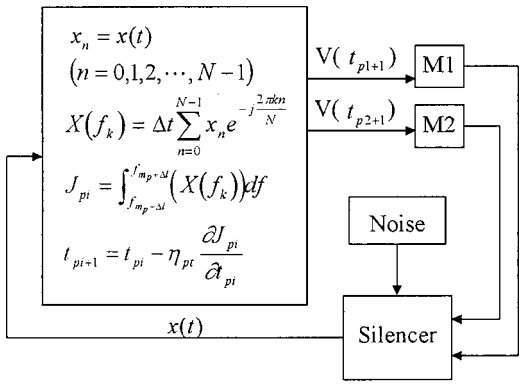


FIG. 5. Flow chart of the control to drive the control motors.

present article, a method based on the FFT analysis is presented which is applicable to blower noise because constant pressure does not affect the result of the FFT. Moreover, the method is applicable to varying high-frequency noise.

A. Control algorithm

A microphone of a noise meter is installed near the outlet of the silencer (left-hand end of the silencer in Fig. 1), and is used to detect the sound pressure of noise exiting the silencer. The signal is processed by a digital signal processor (DSP). If the sound pressure is $x(t)$, then its Fourier transformation is

$$X(\omega)=\int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt, \tag{1}$$

which is a difficult integration for real systems, so the numerical Fourier transformation is used. If the sampling time is Δt, then Eq. (1) can then be approximated to

$$X(f_k)=\Delta t\sum_{n=0}^{N-1} x_n e^{-j(2\pi kn/N)} \quad (n=0,1,2,...,N-1), \tag{2}$$

where k is the sampling number $(k=0,1,2,...,N-1)$, and f_k is the frequency.

The frequency range used in this experiment is 0.1–1.0 kHz for the first stage, and 1.0–3.0 kHz for the second stage. Hence, the following is taken to be the cost function:

$$J_{pi}=\max[X(f_k)] \quad \begin{cases} p=1 & (100\leqslant f_k\leqslant 1000), \\ p=2 & (1000\leqslant f_k\leqslant 3000), \end{cases} \tag{3}$$

where i is the sample number, $p=1$ denotes the first stage, and $p=2$ the second stage. The frequency of the peaks in blower noise varies over time, which makes it difficult to control the silencer using the cost function in Eq. (3) directly. In order to deal with this frequency variation, the cost function in Eq. (3) is integrated over a small range $2\Delta l$, which corresponds to the range of frequency variation, as shown in Fig. 4. Then, the cost function, allowing for frequency variation, is given by

$$J_{pi}=\int_{f_{m_p}-\Delta l}^{f_{m_p}+\Delta l}[X(f_k)]df \quad \begin{cases} p=1 & (100\leqslant f_k\leqslant 1000), \\ p=2 & (1000\leqslant f_k\leqslant 3000), \end{cases} \tag{4}$$

where f_{m_p} is the peak frequency. The motor is controlled so as to minimize the cost function J_p. DC motors are used for controlling this resonator, as shown in Fig. 1. Control is applied by supplying a fixed-voltage rectangular pulse signal to the motor. The angle of rotation of the motor is controlled by the pulse width. In order to achieve the optimal rotation angle of the motor, the following algorithm is applied:

$$t_{pi+1}=t_{pi}-\eta_{pt}\frac{\partial J_{pi}}{\partial t_{pi}} \quad (p=1,2,\ i=1,2,3,...,\infty), \tag{5}$$

where t is the pulse width, and η_{pt} is the weight of calculation. The pulse width is determined from Eq. (5). A pulse signal with the calculated pulse width defines the angle of rotation of the motor, which in turn controls the angle of rotation of the cylinder (cylinder C in Fig. 2) and changes the area of the holes between cylinders A and B. In this experiment, the voltage was 9 V. Continuing the operation of the motor using the pulse width obtained from Eqs. (2)–(5) minimizes the cost function. When t_{pi+1} becomes negative, the direction of the rotation of the motor should be in the opposite direction, and so the input voltage becomes negative.

B. Control program

The block diagram of this control system is shown in Fig. 5. Adaptive control is applied based on the preceding analysis. The flow of the computer program is as follows:

(1) The noise wave is sampled at a constant rate, and the mean noise is calculated. The mean value is subtracted from the noise wave. This effectively removes the dc component of the noise.
(2) The initial rectangular pulse of length t_0 (s) and voltage 9 (V) is input to the motor.
(3) After the cylinder has rotated, the noise is sampled again, and the Fourier transformation based on Eq. (3) is performed.
(4) The cost function J_0 is calculated using Eq. (4).
(5) Operations (2)–(4) are repeated for the period t_1 (s), and cost function J_1 is calculated. Operations (2)–(5) comprise the initial operation.
(6) Period t_{i+1} (s) is calculated using Eq. (5).

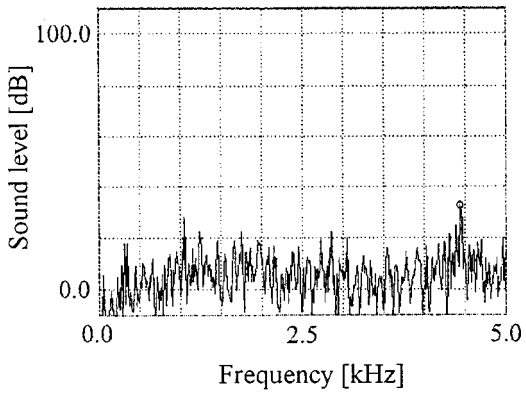


FIG. 6. Frequency spectrum of A-weighted sound pressure level in the sound isolation box.

(7) The rotation of the motor often diverges. Hence, when the calculated pulse width is longer than the maximum value A (s), a shorter pulse width t_0 (s) is chosen and the operation is restarted from (2).
(8) The voltage of the control signal is positive when t_{i+1} (s) is positive, and negative when t_{i+1} (s) is negative.
(9) The operation for t_{i+1} (s) ($i=1,2,...$) is continued in the same way.
(10) When the cost function J_{i+1} becomes less than the set value, the operation is finished.

IV. EXPERIMENTAL SYSTEM

Noise from the silencer was measured in a noise isolation box made of wood of thickness 10 mm, with dimensions 1020 mm×620 mm×720 mm. Sponge rubbers of 200 mm thickness were fixed to the inside surface of the box. The frequency spectrum of the sound pressure level in the box is shown in Fig. 6. The A-weighted sound pressure level in the box is 30 dB, and hence, sound pressure greater than 40 dB(A) can be measured in the box without correction.

The experimental setup is shown in Fig. 7. The microphone of the sound meter was installed near the outlet (10 cm from the central axis, and 10 cm from the left-hand end of the silencer). In this system, the FFT is used, and the cost function is taken to be a peak value of the A-weighted sound pressure level. Hence, the effect of wind from the dryer on the cost function is small because the pressure due to the wind is almost constant, and does not affect the peak frequency. This implies that the microphone can be installed inside the duct, enabling it to be isolated from external noise sources in practical use. This is one of the merits of our control method when attenuating noise from sources with constant wind pressure, such as a blower.

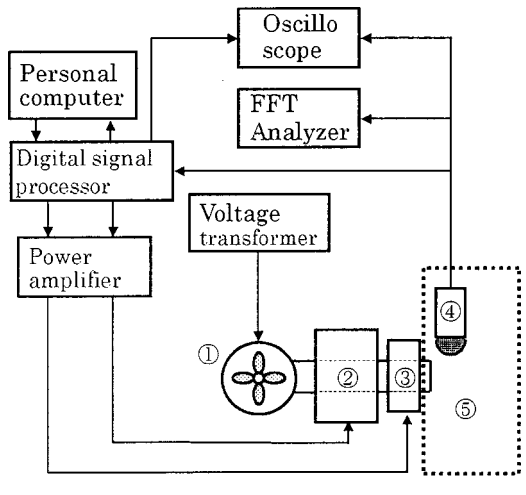


FIG. 7. Experimental setup.

The noise meter includes a well-known circuit for transforming the measured sound pressure into an A-weighted signal based on the Japan Industrial Standard (JIS C 1502). The A-weighted signal was processed by the DSP, which calculated the control signal using the algorithm described above. The control signal was input to the control motors via a power amplifier. The A-weighted signal from the microphone was also input to the FFT analyzer.

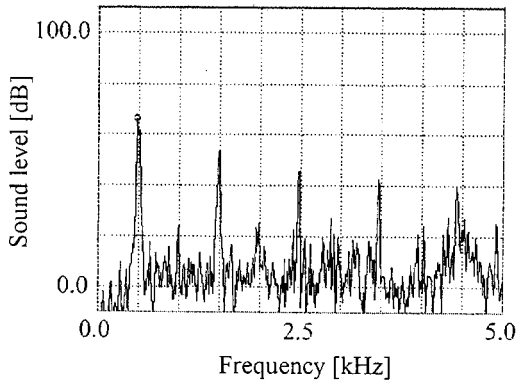
V. EXPERIMENTAL RESULTS

A. Noise control for triangular wave noise

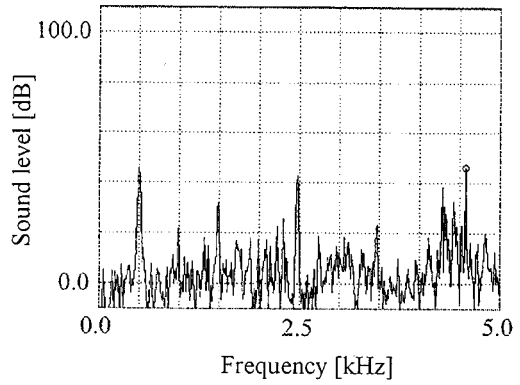
Noise with high-frequency components was investigated in order to verify the effectiveness of this control system. A function generator was used to create a noise signal, which was amplified and emitted from a speaker. The speaker was attached to the inlet of the silencer (right-hand end of the silencer in Fig. 1), into the cavity of inner cylinder A (see Fig. 2). The frequency spectrum of the A-weighted sound pressure level without the silencer fitted is shown in Fig. 8(a), and with the silencer in Fig. 8(b). It can be seen that the suppression of noise by the proposed silencer is about 25 dB(A) for the first peak and the third peak (the second significant peak). In previous devices,[7–9] only a single low-frequency sine wave was suppressed. This silencer is able to attenuate two significant peaks due to its two-stage design, and is applicable to high frequencies. This is one of the advantages of our silencer.

B. Control of blower noise

The blower used in this experiment is a hair dryer. The power source for the dryer is mains supply varied by a voltage transformer. The velocities of wind from the blower in this experiment were 6.6 m/s at 80 V, 8.0 m/s at 100 V, and 9.2 m/s at 130 V. As our method uses the FFT, the effect of wind is small. The constant pressure due to wind flow does not affect the peaks in the frequency domain obtained by the FFT. Moreover, the frequency of the peaks changes over time due to pressure fluctuations and other external effects. The dryer was installed at the inlet of the silencer (right-hand end of the silencer in Fig. 1). The sound pressure level spectrum, in the frequency domain, of noise generated by the dryer is shown in Fig. 9, without control. The results of two tests at the same voltage (130 V) are shown in the upper and lower figures. There is a frequency difference between the peaks in the upper and lower figures. The significant peaks are in the range 100–2000 Hz, and the variation of the peak

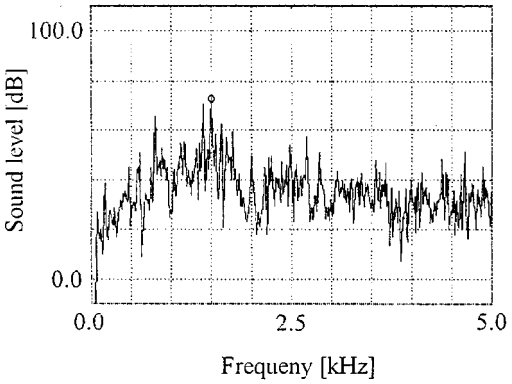


(a) Sound pressure without control

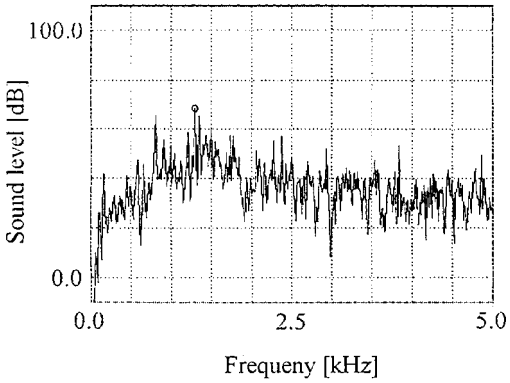


(b) Sound pressure with control

FIG. 8. Frequency spectrum of the A-weighted sound pressure level for the triangular wave noise (frequency=500 Hz).



(a) Sound pressure without control



(b) Sound pressure without control

FIG. 9. Frequency spectrum of the A-weighted sound pressure level for the blower noise [driving voltage=130(V) for the blower].

frequencies is about 200 Hz. It is difficult to use a Helmholtz-type silencer when there are frequency variations like this, however, our control method accounts for such variation.

The A-weighted sound pressure levels for three voltages of 80, 100, and 130 V, with the silencer fitted, are shown in Fig. 10. It can be seen that the noise is attenuated by about 7 dB(A) in every case. Time response of signals from the noise meter in V (signal without translation in dB), with and without the silencer fitted, are shown in Fig. 11, represented by solid and dashed lines, respectively. The amplitude of the signal is reduced by 70% with the silencer fitted. The amplitude reduces with time as a result of the adaptive control, requiring about 2 min for sufficient convergence. Although it takes some time to tune the silencer, as a result of using the FFT, control is not required after the silencer is tuned.

The sound pressure levels in the frequency domain are shown in Figs. 12 and 13 for two voltages. The rotation speed is 7400 rpm at 80 V. Two significant peaks occur at about 600 and 1200 Hz. Subharmonic peaks are also observed below the fundamental frequency. The dryer has five blades, and hence, the fundamental frequency due to blade rotation is about 600 Hz, corresponding to the fundamental frequency of 615 Hz in Fig. 12(a). The two most important peaks are attenuated and the spectrum flattens under silencer control, as shown in Fig. 12(b). The degree of attenuation is about 30 dB(A) at 600 Hz and 10 dB(A) at 1200 Hz. The first stage of the resonator in the proposed silencer was designed to have a resonant frequency between 10 and 1000 Hz, and that of the second stage to be between 1000 and 3000 Hz. The first peak is suppressed by the first resonator stage, and the second peak is suppressed by the second resonator stage. Frequencies over 3 kHz are not suppressed as such frequencies are outside the design specifications, however, the addition of a third stage would extend control to this range. At 130 V, the rotation speed is 9500 rpm and the fundamental frequency is 790 Hz, which corresponds to a fundamental frequency of 800 Hz, as shown in Fig. 13(a).

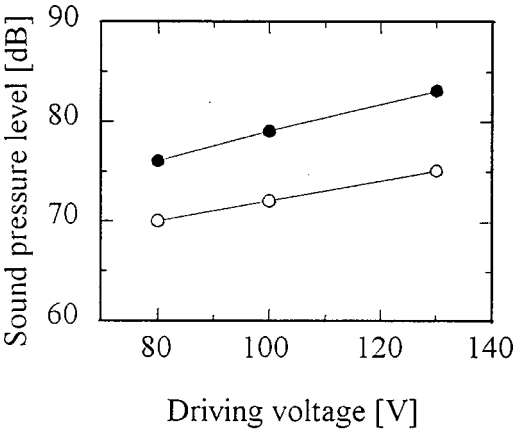


FIG. 10. A-weighted sound pressure level versus driving voltage for the blower noise (○: with the present control and ●: without control).

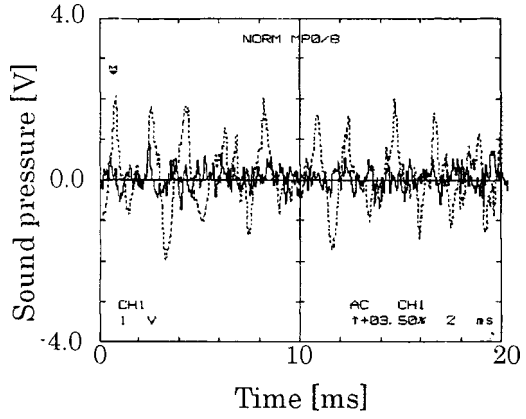


FIG. 11. Time response of the blower noise (—— with the present control, ---- without control) [driving voltage=130(V) for the blower].

Although the second significant frequency does not correspond to a harmonic of rotation speed [1400 Hz in Fig. 13(b)], the silencer still works well and the two significant peaks are adequately attenuated, as shown in Fig. 13(b). The degree of attenuation is about 10 dB(A) for the first peak, and 20 dB(A) for the second peak, however, the third peak, at over 1.5 kHz, is not attenuated.

The control voltage was 9 V and the control current for the motor is 0.075 A, thus 0.675 W per pulse in the control signal, for each motor, is required by this adaptive control method, significantly less than that required for active control.

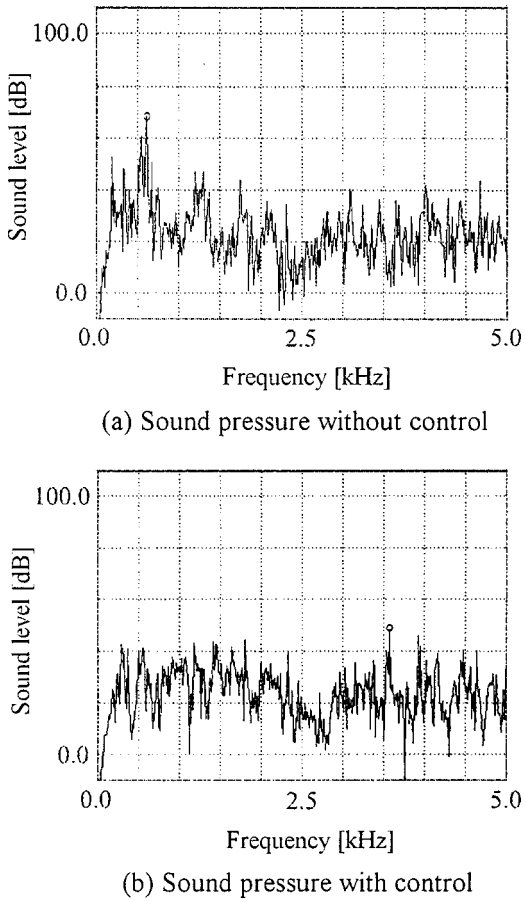


FIG. 12. Frequency spectrum of the A-weighted sound pressure level for the blower noise [driving voltage=80(V) for the blower].

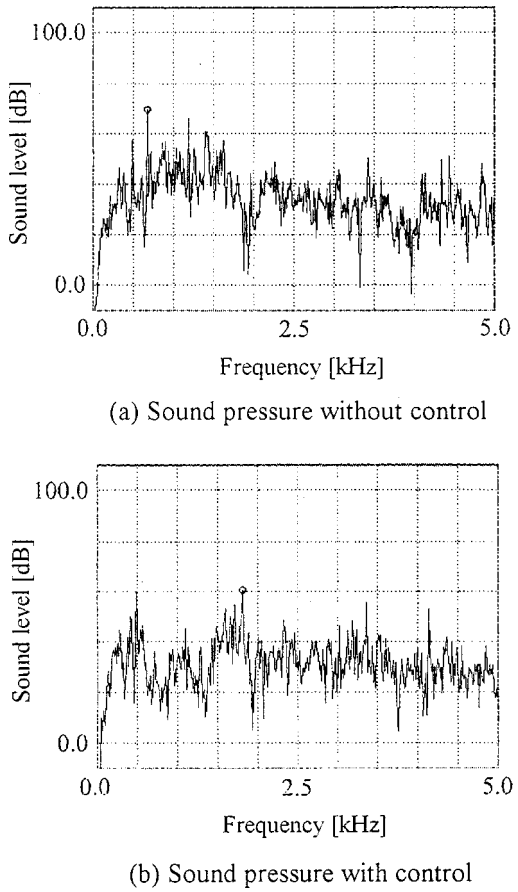


FIG. 13. Frequency spectrum of the A-weighted sound pressure level for the blower noise [driving voltage=130(V) for the blower].

VI. CONCLUSION

A silencer consisting of a two-stage Helmholtz resonator with auto-tuning control was presented.

(1) An algorithm using the FFT was presented for controlling the silencer. In order to validate the system and the control method, experimental tests were carried out.
(2) Experimental results were obtained for triangular wave noise. It was ascertained that the silencer is effective for attenuating noise with high-frequency components.
(3) The frequencies of peaks in blower noise vary over time. Hence, it is difficult to control noises with tuning control. A control method that deals with this variation was presented, and it was ascertained that the silencer is effective in attenuating such noise.

APPENDIX: SELF-TUNING SINGLE-STAGE SILENCER WITH CONVENTIONAL CONTROL

A single-stage self-tuning silencer was discussed in Refs. 9–13. In order to demonstrate the advantages of our silencer, a conventional silencer is discussed. A single-stage silencer consists of outer, middle, and inner cylinders, similar to the proposed cylinder. The outer and middle cylinders are connected by a plate at both ends, with a cavity between

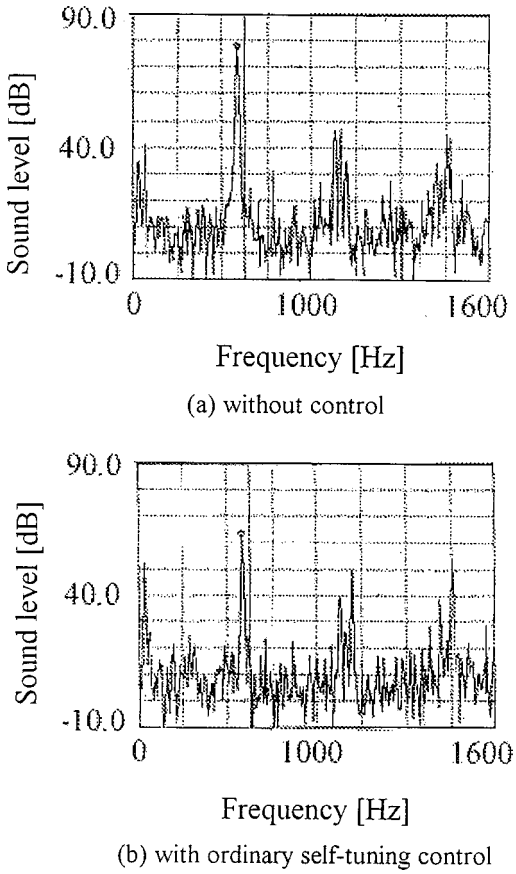


FIG. A1. Frequency spectrum of the A-weighted sound pressure level for triangular wave noise under conventional tuning control (frequency $=500$ Hz).

them. There is no gap between the middle and inner cylinders, however, they are not connected, and both have holes at regular intervals. The inner cylinder is driven in the axial direction by a ball screw and motor. Hence, the area of the holes of the middle cylinder varies with the movement of the inner cylinder. Sound from a speaker is directed into one end of the silencer.

The control algorithm here is a little bit different from the conventional control. In the conventional control, the external frequency is detected, and the resonance frequency of the silencer is tuned. This needs two microphones. In order to make the system compact, this study uses the cost function minimization method, which requires only one microphone. The effect will be the same as that of conventional methods. Hence, a microphone of a noise meter is installed at the other end of the silencer, and the following cost function is applied to the output signal:

$$J=\int_0^T f(t)^2 dt, \tag{A1}$$

where $f(t)$ is the signal from the noise meter, and T is time interval under consideration. Adaptive control is applied to the motor of the silencer in such a way as to minimize the cost function J in Eq. (A1). The inner cylinder continues to move until the sound pressure of the target frequency is minimized.

The sound pressure level spectrum for triangular wave noise is shown in Fig. A1 for the frequency domain. The peak corresponding to the principal frequency of the noise is attenuated with control [Fig. A1(b)], compared to that without control [Fig. A1(a)]. However, modes higher than the second are not attenuated. Hence, this method is only applicable to the principal mode. In particular, when the frequency of the target peaks varies, this form of control does not converge. Hence, other mechanisms and control methods are required for attenuating noise with high-frequency components and variations.

[1] J. C. Shipps, S. K. Miller, and J. Shaver, ''Active exhaust silencers quiet diesel generator sets,'' J. Power Div. (Am. Soc. Civ. Eng.) **98**, 42–45 (1994).
[2] K. Nagayasu and S. Suzuki, ''Active control application to noise reduction using adaptive signal processing,'' J. Inst. Electron., Inf. Commun. Eng. **J75B-II**, 862–870 (1992).
[3] S. Suzuki, ''Active noise control system for refrigerator noise,'' Noise Control, **15**, 292–295 (1991).
[4] J. N. Denenberg, S. C. Miller, and J. Mark, ''Active compressor engine silencer reduces exhaust noise,'' Pipe Line Ind. **77**, 64–67 (1994).
[5] T. Kelsall, T. Gerritsen, and T. Landon, ''Venturis as silencers in a BOF Stack,'' Iron Steel Eng. **72**, 37–39 (1995).
[6] M. E. Schuchardt, T. A. Dear, and K. Ingard, ''Four-cylinder air induction,'' Automot. Eng. **102**, 105–108 (1994).
[7] R. Baosheng, H. Matsuhisa, T. Abe, and S. Sato, ''Semiactive control of duct noise by volume-variable resonator,'' J. J. Soc. Mech. Eng. C. **56-531**, 2957–2962 (1990).
[8] T. Izumi and T. Narikiyo, ''Muffler system controlling the volume of resonant cavity,'' J. Acoust. Soc. Jpn. **46-6**, 479–485 (1990).
[9] T. Izumi, H. Takami, and T. Narikiyo, ''Muffler system controlling an aperture neck of a resonator,'' J. Acoust. Soc. Jpn. **47-9**, 647–652 (1991).
[10] C. R. Cheng, J. D. McIntosh, M. T. Zuroski, and J. Eriksson, ''Tunable acoustic system,'' U.S. Patent No. 5,930,371 (1999).
[11] A. von Flotow, A. Beard, and D. Bailey, ''Adaptive tuned vibration absorbers: Tuning laws, tracking, agility, sizing, and physical implementations,'' Noise-Con 94, Ft. Lauderdale, FL (1994), pp. 437–454 (unpublished).
[12] R. J. Bernhard, H. R. Hall, and D. Jones, ''Adaptive-passive noise control,'' Inter-Noise, Toronto (1992), pp. 427–430 (unpublished).
[13] J. M. Bedout, M. A. Franchek, R. J. Bernard, and L. Mongeau, ''Adaptive-passive noise control with self-tuning Resonators,'' J. Sound Vib. **202**, 109–123 (1997).

Zones of quiet in a broadband diffuse sound field

Boaz Rafaely[a)]
Institute of Sound and Vibration Research, University of Southampton, Southampton SO17 1BJ, United Kingdom

(Received 22 January 2001; accepted for publication 10 April 2001)

The zones of quiet in pure-tone diffuse sound fields have been studied extensively in the past, both theoretically and experimentally, with the well-known result of the 10-dB attenuation extending to about a tenth of a wavelength. Recent results on the spatial-temporal correlation of broadband diffuse sound fields are used in this study to develop a theoretical framework for predicting the extension of the zones of quiet in broadband diffuse sound fields. This can be used to study the acoustic limitations imposed on local active sound control systems such as an active headrest when controlling broadband noise. Spatial-temporal correlation is first revised, after which derivations of the diffuse field zones of quiet in the near-field and the far-field of the secondary source are presented. The theoretical analysis is supported by simulation examples comparing the zones of quiet for diffuse fields excited by tonal and broadband signals. It is shown that as a first approximation the zone of quiet of a low-pass filtered noise is comparable to that of a pure-tone with a frequency equal to the center frequency of the broadband noise bandwidth. [DOI: 10.1121/1.1377632]

PACS numbers: 43.50.Ki, 43.55.Cs [MRS]

I. INTRODUCTION

Active control of sound has been studied intensively in the past two decades, both theoretically and experimentally.[1,2] Global control of sound in enclosures was shown to be limited to the very low frequencies, where only few acoustic modes dominate the sound field,[1] and so in many cases active control is practical only locally, generating limited zones of quiet. A typical application for local active sound control is a noise-reducing headrest in a passenger seat, attenuating noise around the passenger's ears.[3–8] Since local control would usually be performed in enclosures, a model which was often used is that of diffuse primary sound field, and a decaying near field to model the secondary pressure from a closely located source. Pure-tone sound fields have been studied extensively for such local control, developing theoretical limits on the spatial extension of the zone of quiet,[9,10] and verifying the results with experiments.[8] It was shown that the 10-dB zone of quiet is extended to about one-tenth of a wavelength for pure-tone sound fields.[11] The analysis of zones of quiet in diffuse fields used the well-known spatial correlation function of pure-tone diffuse fields,[9,10] derived by Cook *et al.* in the 1950s.[12]

Although the theoretical and experimental results for pure-tone local control were useful to predict the performance of active headrest attenuating low-frequency tonal noise in propeller aircraft, for example, in many cases the nature of the noise is broadband, such as in most jet passenger aircraft, and so pure-tone results will be of limited use in this case. For a broadband local active control system a useful measure of performance would be the spatial extent of the overall sound attenuation, which requires the analysis of broadband sound fields and so cannot make use of the pure-tone results. Previous studies of broadband local control systems were performed experimentally, by including, for example, the effect of the feedback control system, which would usually be used in this case.[4,5] It was shown that broadband local active control could be useful in practice, although in addition to the limitations imposed by the acoustics, other limitations are also imposed by the control system, due to, for example, the delay in the response between the loudspeaker and the cancellation point.

The aim of this article is to develop a theoretical framework for predicting the spatial extent of the zones of quiet in broadband diffuse sound fields. This can then be used to predict performance limitations of broadband active headrest systems as imposed by the acoustics, and can complement previous experimental results. Similar to the pure-tone zones of quiet case, the analysis of broadband zones of quiet presented here employ spatial correlation of diffuse sound fields. However, in this work *broadband* spatial-temporal correlation is used, as developed recently by Rafaely.[13]

The article is organized as follows. First the diffuse sound field and spatial correlation are introduced, after which theoretical results for broadband local control are developed, for near-field control, but also for far-field control, where the secondary source is located away from the cancellation point. Finally simulation results for various broadband sound fields are presented and compared to the pure-tone case. It is shown that as a first approximation, broadband zone of quiet can be predicted from that of tones at the mid-frequency of the broadband noise bandwidth.

II. THE DIFFUSE SOUND FIELD

The plane wave model of a diffuse sound field assumes an infinite number of plane waves, arriving uniformly from all directions, with random phases.[14,15] Although a perfect diffuse field rarely exists, the model is widely used for reverberant sound field analysis, where the field is assumed to

a)Electronic mail: br@isvr.soton.ac.uk

 0001-4966/2001/110(1)/296/7/$18.00

be sufficiently diffuse. A commonly used definition for sufficiently diffuse field is that by Schroeder,[16] which defines the field being diffuse above the Schroeder frequency. This corresponds to the frequency above which there exists at least three room modes within the 3-dB bandwidth of any one mode.[14] This complements the wave model of a diffuse field if it is assumed that each mode can be represented by eight plane waves,[17] and so a large number of significant modes implies a large number of plane waves, which in the limit approaches the definition of a perfect diffuse field. The pressure in a perfect diffuse field can therefore be written as a function of space and time in spherical coordinates $\mathbf{r}=(r,\theta,\phi)$ as[14]

$$p(\mathbf{r},t)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}p_n(\mathbf{r},t), \tag{1}$$

where $p(\mathbf{r},t)$ is the total pressure at position $\mathbf{r}$ and time t, N is the number of plane waves which approaches infinity, and p_n is the nth plane wave. The spatial correlation in the pure-tone diffuse field was studied both theoretically and experimentally by Cook *et al.*,[12] who showed that it behaves as a sinc function,

$$\rho(\Delta\mathbf{r})=\mathrm{sinc}\,(k\Delta\mathbf{r})=\frac{\sin(k\Delta\mathbf{r})}{k\Delta\mathbf{r}}, \tag{2}$$

where k denotes the wave number and ρ is the correlation coefficient, which can be defined, assuming the sound field is stationary over both space and time,[1] as

$$\rho(\Delta\mathbf{r},\Delta t)=\frac{E[p(\mathbf{r}_1,t_1)p(\mathbf{r}_0,t_0)]}{E[p^2]}, \tag{3}$$

where $\Delta\mathbf{r}$ denotes the distance between the two points, $\Delta\mathbf{r}=|\mathbf{r}_1-\mathbf{r}_0|$, Δt denotes the time lag given by $\Delta t=t_1-t_0$, $E[\cdot]$ denotes the expectation operation which is calculated as the average over many samples of diffuse sound fields, and $E[p^2]$ is the variance of the pressure which is not dependent on $\mathbf{r}$ or t due to the stationarity assumption. As discussed earlier, (2) was widely used in the theoretical analysis of zones of quiet in pure-tone diffuse sound fields. Rafaely[13] recently developed an expression for the correlation which can incorporate both pure-tone and broadband sound fields, and which depends on the power spectral density of the signal exciting the diffuse field,

$$\rho(\Delta\mathbf{r},\Delta t)=\frac{1}{2\pi E[p^2]}\int_{-\infty}^{\infty}S(\omega)\,\mathrm{sinc}\left(\frac{\omega\Delta\mathbf{r}}{c}\right)e^{j\omega\Delta t}\,d\omega, \tag{4}$$

where $E[p^2]$ is equal to the integral over $S(\omega)$, i.e., the signal power. Equation (4) enables the extension of the pure-tone local control results to broadband sound fields, as shown in the following sections.

III. NEAR-FIELD BROADBAND ACTIVE SOUND CONTROL

Local active sound control in a diffuse sound field can be achieved by introducing a secondary source and cancelling the total pressure in the near field of the source. A simple model used to theoretically study such an approach is that of a monopole secondary source in a primary diffuse sound field. This arrangement was used by Joseph *et al.*[10] to study zones of quiet in pure-tone diffuse fields, and provided a useful insight into the performance of more practical near-field active sound control systems such as an active headrest system. A derivation of the spatial extension of the zones of quiet in a *broadband* diffuse sound field for local active control is presented in this section. This is a novel result which can be used to predict the spatial extent of the overall attenuation of the broadband noise in diffuse sound fields.

Consider a secondary monopole source placed at the origin of a spherical coordinate system, $\mathbf{r}=(r,\theta,\phi)$, with the resulting pressure denoted by $p_s(\mathbf{r},t)$. The primary sound field is diffuse and is denoted by $p_p(\mathbf{r},t)$. The total pressure is a superposition of the primary and secondary pressure contributions and is given by

$$p(\mathbf{r},t)=p_p(\mathbf{r},t)+p_s(\mathbf{r},t). \tag{5}$$

The pressure at position $\mathbf{r}_0=(r_0,\theta_0,\phi_0)$ is cancelled, i.e., $\mathbf{r}_0$ is assumed to be the cancellation point, such that

$$p_p(\mathbf{r}_0,t)+p_s(\mathbf{r}_0,t)=0. \tag{6}$$

It is now assumed that position $\mathbf{r}_0$ is in the near field of the secondary source, such that the indirect secondary sound field resulting from reflections is negligible. The distance from the source at which the direct field dominates is referred to as the ''reverberation distance,'' which depends on the room volume and reverberation time.[16] The spatial extent of the zone of quiet depends on how well the primary pressure is attenuated around the cancellation point. The averaged squared total pressure at position $\mathbf{r}_1=(r_1,\theta_1,\phi_1)$ near the cancellation point is therefore calculated, where the expectation operation $E[\cdot]$ is used as a statistical average over many samples of diffuse sound fields. The variance of the total pressure at position $\mathbf{r}_1$ can therefore be written using (5) as

$$E[p^2(\mathbf{r}_1,t)]=E[p_p^2(\mathbf{r}_1,t)^2]+E[p_s^2(\mathbf{r}_1,t)]+2E[p_p(\mathbf{r}_1,t)p_s(\mathbf{r}_1,t)]. \tag{7}$$

Note that the variance of the total pressure at $\mathbf{r}_1$ depends on the variance of the primary and secondary fields at the same point, but also on the correlation between the primary and secondary fields at $\mathbf{r}_1$. Since we assumed in (6) that both fields are equal with opposite phase at the cancellation point $\mathbf{r}_0$, this correlation will depend on how both fields change from $\mathbf{r}_0$ to $\mathbf{r}_1$, which will be developed later. We next expand each of the terms in Eq. (7), and reformulate the equation.

Assuming the diffuse primary sound field is stationary, such that the variance of the pressure is the same for all $\mathbf{r}$ and t, the following equality can be written:

$$E[p_p^2(\mathbf{r}_1,t)]=E[p_p^2(\mathbf{r}_0,t)]=E[p_p^2]. \tag{8}$$

It is now assumed that the secondary source is generated by a monopole point source. Although the monopole source is not an accurate representation of more practical secondary sources such as loudspeakers, under some assumptions the pressure produced by a monopole behaves in a similar way

to that produced by a piston in a baffle, which is often used to model sound radiation from loudspeakers. These assumptions are[18] (1) $ka<0.5$, or $a<\lambda/4\pi$, which means that the source radius a is much smaller than a wavelength, and the source can therefore be considered omni-directional, and (2) $r>a$, which suggests that only pressure further away than one source radius is considered. For example, these assumption will hold for a 4-in. ($a=5$ cm) loudspeaker, for frequencies below about 500 Hz, further than 5 cm from loudspeaker. These are reasonable assumptions considering a practical local active control system such as active headrest,[4,5] and so the monopole model should provide useful insight into the behavior of more practical systems.

The secondary sound field produced by a monopole point source in the near field is assumed to generate spherical waves, which propogate away from the source and decay in amplitude:[17,19]

$$p_s(r,t)=\frac{\rho_0}{4\pi r}\dot{q}\left(t-\frac{r}{c}\right), \tag{9}$$

which is now dependent only on the distance from the source, r, with q denoting the source strength (volume velocity per unit volume) and $\dot{q}$ its derivative with respect to time. The secondary pressure at $\mathbf{r}_1$ can now be written in terms of the secondary pressure at $\mathbf{r}_0$ using (9) as

$$p_s(\mathbf{r}_1,t)=\frac{r_0}{r_1}p_s\left(\mathbf{r}_0,t-\frac{\Delta r}{c}\right), \tag{10}$$

where $\Delta r=r_1-r_0$, which is the difference in the distances of the two points $\mathbf{r}_1$ and $\mathbf{r}_0$ to the source.

The averaged squared secondary pressure at $\mathbf{r}_1$ can now be written using (10), (6) and (8) as

$$\begin{aligned}E[p_s^2(\mathbf{r}_1,t)]&=\left(\frac{r_0}{r_1}\right)^2E\left[p_s^2\left(\mathbf{r}_0,t-\frac{\Delta r}{c}\right)\right]\\&=\left(\frac{r_0}{r_1}\right)^2E\left[p_p^2\left(\mathbf{r}_0,t-\frac{\Delta r}{c}\right)\right]\\&=\left(\frac{r_0}{r_1}\right)^2E[p_p^2].\end{aligned} \tag{11}$$

The last term in (7) can also be written using (10), (6), and (3) as

$$\begin{aligned}E[p_p(\mathbf{r}_1,t)p_s(\mathbf{r}_1,t)]&=E\left[p_p(\mathbf{r}_1,t)\frac{r_0}{r_1}p_s\left(\mathbf{r}_0,t-\frac{\Delta r}{c}\right)\right]\\&=-\frac{r_0}{r_1}E\left[p_p(\mathbf{r}_1,t)p_p\left(\mathbf{r}_0,t-\frac{\Delta r}{c}\right)\right]\\&=-\frac{r_0}{r_1}\rho\left(\Delta\mathbf{r},\frac{\Delta r}{c}\right)E[p_p^2].\end{aligned} \tag{12}$$

The variance of the total pressure at position $\mathbf{r}_1$ in (7) can now be written in terms of the variance of the primary pressure by substituting Eqs. (8), (11), and (12) in Eq. (7),

$$E[p^2(\mathbf{r}_1,t)]=E[p_p^2]+\left(\frac{r_0}{r_1}\right)^2E[p_p^2]-2\frac{r_0}{r_1}\rho\left(\Delta\mathbf{r},\frac{\Delta r}{c}\right)E[p_p^2]. \tag{13}$$

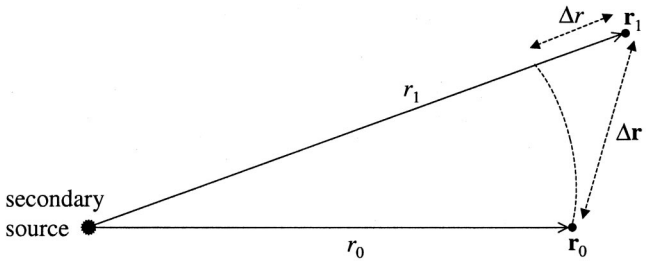


FIG. 1. Graphical representation of the cancellation point, $\mathbf{r}_0$, and a position near the cancellation point, $\mathbf{r}_1$, relative to the secondary source. The distances $\Delta\mathbf{r}$ and Δr are also illustrated in the figure.

Dividing (13) by the variance of the primary pressure, an expression for the sound attenuation ϵ at $\mathbf{r}_1$ assuming cancellation at $\mathbf{r}_0$ is derived as follows:

$$\epsilon(\mathbf{r}_1,\mathbf{r}_0)=\frac{E[p^2(\mathbf{r}_1,t)]}{E[p_p^2]}=1+\left(\frac{r_0}{r_1}\right)^2-2\frac{r_0}{r_1}\rho\left(\Delta\mathbf{r},\frac{\Delta r}{c}\right), \tag{14}$$

where the sound attenuation in dB is given by $10\log_{10}\epsilon$. It is important to note that in (14) $\Delta\mathbf{r}=|\mathbf{r}_1-\mathbf{r}_0|$ is the distance from position $\mathbf{r}_1$ to the cancellation point $\mathbf{r}_0$, while $\Delta r=r_1-r_0$ is the difference between the distances of the two points $\mathbf{r}_1$ and $\mathbf{r}_0$ to the secondary source, as illustrated in Fig. 1. In the simplified case of on-axis attenuation, the two distances are equal, i.e., $\Delta\mathbf{r}=\Delta r$. Equation (14) together with the expression for the spatial-temporal correlation function in a diffuse field [Eq. (4)] can be used to calculate the attenuation of broadband noise in the near field of a monopole point source, a distance $\Delta\mathbf{r}$ away from the cancellation point. Examples of near-field zones of quiet in a broadband diffuse sound field are presented later.

IV. FAR-FIELD BROADBAND ACTIVE SOUND CONTROL

The previous section described active sound control in a diffuse field where the secondary source was placed close to the cancellation point. The latter was therefore in the near field of the secondary source, and the derivation that followed employed this assumption. In this section it is assumed that the cancellation point is far from the secondary source, such that the resulting secondary field at the cancellation point is assumed to be diffuse. In practice this means that the cancellation point is further than a ''reverberation distance'' or ''radius of reverberation'' [14] away from the secondary source. In this case both the primary and the secondary sound fields are diffuse. Nevertheless, the two diffuse fields are assumed to be uncorrelated, which is achieved in practice if the primary and secondary sources are positioned sufficiently far away from each other (more than a wavelength away for pure-tone fields[1]).

Unlike the case of near-field sound control which provides an insight into the performance of practical active sound control systems, such as an active headrest, broadband sound control using a secondary source in the far field is less practical. This is because a practical feedforward control system will require a good reference of the noise signal in advance, which is rarely available for broadband noise, e.g., jet

turbulence noise, while a feedback control system will have poor performance due to the long delay from the secondary source to the cancellation point. It is important to note that such a limitation is not applicable to pure-tone sound fields where system delay does not affect performance. In addition, placing the secondary source far from the cancellation point could result in large increase in the pressure at other locations in the enclosure,[1] which is an undesirable side effect. Although of less practical relevance, the derivation of far-field broadband active sound control is presented here for theoretical completeness.

Joseph[20] derived an equation for the average mean square pressure away from the cancellation point under similar conditions but when a pure-tone sound field was assumed,[1]

$$E[p^2(\mathbf{r}_1)]=(E[p_p^2]+E[p_s^2])(1-\rho^2(\Delta\mathbf{r})). \tag{15}$$

The sound attenuation can now be derived by dividing (15) with the variance of the primary pressure

$$\epsilon(\Delta\mathbf{r})=\frac{E[p^2(\mathbf{r}_1)]}{E[p_p^2]}=\left(1+\frac{E[p_s^2]}{E[p_p^2]}\right)(1-\rho^2(\Delta\mathbf{r})). \tag{16}$$

Elliott *et al.*[9] noted that the value of $E[p_s^2]/E[p_p^2]$ can only be defined in statistical terms, and does not have a finite mean value. In practice, however, the secondary source strength will be limited, and in an example simulation,[9] a value of $E[p_s^2]$ was used which is three time larger than $E[p_p^2]$, and so for this example (16) can be written as

$$\epsilon(\Delta\mathbf{r})=4(1-\rho^2(\Delta\mathbf{r})). \tag{17}$$

Equation (4) can now be used in (17) to compute the attenuation or the extent of the far-field zones of quiet for a broadband sound field. Examples of such zones of quiet are presented below.

V. EXAMPLES OF NEAR-FIELD ZONES OF QUIET

Examples of near-field zones of quiet calculated using the results derived earlier are presented in this section. The primary field is assumed to be diffuse while the secondary field is excited by a monopole point source. The cancellation point where both fields are equal but opposite in phase is located in the near field of the monopole source. A pure-tone diffuse sound field, which has been well studied previously, is compared to broadband diffuse sound fields using the results derived in this work. The diffuse sound fields in the examples presented here are excited by the signals as described in Table I.

Figure 2 shows the power spectral density of the signals used in the simulation examples as described in Table I. The spatial correlation of the various primary diffuse sound fields are compared next, after which the correlation functions between the primary and secondary sound fields away from the cancellation point are evaluated, which then leads to a comparison of the zones of quiet. The spatial-temporal correlation function for the pressure in a diffuse sound field is calculated in MATLAB using (4) by generating the appropriate signals, sampled at $F_s=2$ kHz, with discrete power spectral densities calculated using the discrete Fourier transform (DFT) having $M=4096$ points. The integral in (4) was approximated by a summation over frequency, as follows:

$$\rho(\Delta\mathbf{r},\Delta t)\approx\frac{1}{\sum_{m=0}^{M-1}S(m)}\sum_{m=0}^{M-1}S(m)\,\mathrm{sinc}\left(\frac{2\pi mF_s}{M}\frac{\Delta\mathbf{r}}{c}\right)\times e^{j(2\pi mF_s/M)\Delta t}. \tag{18}$$

TABLE I. Description of the signals used in the simulation examples.

Signal	Description
300 Hz tone	A 300-Hz pure tone
300 Hz LPF	Broadband signal generated by passing white noise through a 32nd-order Butterworth low-pass filter with a cutoff frequency of 300 Hz
600 Hz LPF	Broadband signal generated by passing white noise through a 32nd-order Butterworth low-pass filter with a cutoff frequency of 600 Hz
BPF	Broadband signal generated by passing white noise through an eighth-order Butterworth low-pass filter with a cutoff frequency of 400 Hz, and another second-order Butterworth high-pass filter with a cutoff frequency of 600 Hz, as used by Rafaely *et al.*[4] to analyze the performance of a laboratory active headrest system

Figure 3 shows the spatial correlation of the primary diffuse field $\Delta\mathbf{r}$ away from the cancellation point, evaluated using (18) as $\rho(\Delta\mathbf{r},0)$, for the sound fields described in Table I. The figure shows that the spatial correlation for the 300-Hz pure-tone sound field behaves as a spatial sinc function, as expected,[12] with the 600-Hz lowpass filtered noise having similar correlation for small $\Delta\mathbf{r}$. This observation that the spatial correlation for a band of frequencies can be approximated by that of a pure tone at the center frequency has been previously observed.[12,21] The 300-Hz low-pass filtered noise has higher spatial correlation, as expected, since it is composed of lower frequencies. It is also interesting to note that the sound field composed of the band-pass filtered noise has a similar spatial correlation to the 600-Hz low-pass filtered noise, for small $\Delta\mathbf{r}$, since it has a similar bandwidth.

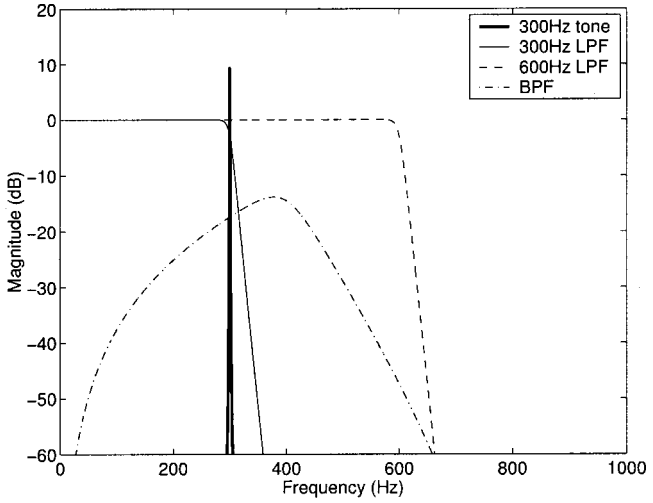


FIG. 2. Power spectral density of the signals used in the simulations as described in Table I.

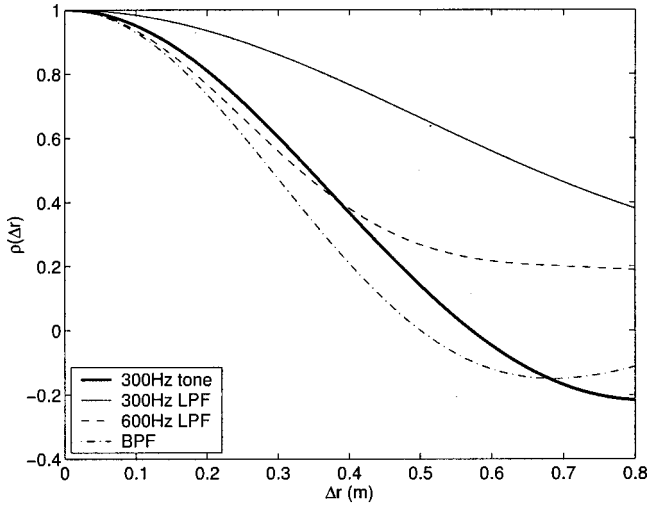


FIG. 3. The spatial correlation of the primary pressure in a diffuse sound field for the excitation signals described in Table I.

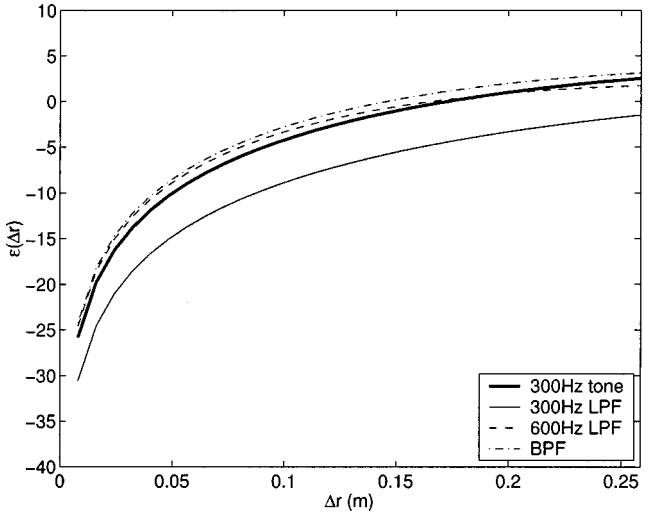


FIG. 5. Attenuation as a function of the distance from the cancellation point $\Delta\mathbf{r}$ for the signals described in Table I.

As shown in (7), the cross-correlation between the primary diffuse field and the secondary near field when evaluated at $\mathbf{r}_1$, i.e., $\Delta\mathbf{r}$ away from the cancellation point, is used in the calculation of the total pressure and then the attenuation at $\mathbf{r}_1$. This cross-correlation is evaluated here for the sound fields described in Table I, using (12) and (18), by substituting $\Delta\mathbf{r}$ and $\Delta t=\Delta r/c$ in the spatial-temporal correlation function of the primary diffuse field. Figure 4 shows this cross-correlation for the signals described in Table I, where it was assumed that the cancellation point is sufficiently far from the secondary source such that $\mathbf{r}_1\approx\mathbf{r}_0$ in order to present the limit of the correlation values. The figure shows that the correlation values are negative for small $\Delta\mathbf{r}$ since the primary and secondary fields are equal but with opposite phase at $\mathbf{r}_0$. Also, comparing the results to Fig. 3, it is clear that the cross-correlation between the primary and the secondary sound fields at $\Delta\mathbf{r}$ away from the cancellation point is smaller than the auto-correlation of the primary diffuse field for a spacing of $\Delta\mathbf{r}$. This can be explained by the fact that when moving from position $\mathbf{r}_0$, where both fields are equal with opposite phase, to position $\mathbf{r}_1$, the primary field reduces correlation according to (4), while the secondary field also reduces correlation according to the near-field behavior described in (10). The total equivalent spacing between the two fields is therefore $2\Delta\mathbf{r}$, compared to only $\Delta\mathbf{r}$ in Fig. 3, resulting in a greater reduction in the cross-correlation compared to diffuse field auto-correlation.

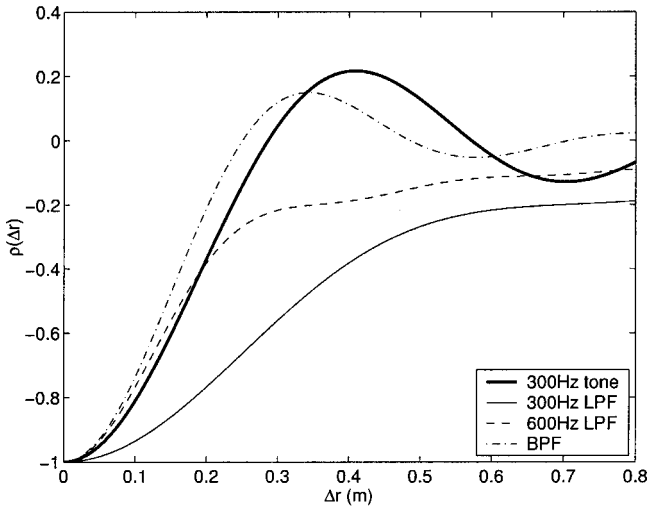


FIG. 4. The spatial cross-correlation function between the primary and the secondary pressures $\Delta\mathbf{r}$ away from the cancellation point, $-(r_0/r_1)\rho(\Delta\mathbf{r},\Delta r/c)$, assuming $r_0/r_1\approx 1$ for the excitation signals described in Table I.

The attenuation $\Delta\mathbf{r}$ away from the cancellation point can be calculated using (14) and (18). Figure 5 shows the calculated attenuation for the sound fields as described in Table I, as a function of the distance from the cancellation point $\Delta\mathbf{r}$. Again, it was assumed that $\mathbf{r}_1\approx\mathbf{r}_0$ to present the limits of the attenuation values. The figure shows that the 300-Hz pure-tone has similar zone of quiet to the 600-Hz low-pass filtered noise and the band-pass filtered noise, whereas the 300-Hz low-pass filtered noise shows larger zones of quiet. It is important to note that the size of the zone of quiet for the 300-Hz pure-tone, defined by $2\Delta\mathbf{r}$ for $\epsilon=0.1$, i.e., 10-dB attenuation, is about 0.088λ, which is slightly smaller than the 0.1λ rule derived by Nelson and Elliott.[1] This is explained by the fact that in the derivation presented in Nelson and Elliott[1] the change in the secondary sound field around the cancellation point was approximated by a first-order function, with higher orders neglected, whereas in this work no such approximation was made.

The 10-dB zone of quiet is presented next, which is the attenuation contour with a 10-dB value. Figure 6 shows the calculated 10-dB attenuation contours, or two-dimensional zones of quiet for the sound fields described in Table I. In this case the monopole secondary source is located at the origin, while the cancellation point $\mathbf{r}_0$ is positioned at (0.2, 0), i.e., 20 cm away from the source. The attenuation as a function of position was calculated using (14) and (18), with only the 10-dB attenuation contour shown. The result for the 300-Hz tone is comparable with that of Garcia-Bonito and Elliott,[8] while again it is clear that the 300-Hz tone has a zone of quiet similar to the 600-Hz low-pass filtered noise and the band-pass filtered noise. These results suggest that

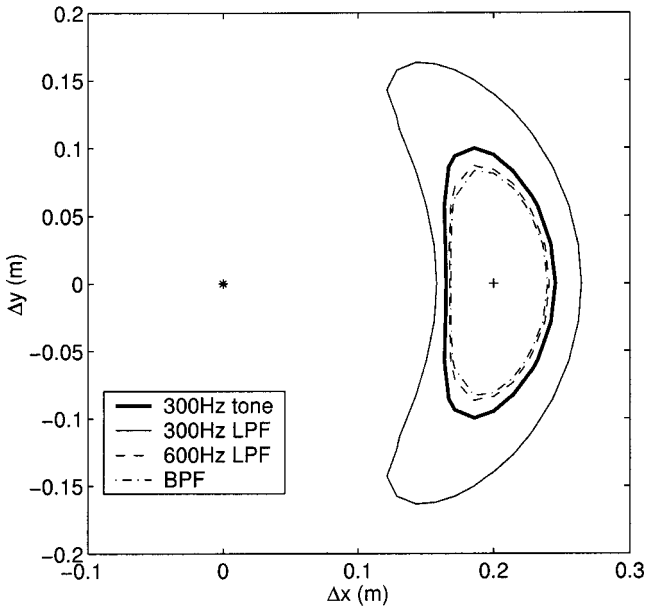


FIG. 6. The 10-dB attenuation contours as a function of Δx and Δy, for the signals described in Table I, with the secondary source denoted by ''*'' and the cancellation point denoted by ''+'' located at (0.2, 0).

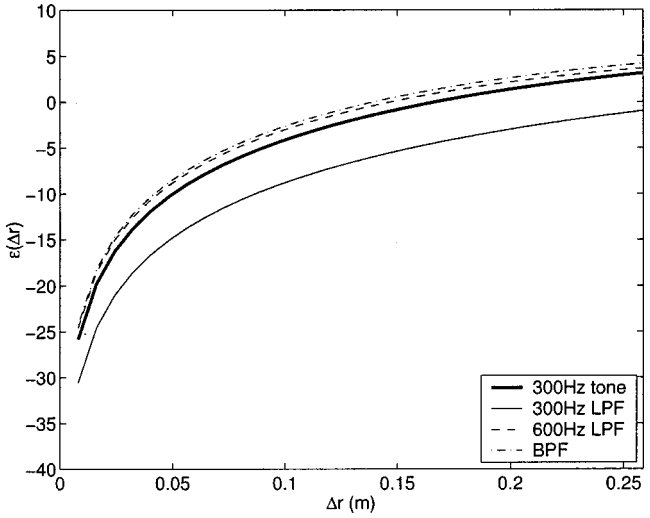


FIG. 7. Far-field attenuation as a function of $\Delta \mathbf{r}$ for the signals described in Table I.

the size of the zone of quiet for a broadband noise of a given bandwidth will be similar to that of a pure tone at the middle frequency range of the broadband noise. Nevertheless, the zone of quiet for a more general spectrum can be calculated more accurately as described earlier.

The 10-dB zone of quiet for the band-pass filtered noise is shown to be about 8 cm. This is slightly higher but comparable to the zone of quiet presented by Rafaely and Elliott,[4] for a laboratory active headrest system, which used a more realistic experiment including a loudspeaker as a source, a Manikin as a head, and feedback control to generate the secondary source signal.

VI. EXAMPLES OF FAR-FIELD ZONES OF QUIET

An example of far-field zones of quiet are presented in this section, where it is assumed that the secondary source is placed far from the cancellation point, such that both the primary field and the secondary field are diffuse. Similar signals as for the previous example were used here to excite the sound fields, which are described in Table I. Equations (17) and (18) with $\Delta t = 0$ were used to calculate the spatial correlation and then the attenuation for the diffuse sound fields in this example.

Figure 7 show the attenuation as a function of distance from the cancellation point for all four diffuse sound fields. Results are very similar to the near-field case, and here, as well, the zone of quiet for the broadband noise can be approximated by that of a tone at the middle frequency.

VII. CONCLUSIONS

The zones of quiet for broadband diffuse sound fields were derived theoretically and then demonstrated using simulation examples. Both near-fields zones of quiet, where the cancellation point is in the near-field of the secondary source, and far-field zones of quiet, where the cancellation point is in the far-field of the secondary source, were considered. This article demonstrated how to calculate the zones of quiet for sound fields excited by broadband signals, and has presented examples with several low-pass- and band-pass-type random signals, comparing these to the well-known results for tonal excitations. It was shown that for simple low-pass filtered noise, the spatial correlation and the zone of quiet are comparable to those of a tone at the middle bandwidth frequency. Simulation results for near-field zones of quiet for a band-pass noise were comparable to a previous experiment[4] which used experimental study with a laboratory headrest system. The theory and tools developed here could be used to simulate and predict broadband zones of quiet more accurately in more realistic acoustic configurations which include real sources, a head and a control system, for example, but this is suggested for future work.

[1] P. A. Nelson and S. J. Elliott, *Active Control of Sound* (Academic, London, 1992).
[2] S. J. Elliott, ''Down with noise,'' IEEE Spectrum Magazine, **June 1999**, 51–61 (1999).
[3] H. F. Olson and E. G. May, ''Electronic sound absorber,'' J. Acoust. Soc. Am. **25**, 1130–1136 (1953).
[4] B. Rafaely, S. J. Elliott, and J. Garcia-Bonito, ''Broadband performance of an active headrest,'' J. Acoust. Soc. Am. **106**, 787–793 (1999).
[5] B. Rafaely and S. J. Elliott, ''H_2/H_∞ active control of sound in a Headrest: design and implementation,'' IEEE Trans. Control Syst. Technol. **7**(1), 79–84 (1999).
[6] J. L. Nielsen, A. Sabo, G. Ottesen, T. A. Reinen, and S. Sorsdal, ''A local active noise control system for locomotive drivers,'' Proceedings of Internoise 2000 conference, Nice, August 2000, pp. 2901–2905.
[7] C. Carme, D. Derrien, and G. Valentin, ''The ANCAS seat: extentions and industrial applications,'' in *Proceedings of ACTIVE 97 Conference*, Budapest, Hungary, August 1997, pp. 381–390.
[8] J. Garcia-Bonito and S. J. Elliott, ''Local active control of diffracted diffuse sound fields,'' J. Acoust. Soc. Am. **98**, 1017–1024 (1995).
[9] S. J. Elliott, P. Joseph, A. J. Bullmore, and P. A. Nelson, ''Active cancellation at a point in a pure tone diffuse sound field,'' J. Sound Vib. **120**(1), 183–189 (1988).
[10] P. Joseph, S. J. Elliott, and P. A. Nelson, ''Near field zones of quiet,'' J. Sound Vib. **172**(5), 605–627 (1994).
[11] P. Joseph, S. J. Elliott, and P. A. Nelson, ''Statistical aspects of active control in harmonic enclosed sound fields,'' J. Sound Vib. **172**(5), 629–655 (1994).

[12] R. K. Cook, R. V. Waterhouse, R. D. Berendt, E. Seymour, and M. C. Thompson, ''Measurement of correlation coefficients in reverberant sound fields,'' J. Acoust. Soc. Am. **27**, 1072–1077 (1955).
[13] B. Rafaely, ''Spatial-temporal correlation of a diffuse sound field,'' J. Acoust. Soc. Am. **107**, 3254–3258 (2000).
[14] A. D. Pierce, *Acoustics—An Introduction to its Physical Principles and Application* (McGraw-Hill, New York, 1981).
[15] F. Jacobsen, ''The diffuse sound field,'' Report No. 27, The Acoustic Laboratory, Technical University of Denmark (1979).
[16] M. R. Schroeder, ''The 'Schroeder frequency' revised,'' J. Acoust. Soc. Am. **99**(5), 3240–3241 (1996).
[17] L. E. Kinsler, A. R. Frey, A. B. Coppes, and J. V. Sanders, *Fundamentals of Acoustics*, 3rd ed. (Wiley, New York, 1982).
[18] L. L. Beranek, *Acoustics* (American Institute of Physics, New York, 1986).
[19] P. A. Nelson, J. K. Hammond, P. Joseph, and S. J. Elliott, ''Active control of stationary random sound fields,'' J. Acoust. Soc. Am. **87**, 963–975 (1990).
[20] P. Joseph, ''Active control of high frequency enclosed sound fields,'' Ph.D. thesis, University of Southampton, England, 1990.
[21] H. Nelisse and J. Nicolas, ''Characterization of a diffuse field in a reverbrant room,'' J. Acoust. Soc. Am. **101**, 3517–3524 (1997).

A new transducer holder mechanism for pipe inspection

Dongshan Guo and Tribikram Kundu
Department of Civil Engineering and Engineering Mechanics, University of Arizona, Tucson, Arizona 85721

(Received 16 October 2000; accepted for publication 13 April 2001)

A new transducer holder mechanism has been designed and fabricated for pipe inspection by cylindrical guided waves. Commercially available ultrasonic transducers have been used to generate compressional ultrasonic waves in the coupling medium. Those waves are converted to cylindrical guided waves in the pipe by the new coupling mechanism. A number of advanced coupling mechanisms developed recently for large plate and pipe inspection require the presence of a coupling fluid between the ultrasonic transducer and the pipe or plate specimen. These mechanisms can be used for inspecting horizontal pipes and plates. The new coupling mechanism presented in this article uses solid material as the coupler and can be used equally well for inspecting horizontal as well as inclined or vertical pipes. The new coupling mechanism has been designed to generate efficiently different Lamb modes in the pipe. The new design has been used to inspect different kinds of anomalies in copper and aluminum pipes. The preliminary results show that a number of Lamb modes, when generated properly by the new coupling mechanism, are very sensitive to the pipe defects. These experimental results along with the new design of the coupling mechanism are presented in this article. [DOI: 10.1121/1.1377289]

PACS numbers: 43.58.Vb, 43.35.Zc [SLE]

I. INTRODUCTION

Pipeline operators use a variety of methods to check pipes periodically. All these methods use a ''smart pig,'' a robotic device that crawls through the pipe carrying sensors. Smart pigs are expensive to run (several thousand US dollars per mile[1]) and the results sometimes are not reliable. Furthermore, to use pigs, pipelines must allow free passage and must have pig launch and retrieval facilities that are not always available. The smart pig technique also does not work for small diameter pipes.

Corrosion and pitting damages in small diameter pipes are detected by flash radiography and eddy current technique by measuring the wall thickness loss. However, because many pipes in industries are insulated, the conventional NDT (nondestructive testing) techniques are expensive for the insulation removal requirement.

Because of the limitations of the current ''smart pig'' technique and the radiography and eddy current techniques for inspecting insulated pipes, it is necessary to develop a technology to assess the integrity of the pipeline in a nonintrusive manner without removing insulation or disrupting its normal operation. One way of doing it is by launching Lamb waves or cylindrical guided waves in the pipe. Lamb waves can be launched in the pipeline by ultrasonic exciters placed either on the outer wall or inner wall (for large diameter pipes). Several investigators[2–8] have used Lamb waves for inspecting corrosion, cracks, and other types of material defects in pipes. In these efforts time histories recorded by the receiver have been carefully analyzed for detecting any small signal reflected by the defect. Some investigators used single symmetric modes[4,5] while others used antisymmetric multi-modes[6–8] for pipe inspection. Most of these efforts have been found to be very useful for detecting defects in pipes. However, often small ''defect signals'' may remain undetected by some Lamb modes.[8]

In this article we developed an alternative technique. We try to detect defects from the change in the shape and amplitude of the $V(f)$ curve. The $V(f)$ curve is the plot of the variation of the received signal voltage with signal frequency. Ideally, the strike angle and the signal frequency should be adjusted to launch the Lamb mode that is most sensitive to the pipe damage and least prone to attenuation and can propagate a long distance along the pipeline before losing its strength. The Lamb wave frequency should be set such that it is not affected by the noise generated by the fluid flowing through the pipe so that the pipes can be inspected under its normal operating conditions. We have designed and fabricated a number of new coupling mechanisms for efficiently generating different cylindrical Lamb modes in pipes. These new designs are described in this article and some experimental results generated by the new coupling mechanism are presented. The new design uses commercially available conventional ultrasonic transducers and does not require any special transducer like the ''comb transducer'' [8] for generating multiple Lamb modes in a pipe.

II. EVOLUTION OF THE COUPLING MECHANISM DESIGN

Kundu and associates[9–11] among others used Lamb waves for detecting internal defects in small composite plate specimens. Different Lamb modes in the plate were generated by rotating the transducer in the coupling fluid, thus changing the incident angle of the striking beam, and fine tuning the frequency of excitation. Ghosh and Kundu[12] then extended this technique and developed a conical container of the coupling fluid for inspecting large plates. Ghosh and Kundu's design is found to be very effective for detecting defects in large plates.[13] Guo and Kundu[14] followed similar steps and designed conical containers for coupling fluids for the horizontal pipe inspection. In the present article Guo and

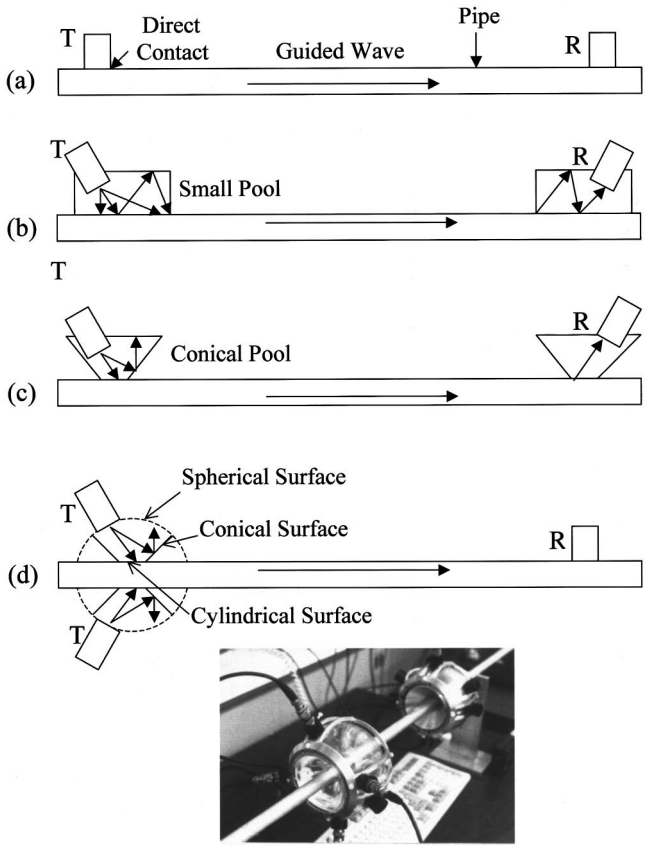


FIG. 1. Evolution of the coupler design: (a) direct contact, (b) small pool of coupling fluid, (c) conical pool, and (d) annular plexiglas holder. Bottom image: photograph of the coupler shown in (d).

Kundu's design has been modified and the coupling fluid has been replaced by a coupling solid for inspecting pipes oriented at any inclination (0 to 90 degrees). This solid coupler can efficiently generate axisymmetric and non-axisymmetric Lamb modes in a pipe. The new coupler can accommodate multiple transmitters as opposed to the single transmitter excitation in the previous design.[14] The advantage of using multiple transducers will be discussed later in Sec. V.

Figure 1 outlines the evolution of the coupler design for pipe inspection. Our first attempt was to place the transducer and receiver in direct contact with the pipe wall [Fig. 1(a)]. However, in this manner one does not have much control over which Lamb mode is generated. Then two small open-ended water containers were placed directly over the pipe wall as shown in Fig. 1(b) and the transducers were placed in those containers immersed in the coupling fluid. Thus the transducers could freely rotate and the incident beam could strike the pipe wall at any desired angle. The containers were placed on the pipe wall with putty to have a watertight contact. The container can be of any shape, as long as its bottom opening is small enough to fit on the pipe wall and the top part is large enough to accommodate the transducer. The received signal strength improved significantly with this arrangement. However, consistency of the received signal was not very satisfactory. The water surface disturbance caused the signal to change (which could be observed on the computer or oscilloscope screen). This was attributed to the multiple reflections of the signal within the water pool as shown in Fig. 1(b). Diverging signals from the transducer strike the pipe at various angles [see Fig. 1(b)], generating different Lamb modes in the pipe; this is not desirable. It contributes to the inconsistency in the experimental results.

Next, the water container shape was made conical, as shown in Fig. 1(c).[14] The diverging signals from the transducer after being reflected by the container wall could not reach the specimen with the conical pool arrangement of Fig. 1(c). The inconsistency problem arising from the diverging beam and the water surface disturbance could be avoided by this conical water container design. Ghosh and Kundu[12] have discussed in detail the advantages of conical water containers for large plate inspections.

Although these conical water containers gave significant improvement in comparison to the direct contact type arrangement, the main constraint of this arrangement is that the pipe must be kept horizontal.[14] Second, to inspect another segment of the pipe, the water container must be removed from its previous position and placed very carefully (should be made watertight with putty) in the new position. Use of putty at the water-container–pipe-wall junction to make it watertight is time consuming. Maximum inclination angle of the transducer was limited by the angle of the conical container.

To avoid these shortcomings associated with the conical containers a new transducer mounting geometry is envisioned. It has annular shape with cylindrical and conical inner walls and spherical outer wall as shown in Fig. 1(d). The pipe tightly fits inside the cylindrical inner hole of the annular shape. The transducers are mounted on the spherical outer surface. One or more transducers can be placed on the spherical surface. A photograph of this coupler is also shown in Fig. 1. With this coupler the pipe can be excited by multiple transducers, vibrating in the same phase. The striking angle can be changed simply by sliding the transducers on the spherical outer surface. The inner conical surface reflects away the diverging beams [Fig. 1(d)], and those beams cannot strike the pipe specimen.

III. SPECIMENS

Pipes made of aluminum and copper, with different types of anomaly or fabricated defect, have been inspected. Specimens include a copper pipe with (and without) an outer coating over a small region and a number of aluminum pipes with notches or grooves and milled slots cut on its outer and inner surfaces.

First a copper pipe [Fig. 2(a)] of 22.23-mm (0.875 in.) outer diameter and 20.6-mm inner diameter was inspected using the coupler geometry, shown in Fig. 1(d). The receiver was placed at a distance of 2 m from the transmitter. Received signals were first generated for a free pipe. Then these signals were again recorded after a silver tape was wrapped around a small region of the pipe between the transmitter and the receiver. The tape simulates a coating on the outer wall of the pipe. The tape width is 50.8 mm and the outer diameter of the tape-wrapped region is 23.4 mm.

A second set of pipe specimens made of aluminum was fabricated to study the sensitivity of Lamb modes to notches and slots. Notches and slots in pipes represent corrosion and

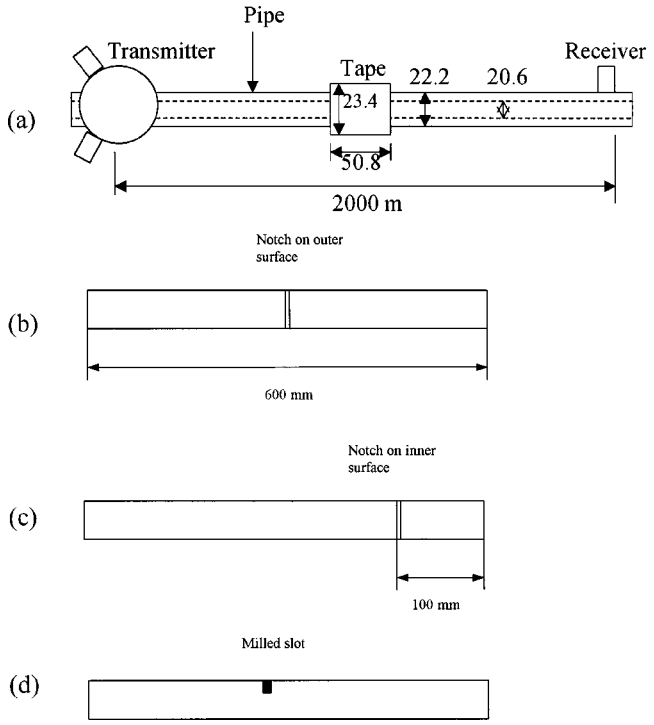


FIG. 2. (a) Geometry of the copper pipe, shown with the taped region and transmitter/receiver arrangement. All dimensions are in mm. (b), (c) and (d) Aluminum pipes (o.d. = 22.23 mm, wall thickness = 1.59 mm) with circular notches (100% circumference) and milled slot (10% circumference).

pitting defects, respectively. Figure 2(b) shows an aluminum pipe (600 mm long, 22.23-mm outer diameter, and 1.59-mm wall thickness) with a notch (5 mm wide and 0.7 mm deep) on the outer surface. Figure 2(c) shows a second aluminum specimen with a notch of the same dimension as in Fig. 2(b), but cut on the inner surface of the aluminum pipe. Since it is difficult to machine the notch on the inner surface, the notch is placed only 100 mm from one end of the pipe. Figure 2(d) shows a third aluminum specimen with a milled slot (5 mm wide and 0.7 mm deep) cut on the outer surface over only 10% of the circumference.

IV. EXPERIMENTAL SETUP

A schematic of the experimental setup is shown in Fig. 3. A personal-computer-based data acquisition software controls a model 395 Wavetech waveform generator, which sends out tone-burst signals in the sweeping frequency mode, continuously varying from the lowest frequency to the highest frequency. The signal from this generator is amplified by a model 310 Matec broadband gated amplifier and then used to excite the transmitting transducers, mounted on the new coupling geometry. The receiving transducer is located at a distance (2 m for the copper pipe and 0.5 m for the aluminum pipe) from the transmitter and in direct contact with the pipe wall. The received signal is amplified by a Tektronix TM 506 amplifier and is routed back to the computer through an interface. A Gage 40 MHz data acquisition A/D card converts the analog signal into digital signal. A software module then processes the signals and produces the received voltage

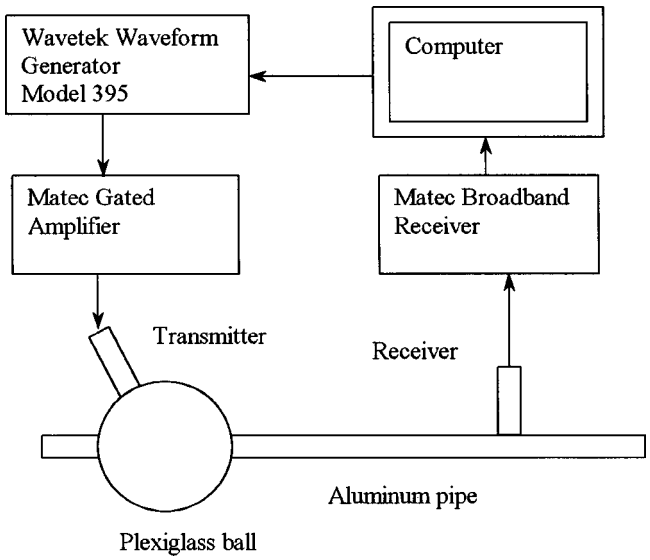


FIG. 3. Schematic of the experimental setup.

versus signal frequency or the $V(f)$ curve that can be displayed on the computer screen or can be stored in a file for future analysis.

V. EXPERIMENTAL RESULTS

One should first obtain the optimum incident angle (θ) for generating strong cylindrical guided waves in a pipe. One can obtain θ theoretically or experimentally. To determine θ experimentally one must record the received signal amplitudes for different incident angles of the striking beam and see for what inclination the received signal is strongest.

On the other hand, to predict θ theoretically one should compute the phase velocity dispersion curves for the pipe knowing its geometry and material properties. Then one should evaluate the optimum angle from Snell's law,

$$\theta = \sin^{-1}\left(\frac{\nu_c}{\nu_{ph}}\right), \tag{1}$$

where ν_c is the longitudinal wave speed in the coupling medium between the transducer and the pipe (for Plexiglas $\nu_c = 2.77$ km/s), and ν_{ph} is the phase velocity of the cylindrical guided wave in the pipe. Here θ is obtained experimentally.

Before inspecting the specimens shown in Fig. 2, it is first investigated how the number of transmitters and their inclination angles affect the received signal strength. To this aim $V(f)$ curves for a copper pipe in absence of any anomaly are generated using different numbers (one to four) of transmitters for four different incident angles, 0, 8, 16, and 24 degrees. The angle is measured relative to a plane normal to the pipe axis. When multiple transducers are used, then those are placed at uniform spacing around the periphery of the pipe. In other words, when two transducers are used, those are placed at 180 degrees apart, three transducers are placed at 120 degrees apart, and four transducers are placed at 90 degrees apart. Results are shown in Fig. 4. From Fig. 4 one can see that the strength of the received signal increases

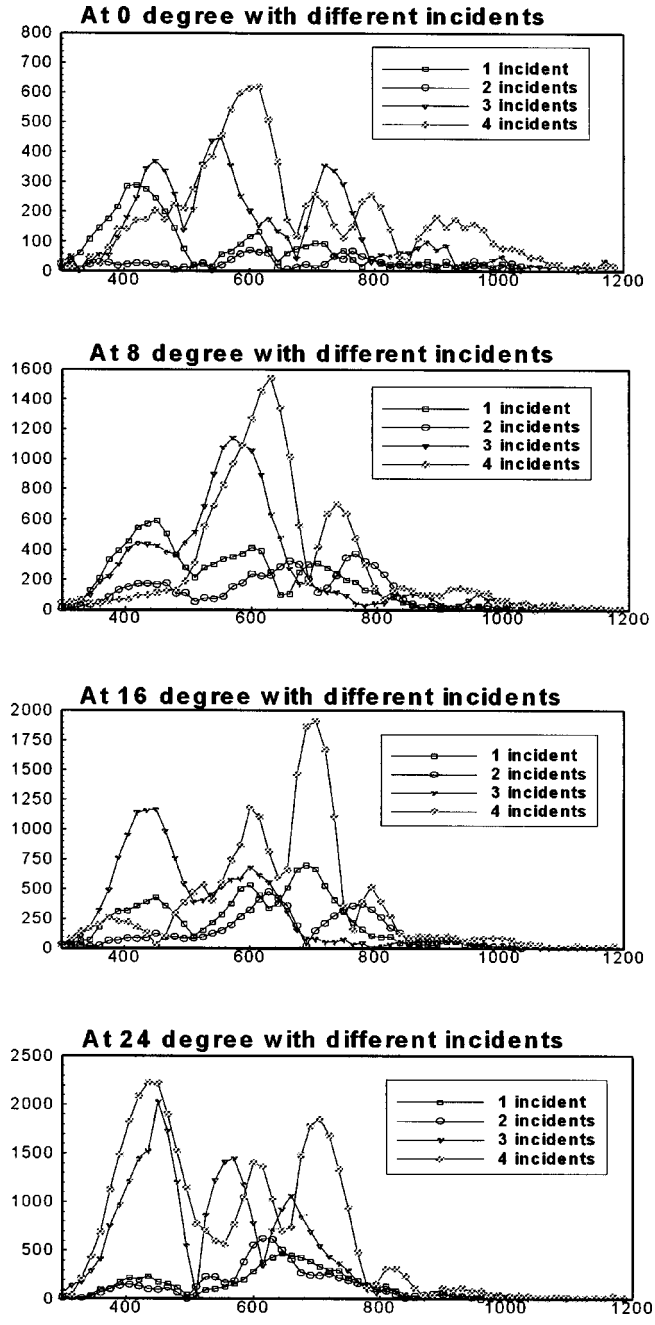


FIG. 4. Four curves in each graph show $V(f)$ curves generated by one, two, three, and four transducers inclined at 0- (top), 8- (second from top), 16- (third from top), and 24-degree (bottom) angles for a copper pipe (abscissa is in kHz).

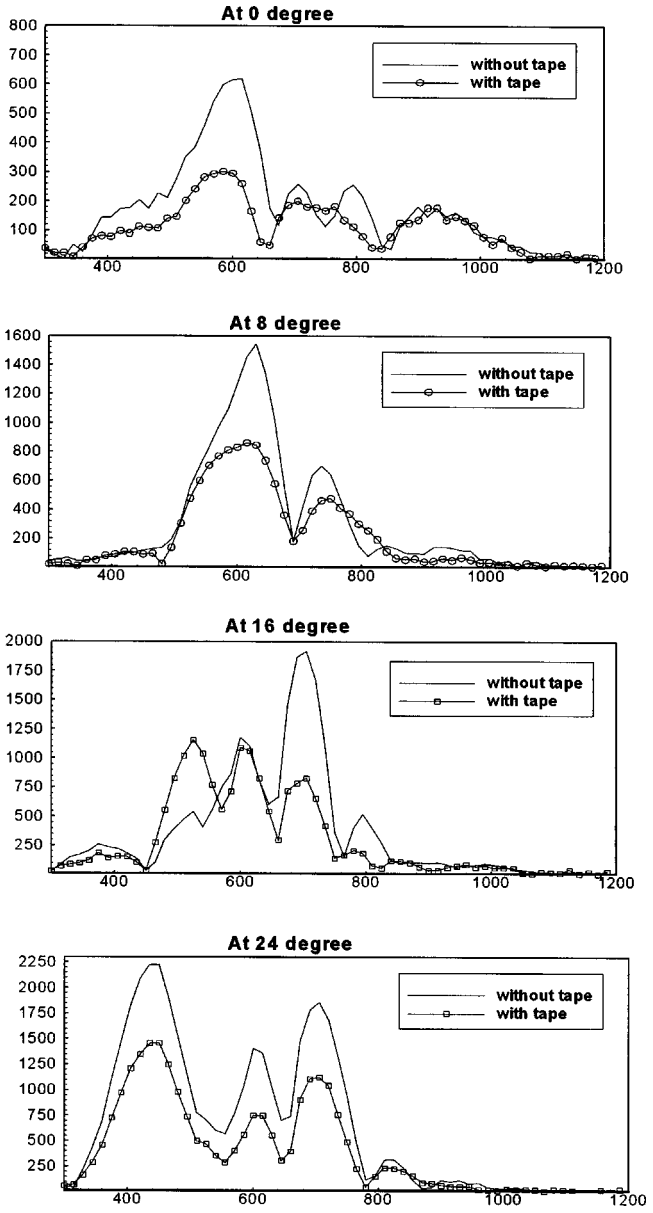


FIG. 5. Two curves in each graph show $V(f)$ curves generated by copper pipes with (line with circles) and without (continuous line without circles) the tape [Fig. 2(a)]; transducers are inclined at 0 (top), 8, 16, and 24 degrees (bottom), (abscissa is in kHz).

significantly as the number of transmitters increases. Hence, multiple transducer excitation is preferable to single transducer excitation.

A comparison of the strengths of the strongest modes generated by four transducers inclined at 0, 8, 16, and 24 degrees shows that for the copper pipe the strongest signal is obtained when the transducers are inclined at 24 degrees. In Fig. 4 one can see that the strongest peak values are approximately 650, 1600, 1950, and 2250 for 0, 8, 16, and 24-degree angles of inclination. From this exercise we conclude that for generating strong cylindrical guided waves in a pipe that are capable of propagating a long distance along the pipe, one must use multiple transducers inclined at an optimum angle. However, to generate non-axisymmetric Lamb modes in the pipe a smaller number (one or two) of transmitters is recommended. A larger number (four or more) of transducers uniformly placed around the pipe and excited in same phase will generate strong axisymmetric modes and weak non-axisymmetric modes.

To investigate if these strong Lamb modes are sensitive to pipe anomalies such as coatings, $V(f)$ curves have been generated for a copper pipe in the absence of any coating and when a small part of the pipe is wrapped by a silver tape [Fig. 2(a)]. Four transmitters are used and the inclination angle is varied from 0 to 24 degrees. The experimental results are shown in Fig. 5. One can see from this figure that for all four incident angles the peak values are significantly reduced in the presence of the tape.

After the initial success of our effort to generate in a

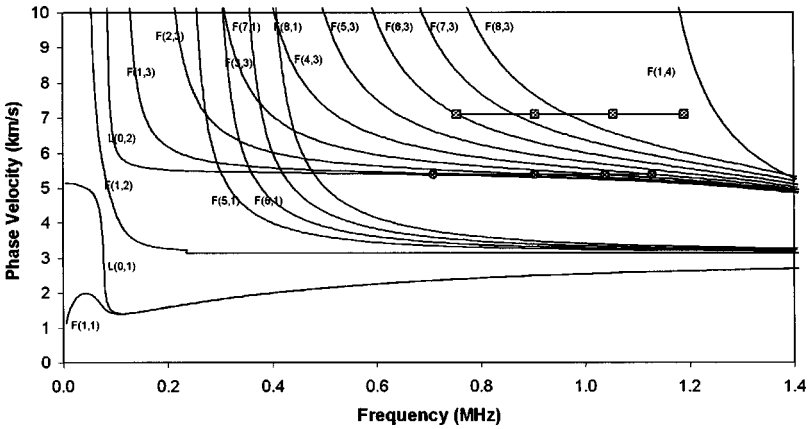


FIG. 6. Phase velocity dispersion curves for aluminum pipe—22.23-mm outer diameter and 1.59-mm wall thickness. Squares and circles show the frequency–phase velocity combinations where peaks in the experimental $V(f)$ curves, shown in Figs. 8 and 7, are obtained.

pipe strong Lamb modes that are sensitive to pipe anomalies, a careful investigation is carried out on the defect detection by Lamb modes in aluminum pipes. It is investigated which Lamb mode and what frequency-incident angle combination should be selected for detecting a specific defect in the pipe. To this aim aluminum pipes with milled slots and notches cut on the outer and inner surfaces are inspected. Specimen geometries for this investigation are shown in Figs. 2(b), (c), and (d).

Experiments were carried out after positioning the transmitting transducer at two different angles, 23 and 31 degrees for which strong Lamb modes could be generated in the aluminum pipe. The angle was measured from the line perpendicular to the pipe axis. For this experiment one transmitter was used so that both axisymmetric and non-axisymmetric (or flexural) modes could be generated.

The phase velocity dispersion curves for a homogeneous elastic aluminum pipe are obtained by Gazis' technique[15,16] and shown in Fig. 6. These were generated by taking the longitudinal and shear wave speeds in aluminum equal to 5.62 and 3.34 km/s, respectively. The voltage amplitude versus frequency or $V(f)$ curves (Figs. 7–9) were obtained by continuously changing the frequency of 15 cycles of the tone burst signal and synchronous detection of the amplitude value of the received signal. The peaks in the $V(f)$ curves correspond to Lamb modes. For Fig. 7 the incident angle is 31 degree and from Snell's law [Eq. (1)] one can see that 31 degrees corresponds to a phase velocity of 5.38 km/s. The peak positions and phase velocity for Fig. 7 are shown on the dispersion curve plot (Fig. 6) by solid circles.

Figures 7–9 show $V(f)$ curves generated in four aluminum pipe specimens. Geometries of three specimens are shown in Figs. 2(b), (c) and (d); the fourth specimen is a defect-free aluminum pipe having the same outer diameter and wall thickness as other three specimens. The notch is used to simulate the corrosion and the milled slot simulates the pitting.

The $V(f)$ curves for defect-free and defective specimens with outer surface notch and inner surface notch are shown in Figs. 7 and 8. Figure 7 is generated for a 31-degree incident angle. This figure shows that lower frequency modes are very sensitive to both outer surface and inner surface notches. The amplitude change of these modes due to the presence of the inner surface notch is even larger. However, in Fig. 8 (transmitter inclination angle is 23 degree, corresponding phase velocity from Snell's law is 7.09 km/s) Lamb modes (peaks) show little sensitivity to the presence of pipe notches although the signal frequency range (500 to 1400 kHz) is same for Figs. 7 and 8.

When guided waves travel through a defect, mode conversion often occurs. In Fig. 7 one can see that the first peak that corresponds to the $L(0,2)$ mode almost disappears in presence of the defect.

The peak positions and corresponding phase velocity

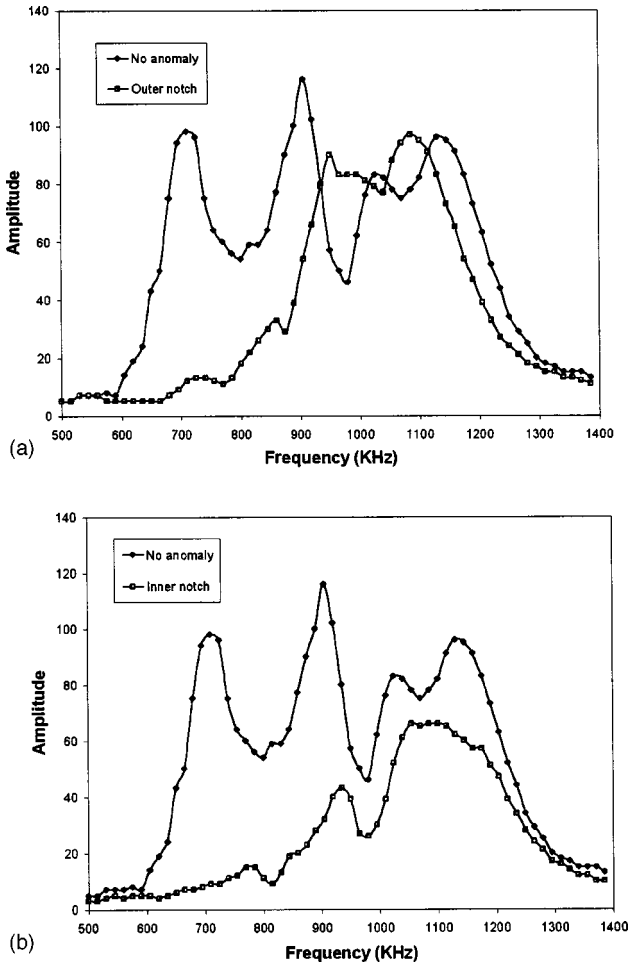


FIG. 7. $V(f)$ curves for defect-free and defective specimens, (a) with outer notch and (b) with inner notch, for 31-degree incident angle.

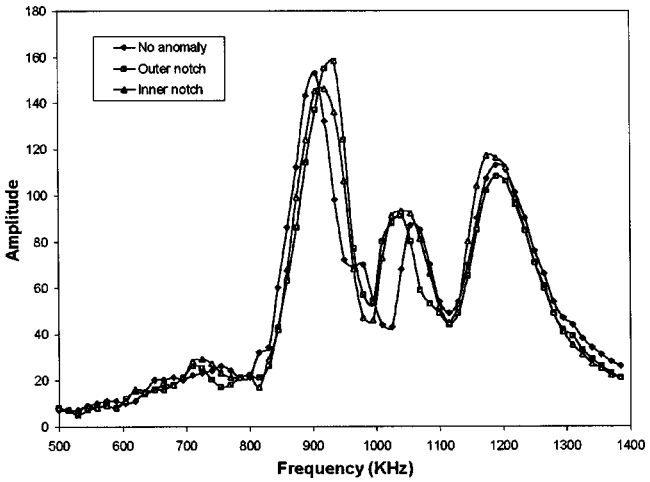


FIG. 8. $V(f)$ curves for defect-free and defective specimens (with outer and inner notches) for 23-degree incident angle.

(7.09 km/s) for Fig. 8 are shown by solid squares in Fig. 6. One can see in Fig. 6 that these peaks correspond to higher-order flexural modes.

The general conclusion of Figs. 7 and 8 is that the striking angle that generates Lamb modes near the high-frequency asymptote of the phase velocity dispersion curves is more effective for detecting defects than the angle that generates Lamb modes away from the high-frequency asymptote. Ghosh *et al.*[13] observed similar phenomenon for plates.

The ability of selected guided wave modes to detect the milled slot [Fig. 2(d), over 10% of the circumference] that simulates the pitting defect is shown in Fig. 9. Here the transmitter is inclined at 31 degrees. The experiments were carried out by positioning the milled slot first at the top, and then rotating the pipe 180 degrees so that the slot is positioned at the bottom of the pipe. In the first setup the transmitting transducer, the receiving transducer, and the milled slot are placed on the top of the pipe. The second setup keeps the transducers at the same positions but the slot is placed at the bottom by rotating the pipe. One can see that the guided

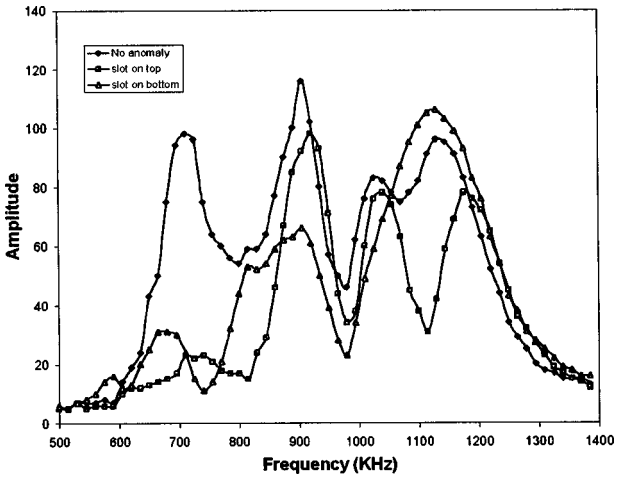


FIG. 9. $V(f)$ curves for defect-free and defective (milled slot defects in two different positions) specimens for 31-degree incident angle.

waves have good sensitivity to the defects even when the defect was not aligned with the transducers.

VI. CONCLUDING REMARKS

In this article it is experimentally shown that cylindrical guided wave modes that are sensitive to pipe defects can be launched in a pipe by a new coupling mechanism. The new design for the coupler proposed here uses commercial ultrasonic transducers and can inspect a pipe inclined at any angle. It has a spherical outer surface, a cylindrical inner surface, and two conical surfaces. The transducers are mounted on the spherical outer surface and the pipe is press fitted against the cylindrical inner surface. The ultrasonic energy is transmitted into the pipe through the cylindrical inner surface. The spherical outer surface allows transducers to be inclined relative to the pipe at any desired angles. The conical surfaces prevent the diverging signals from reaching the pipe wall. It is shown that transducers inclined at an angle produce much stronger guided waves in the pipe compared to those placed normal to the pipe wall. The correct inclination angle and frequency for producing a strong Lamb mode in the pipe can be determined experimentally or theoretically from the dispersion curves of the pipe. This investigation shows that the Lamb modes near the high-frequency asymptote of the dispersion curves generated by the transmitter inclined at a relatively larger angle are more sensitive to defects in comparison to the flexural modes of higher phase velocity generated by the transmitter inclined at a relatively lower angle.

The experiment was carried out on pipes oriented at different angles and, as expected, the inclination angle of the pipe was found to have no effect on the recorded signal. This new coupling mechanism appears to have an excellent potential for pipe inspection by Lamb waves using commercial transducers. However, more study is needed to conclude how small a defect from a given distance can be detected by this technique. Applicability of this technique in the field environment, when the pipe is buried and/or when its surface is rough and mildly corroded, needs to be investigated.

ACKNOWLEDGMENTS

This research was financially supported by the NSF Grant Nos. CMS-9896182 and CMS-9901221 and the EPRI Grant No. EP-P241/C110.

[1] J. F. Kiefner, R. W. Hyatt, and R. J. Eiber, ''NDE Needs for Pipeline Integrity Assurance,'' American Gas Association Catalog Number L51505 (1986).
[2] D. N. Alleyne and P. Cawley, ''The Long Range Detection of Corrosion in Pipes Using Lamb Waves,'' in *Review of Progress in QNDE*, Vol. 14B, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1995), p. 2073.
[3] C. W. Chan and P. Cawley, ''Guided Waves for the Detection of Defects in Welds in Plastic Pipes,'' in *Review of Progress in QNDE*, Vol. 14B, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1995), p. 1537.
[4] D. N. Alleyne and P. Cawley, ''Long Range Propagation of Lamb Waves in Chemical Plant Pipework,'' Mater. Eval. **45**, 504–508 (1997).
[5] H. Kwan and A. Holt, ''Feasibility of under-lagging corrosion detection in steel pipe using the magnetorestrictive sensor technique,'' NDT & E Int. **28**, 211–214 (1995).

[6] J. J. Ditri, J. L. Rose, and G. Chen, ''Mode Selection Criteria for Defect Detection Optimization Using Lamb Waves,'' in *Review of Progress in QNDE*, Vol. 11, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1992), pp. 2109–2115.
[7] J. L. Rose, M. J. Quarry, A. V. Bray, and C. J. Corley, ''Corrosion Under Insulation Detection Using Guided Waves,'' ASNT Fall Conference in Pittsburgh, PA, October 1997, pp. 1–3.
[8] M. J. Quarry and J. L. Rose, ''Multimode Guided Wave Inspection of Piping Using Comb Transducers,'' Mater. Eval. **57**, 1089–1090 (1999).
[9] T. Kundu, K. Maslov, P. Karpur, T. Matikas, and P. Nicolaou, ''A Lamb Wave Scanning Approach for Mapping of Defects in [0/90] Titanium Matrix Composites,'' Ultrasonics **34**, 43–49 (1996).
[10] K. I. Maslov and T. Kundu, ''Selection of Lamb Modes for Detecting Internal Defects in Laminated Composites,'' Ultrasonics **35**, 141–150 (1997).
[11] W. Yang and T. Kundu, ''Guided Waves in Multilayered Anisotropic Plates for Internal Defect Detection,'' J. Eng. Mech. **124**, 311–318 (1998).
[12] T. Ghosh and T. Kundu, ''A New Transducer Holder Mechanism For Efficient Generation and Reception of Lamb Modes in Large Plates,'' J. Acoust. Soc. Am. **104**, 1498–1502 (1999).
[13] T. Ghosh, T. Kundu, and P. Karpur, ''Efficient Use of Lamb Modes for Detecting Defects in Large Plates,'' Ultrasonics **36**, 791–801 (1998).
[14] D. Guo and T. Kundu, ''A New Sensor for Pipe Inspection by Lamb Waves,'' Mater. Eval. **58**(8), 991–994 (2000).
[15] D. Z. Gazis, ''Three Dimensional Investigation of the Propagation of Waves in Hollow Circular Cylinders. I. Analytical Foundation,'' J. Acoust. Soc. Am. **31**, 568–573 (1959).
[16] J. L. Rose, *Ultrasonic Waves in Solid Media* (Cambridge U. P., Cambridge, 1999).

Three-dimensional sound localization from a compact non-coplanar array of microphones using tree-based learning

Juyang Weng and Kamen Y. Guentchev
Department of Computer Science and Engineering, Michigan State University, East Lansing, Michigan 48824

(Received 28 May 1999; accepted for publication 13 April 2001)

One of the various human sensory capabilities is to identify the direction of perceived sounds. The goal of this work is to study sound source localization in three dimensions using some of the most important cues the human uses. In an attempt to satisfy the requirements of portability and miniaturization in robotics, this approach employs a compact sensor structure that can be placed on a mobile platform. The objective is to estimate the relative sound source position in three-dimensional space without imposing excessive restrictions on its spatio-temporal characteristics and the environment structure. Two types of features are considered, interaural time and level differences. Their relative effectiveness for localization is studied, as well as a practical way of using these complementary parameters. A two-stage procedure was used. In the training stage, sound samples are produced from points with known coordinates and then are stored. In the recognition stage, unknown sounds are processed by the trained system to estimate the 3D location of the sound source. Results from the experiments showed under $\pm 3°$ in average angular error and less than $\pm 20\%$ in average radial distance error.
[DOI: 10.1121/1.1377290]

PACS numbers: 43.58.Yb [SLE]

I. INTRODUCTION

A sound produced by a point-source generates acoustic waves with spherical symmetry, assuming uniform density of the surrounding air and absence of obstacles or other sounds. It is known that the location of the source can be established by detecting the front of the propagating wave and computing the center of the sphere.[1,2] Unfortunately acoustic waves are not clearly distinguishable objects and such a task is not trivial in real environments even if real-life sources could be approximated by points.[3] Numerous studies have attempted to determine the mechanisms used by humans to achieve dimensional hearing.[2,4,5] Most phenomena have been reasonably explained in principle, although many aspects of human dimensional hearing need further study. It is known that two of the most important cues used by humans are the interaural differences: in time and level (ITD and ILD).[3,6,7] Other cues relate to the spectral variations caused by diffractions at the head and pinnae.[8] For sounds with longer duration, cognitive processes start playing an important role, including dynamic head adjustments, high-level reasoning, etc.[7] However, the computational steps in the use of spectral variations and other information by the human cognitive process are still not well understood. The purpose of the work presented here is sound localization by machines. We use only two low-level cues, ITD and ILD, in the work presented here.

A. Sound localization by machine

Sound localization can be used in many different applications: robot hearing, human-machine interfaces, monitoring devices, handicappers' aids, etc., where other means fail for different reasons. The obvious importance of building sound localization devices has prompted numerous efforts in the research community and a variety of techniques has been developed. Driven by concrete application needs, sensor setups of different implementations have seldom attempted to follow the human model. The number, size, and placement of the sensors in such devices follow the specific needs of the task and are optimized for accuracy, stability, ease of use, etc. For example, a number of microphone subarrays have been placed on the walls with a goal to pick up the location of a speaker in a room.[9–12] In other studies a human model has been followed to some degree, resulting in constraints in applicability and limited accuracy.[13] A significant amount of work has been devoted to devices with a limited functionality (e.g., constrained to localization in a single half-plane while still using large sensor structures)[12–14] or the help of a nonacoustical modality has been used (e.g., vision).[14]

In contrast to large, fixed sensor arrays for special situations and environments, this work concentrates on a compact, mobile sensor array that is suited for a mobile robot to localize 3D sound sources with moderate accuracy. It can be positioned arbitrarily in space while being capable of identifying the relative position of an arbitrarily located sound source. It is necessary to point out that the human three-dimensional sound localization capabilities, while amazingly accurate in some instances, often have very serious limitations. The precision depends on various characteristics of the perceived sound: spectral contents, envelope variability as a function of time, volume level, reverberation and echo, etc. It can be disappointingly low and in some instances totally inconclusive.[4] Sometimes it can be convincingly wrong (e.g., Franssen effect).[5] One major difference between human and engineering setup is the number of sensors available.

Most authors distinguish a single parameter as the most significant factor for dimensional sound localization. It is the interaural time difference (ITD) of the sound as perceived by two sensors. Numerous studies report the ITD as the main

 0001-4966/2001/110(1)/310/14/$18.00

cue in human dimensional hearing.[16] The clear geometrical representation of the problem makes it a favorite feature to be used when approaching such a task by a machine setup.[9,11,12,14–18] Another cue known to have notable importance in human dimensional hearing is the interaural level differences (ILDs). Surprisingly ILDs have seldom been used in actual system implementations because they are believed to have unfavorable frequency dependence and unreliability.[3,13] Another reason is the lack of an explicit and stable relationship between ILD and source location which would otherwise allow for a simple algorithmic solution to be derived.[3] The learning approach used in this study does not have such limitations and it benefits from the added cues, both ITD and ILD.

Finally, the processing of the extracted features is one of the dominating factors for the success of a localization procedure. Most works determine the ITD and then use either an iterative search algorithm to minimize a certain objective function,[12,13,19] or an approximation model for which a closed-form solution can be derived.[9,10] The former is relatively slow and, thus, it may not reach real time speed. The latter introduces model errors and cannot use more feature types for better accuracy.

To use both interaural time differences (ITDs) and interaural level differences (ILDs) while effectively dealing with the complex nonlinear relationships among these feature measurements and the solution, our work employs a learning based approach. It consists of a training phase and a performance phase. In the training phase, sounds from known 3D positions are generated for training the system, during which a fast retrieval tree is built. In the performance phase, the system approximates the solution by retrieving the top match cases from the retrieval tree. This flexible framework allows for the use of more than one type of feature, and to deal with the 3D localization problem without imposing unrealistic assumptions about the environment, despite the compactness of the sensor structure. As far as we know, this work is the first to use a compact non-coplanar sensor array for full 3D sound localization.

In order to objectively evaluate the performance of the system, initially a linear search algorithm was used when searching for the nearest neighbors in the 12-dimensional input space. The obtained results were used to evaluate the correctness and the performance of the SHOSLIF procedure. (SHOSLIF stands for Self-organizing Hierarchical Optimal Subspace Learning and Inference Framework.) SHOSLIF achieves a high speed of retrieval due to its logarithmic time complexity $O(\log(n))$, where n is the number of cases learned and stored as necessary.[20] It was found that the results produced by SHOSLIF had identical precision with that of the linear search, while its performance speed was nearly five times faster.

B. Related experimental work

Extensive work on sound localization by microphone arrays has been performed by Brandstein *et al.* with a significant effort in the theory of speech-based sound localization and a series of publications.[9–11,15] Novel methods for estimating ITD and using it for the localization are presented, among which are a pitch-based approach to time delay estimation and a closed-form location estimator. The implementations concentrate on room oriented solutions—microphone arrays placed in room walls. The sound localization is applied in three dimensions, with one of the dimensions (the vertical axis) having a lesser span than the other two (the height of the room is much smaller then its width or length). The reported accuracy of localization is very high—the space resolution is on the order of a few centimeters. The placement of the sensors relative to the sound source in Brandstein's experiments is different from our choice, and so is their use of multiple microphone arrays, compared to a single array in our work. This makes it difficult to compare the accuracy of both experiments. In the case of room-oriented placement it is impossible to define a direction or distance from the arrays since the arrays surround the source. An absolute location measure is used instead. Furthermore, the application domain of the room-oriented implementation is constrained to indoor use, while a free-standing compact array can be transported and used in more diverse environments.

A model for 3D sound localization that uses both ITD and ILD is presented by Martin.[13] The work simulates the human auditory system and uses a binaural setup. Some assumptions and approximations of the real-life environment are made and, as a consequence, the system can only estimate angular direction but not distance. The experiments are performed in an idealized environment (e.g., noise is eliminated) and an angular accuracy of around 5 degrees is achieved. All computations were performed offline. In our work estimating the distance to the source is shown to be a challenging problem and it is the parameter we measure with the least accuracy. In our experiment ambient noise is always present and the computations were performed in real time.

One of the applications with hybrid methodologies is the work of Bub *et al.*[14] They use a second modality (vision) to improve the performance of the localization device by refining the location estimate, provided by the sound localization. Their approach involves the use of ITD only as determined by a linear array of 15 microphones for redundancy. This setup limits the localization domain to a single horizontal 2D half-plane in front of the array. With acoustic means only, their system is estimated to achieve an angular accuracy of around 5 degrees and under 10% in distance with background noise only. This error is estimated from a single test point in 2D space in front of the array. No statistical estimates of the accuracy were presented. The authors indicate that competing noise can significantly degrade the accuracy of localization.

Another work with a practical implementation is by Rabinkin *et al.*[12] Their system uses two subarrays of four coplanar microphones each, placed on room walls. It employs a DSP to estimate ITD by using a cross-power spectrum phase algorithm. Then a space search algorithm minimizes the error between estimated and computed delays to produce a location estimate. The performance with the algorithms being run offline was reported as a percentage of deviation from a preset bound of 6 degrees and was reported as being generally lower than 20%. The performance from the

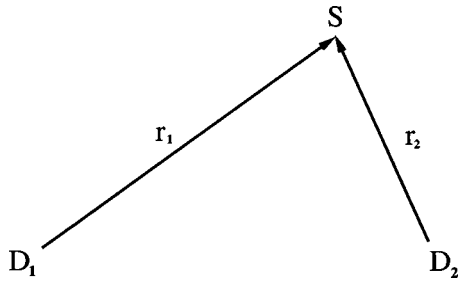


FIG. 1. Two detectors, D_1 and D_2, and the sound source S.

online tests was reported as being from ''moderately well'' to ''quite poor.''

A number of other publications[4,5,7,8,16–19,21] treat theoretical aspects of auditory localization and the various methodologies used in practical sound source localization. However, there is no experimental work or actual implementations reported. The problem of full three-dimensional sound source localization with a compact mobile structure, as examined by this article, has not been studied in any existing works that we are aware of.

II. THEORETICAL ISSUES AND SENSOR STRUCTURE

This work targets versatile applications such as the dimensional hearing of a mobile robot. For this reason we cannot use room-oriented solutions,[10,12] which typically use a large intersensor distance, with all the sensors fixed in the room. In our case the sound source will necessarily be located outside of the sensor structure. Furthermore, the distance to the source will generally be significantly larger than the span of the sensor structure. Most of the sound sources that are of interest for the purposes of sound localization are compact enough to be assumed point sources. If the source cannot be approximated by a point, then the problem of localizing that source is different from what we address here and thus is outside the scope of this work. The same applies to the case of multiple sources with comparable sound intensity. To determine the minimum number of sensors and their optimal placement, we need to look into the geometrical aspects of the problem.

A. Sensory measurements

From the front of the spherical acoustic wave, the two main parameters of measurements for each pair of sensors are ITD and ILD. Assuming that the speed of sound is constant, which is true only for uniform media (density, temperature, chemical and physical contents, etc.), ITD is equal to the difference of the distances between each of the detectors and the sound source, divided by the speed of sound. For simplicity we leave the constant out of the equation:

$$\mathrm{ITD} \sim r_1 - r_2, \tag{1}$$

where r_i is the distance between the sound source and the ith microphone, $i=1,2$, and $\sim$ indicates proportionality (Fig. 1). The constant (the speed of sound in air) is irrelevant to the localization and is left out. Also, since the amplitude of the sound wave, or the intensity of the sound, varies inversely with the square of the distance, the ILD is proportional to the difference between the inverse values of the square of the distance. However, if we take the difference, we will be confronted with high-order terms in the equations which will lead to unnecessary complication of the computations. A much simpler form is provided by the ratio of the two values:

$$\mathrm{ILD} \sim \frac{r_2^2}{r_1^2}. \tag{2}$$

Both parameters in (1) and (2) can be estimated from the signals, detected by a pair of microphones.

B. Uniqueness

The major theoretical issues that need to be discussed here are (1) whether there is a unique solution to the 3D location of the sound source, (2) how many sensors are minimally required for a unique solution, and (3) the reliability of the solution. In this section, we consider the first two issues that are related to uniqueness. The third issue will be discussed in Sec. II C.

It is known that spatial hearing in humans involves spectral analysis, which is beyond the scope of this paper. For the uniqueness issues related to our experimental system, we will consider three cases here: using ITD only, using ILD only, and using both. Further implications are necessary to make the theoretical uniqueness problem tractable. We assume that the sound source and every sensor are all compact enough so that we can consider each as a point source. Further, we also simplify the environment. We assume that the environment is filled with uniform still air, free of objects other than the sound source and microphones whose volume is negligible. Although the above assumptions are generally not exactly true in a realistic application environment, the results from the analysis will give us insight into the important issue of uniqueness. A proper system design may help to approximately satisfy these assumptions. For example, we use only light supporting material for our microphone array to minimize its sound absorption and reflection. It is worth noting that our actual sound localization method is based on learning and it does not use the mathematical relations discussed in this section.

A pure algebraic analysis of this complex problem seems neither tractable nor intuitive either. In the following analysis we choose to use an analysis approach that is largely geometric in nature. Consequently, our analysis considers general placements of sensors and it excludes various degenerate cases that defeat our original purpose, such as various coplanar four-detector placements.

1. Using ITD only

First, consider a sound detector pair which consists of detectors D_1 and D_2. As shown in Fig. 2, without loss of generality, we position our 3D coordinate system in such a way so that D_1 is at $(c,0,0)$ and sensor D_2 is at $(-c,0,0)$. By definition, a hyperbola (in 2D) is the set of points for which the difference of the distances from two fixed points D_1 and D_2 is constant, as shown in Fig. 2. In 3D, such points

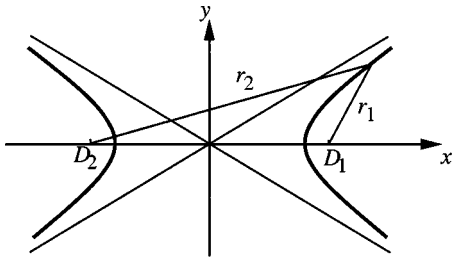


FIG. 2. From two detectors, a given ITD determines one sheet of a hyperbola.

lie in a rotational hyperboloid whose intersection with the *xy* plane is shown in Fig. 2, where the axis of symmetry is the *x* axis.

Suppose that the distances between a point $X=(x,y,z)$ and D_1 and D_2 are denoted by r_1 and r_2, respectively. The ITD gives $r_2-r_1=\text{ITD}=2a$, where we define variable a as equal to half of the ITD. If a is positive, only the right sheet in Fig. 2 can contain the sound source. Otherwise, only the left sheet can. If $a=0$, the two sheets of the hyperboloid are degenerated into a plane $x=0$ in 3D.

Next, consider three noncollinear detectors, D_1, D_2, and D_3, as shown in Fig. 3. The plane in which these three detectors lie is called *the detector plane* of detectors 1-2-3. The hyperboloid determined by the ITD of detectors *i* and *j* is called hyperboloid $i-j$. Thus, hyperboloid $i-j$ is the same as hyperboloid $j-i$. In Fig. 3, hyperboloid 1–2 and hyperboloid 1–3 intersect in 3D to give a 3D curve C_{123}. The third hyperboloid 2–3 (which is a plane in Fig. 3) does not give any additional constraint, since all the three sensors have been considered already by the first two hyperboloids 1–2 and 1–3. We define the *solution set* for the sound source as the set of all the possible 3D locations of the sound source. Therefore, we have the following observation:

Observation 1: All the ITD measures from three noncollinear sound detectors generally define an infinite number of solutions for the 3D location of sound source and the solution set is a 3D curve.

In Fig. 3, the hyperboloid 2-3 is a plane, which is a special case for a hyperboloid. Thus, the solution set is in a plane that is perpendicular to the detector plane 1-2-3.

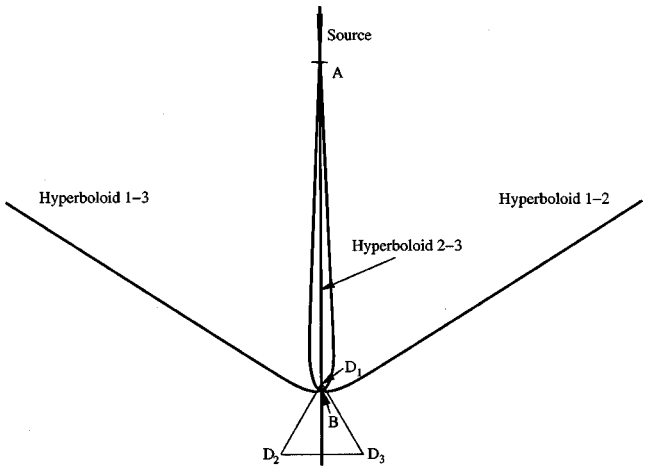


FIG. 3. From three detectors, all possible ITD measurements determine a solution curve C_{123} in the 3D space, which is the intersection of any two hyperboloids. In the above situation, the plane in which the curve C_{123} lies is orthogonal to the page.

Suppose that we add another sound detector D_4 above the detector plane 1-2-3 (e.g., above D_1 when the page is horizontally placed). Note that D_4 cannot be coplanar with the other three noncollinear detectors to avoid degeneracy. The detectors D_1 and D_4 result in a single sheet of hyperboloid 1–4. The intersection of the curve C_{123} and the single sheet of hyperboloid 1–4 gives two 3D points, one is the true sound source and the other is a spurious sound source. If the true sound source *A* in Fig. 3 is coplanar with D_1, D_2, and D_3, the spurious sound source is point *B* in Fig. 3. From Fig. 3, we can see that typically one of the solution points is outside (or far from) the detector array and the other solution point is inside (or near) the detector array. We can view geometrically why there are only two solution points. In Fig. 3, suppose the ITDs between detectors 1 and 4 are zero (by, e.g., adjusting the position of D_4 slightly); the one sheet of the hyperboloid 1–4 degenerates into a plane that goes through source *A* and the mid-point of the line segment connecting D_1 and D_4. The intersection of the planar hyperboloid 1–4 with the closed curve C_{123} gives two points. We cannot expect that other hyperboloids formed by D_4 with other sensors can reduce the multiplicity of the solutions, as we saw in the three-detector case. Thus, we have the following observation:

Observation 2: All the ITD measures from four noncoplanar sound detectors generally define two possible solutions for the 3D location of the sound source.

We have used a geometric method to reach this result. In fact, a direct algebraic proof of this uniqueness involves intersection of three quadrics in 3D and as far as we know there has been no known general closed-form solution for such a problem. Even if a closed-form were available there would probably be no sufficient space in this article to fully present it.

2. Using ILD only

We first determine the solution set from a single ILD measurement. Without loss of generality, we position the coordinate system and scale its axes so that the detectors D_1 and D_2 are at (1, 0, 0) and (−1, 0, 0), respectively. Since ILD is proportional to r_2^2/r_1^2, let

$$\frac{r_2^2}{r_1^2}=\frac{(x-1)^2+y^2+z^2}{(x+1)^2+y^2+z^2}=\alpha,$$

where $\alpha>0$, with an exception $\alpha=1$ when the surface becomes a plane $x=0$. The above equation can be rewritten as

$$(x-c)^2+y^2+z^2=r^2,$$

where

$$c=\frac{1+\alpha}{1-\alpha} \quad \text{and} \quad r^2=\left(\frac{1+\alpha}{1-\alpha}\right)^2-1.$$

In other words, given an ILD, the solution set is a sphere centered at $(c,0,0)$ with a radius *r*, as shown in Fig. 4. Using three ILDs from three detectors, the solution set is the intersection of two spheres, sphere 1–2 and sphere 1–3, which

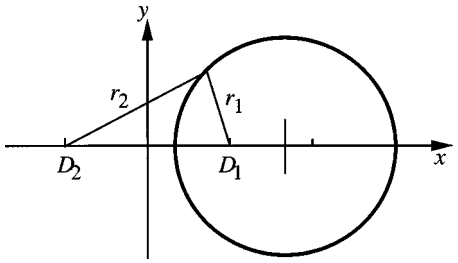


FIG. 4. From two detectors, a given ILD determines a sphere.

gives a circle as shown in Fig. 5. Therefore, we have the following observation.

Observation 3: All the ILD measures from three noncollinear sound detectors generally define an infinite number of solutions for the 3D location of a sound source and the solution set is a 3D circle.

Suppose that a fourth detector D_4 is added. The sphere 1–4 determined by detectors 1 and 4 intersects the circular solution set from detectors 1, 2, and 3 at two points, one being the true location of the sound source and the other being the spurious one. As we can see, the fourth detector D_4 should not be coplanar with the first three detectors. Otherwise, it does not reduce the degrees of freedom of the solution. This leads to the following observation:

Observation 4: All the ILD measures from four noncoplanar sound detectors generally define two possible solutions for the 3D location of the sound source.

3. Using ITD and ILD jointly

Suppose that both ITDs and ILDs from a four noncoplanar sound detector array are used, as shown in Fig. 6. The spurious solution *B* from the ITDs is typically not the same as the spurious one *C* from the ILDs, as illustrated in Fig. 6. Therefore, we have the following observation:

Observation 5: All the ITD and ILD measures from four non-coplanar sound detectors generally define a unique solution for the 3D location of the sound source.

A question remains as to whether ITD and ILD measures from three noncollinear detectors are sufficient to give a unique solution. The answer is no. Consider the plane *P* in which the three detectors lie. Starting with the true 3D position of the sound source *A*, consider mirroring *A* with respect to the plane *P* as a mirror produces a superfluous point A' on the opposite side of the mirror. This superfluous point A' gives exactly the same ITD and ILD measures as *A*. Thus, we have the following observation:

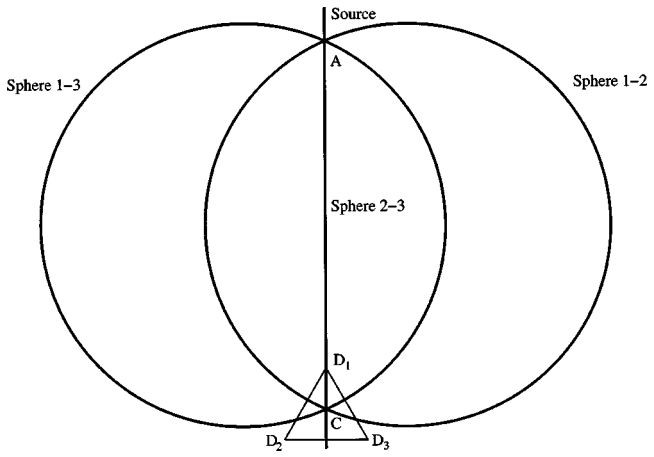


FIG. 5. From three detectors, all possible ILD measurements determine a solution circle in the 3D space, which is the intersection of any two spheres. In the above situation, the plane in which the solution circle lies is orthogonal to the page.

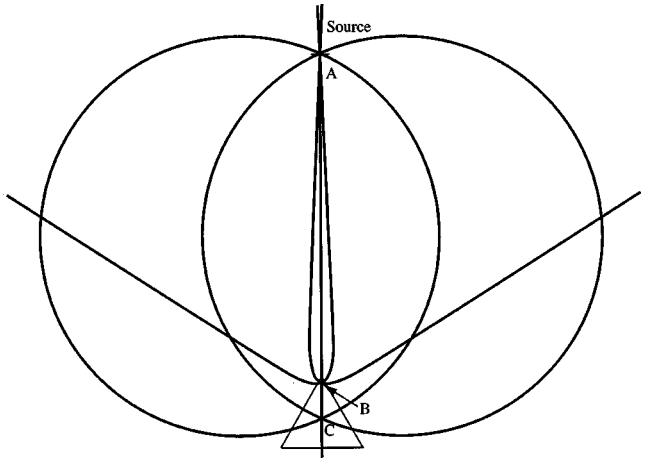


FIG. 6. When ITD and ILD measurements are both used.

Observation 6: All the ITD and ILD measures from any three sound detectors do not define a unique solution for the 3D location of the sound source.

The above reasoning is also applicable to any number of coplanar sensor arrays. By similar reasoning, using the plane in which the sensors lie as the mirror plane, we get the following observation.

Observation 7: All the ITD and ILD measures from any number of coplanar sound detectors do not define a unique solution for the 3D location of the sound source.

We have reached our conclusions about uniqueness using geometric reasoning, which is more intuitive, more informative, and easier to understand for our geometric problem than a purely algebraic approach. We will further see these advantages of geometric reasoning in the following section when we investigate the stability of these solutions. We have tried an algebraic approach to reach a solution, but we were not successful partially due to the fact that a closed form algebraic solution for the intersection of general quadrics does not exist. In other words, we can geometrically reason about these solutions but we cannot express the solutions in an algebraically closed form. As far as we know, our uniqueness analysis presented here is the first complete one for 3D that covers both ITD and ILD.

C. Solution stability

The previous discussion about how the solutions are determined is also very useful for providing insight into the stability of the solution from ITD and ILD features.

We are interested in determining the relative 3D position of the sound source from an array of non-coplanar microphones. The relative position can be determined by the displacement vector from the center of the detector array to the position of the sound source. This displacement vector can be specified by the 3D *direction* of the vector and the length

of the vector. The direction tells the angle and the length tells the *distance* from the detector array to the sound source.

In Fig. 3, we can see that the 3D location of the sound source is determined by three surfaces. With measurement errors, the position of the three surfaces has also a certain amount of error. We can imagine that the amount of error in the position of each surface is represented by the thickness of the surface. In other words, the hyperboloid surfaces in Fig. 3 have a certain degree of thickness. Further, the thickness is not constant for each hyperboloid. The farther the surface patch is away from the detector array, the thicker the part of the surface is. As shown in Fig. 3, the 3D location of the sound source *A* is determined by the intersection of three surfaces that are almost parallel and vertical (assuming that the diagram is a top view when the page is horizontal). Therefore, ITDs give a bulletlike uncertainty region with the long axis of the bullet pointing toward the center of the detector array, as shown in Fig. 3. In other words, the distance in the solution is less reliable than that for direction.

A similar reliability analysis can also be applied to ILD. In Fig. 5, we can see that the 3D location of the sound source is determined by the intersection of three spheres. If only sphere 1–2 and sphere 1–3 are used, the uncertainty region after their intersection will be like a disk, whose rotational axis is aligned with the line of sight from the center of the detector array to the sound source. The additional intersection with sphere 2–3 will probably reduce the width of this uncertainty disk by a minor amount, because the thickness of the sphere 2–3 is large at this long distance. Thus, ILDs give a disklike uncertainty region with the normal of the disk pointing toward the center of the detector array. Comparing the bullet and disk shapes of the uncertainty regions of ITDs and ILDs, respectively, it appears that ITD is relatively better for direction estimate and ILD is relatively better for distance estimate. This observation has been confirmed by our experimental data.

Since each ITD or ILD is separately estimated by comparing the sound signals from two microphones, each of ITD and ILD from every pair of detectors is potentially useful in providing overdetermination for combating noise. This is why we will use all the possible ITD and ILD measures in the estimation of 3D sound location in our system.

D. Number and placement of detectors

From the previous analysis, we know that if four noncoplanar detectors are used, the intersection is unique. A mobile robot requires that the structure of the sensor array be compact, while accuracy consideration requires a large array. Thus, an equidistant structure seems reasonable. In the case of four sensors this suggests a tetrahedron (Fig. 7). In our experiment, an equal-side tetrahedron with a 20-cm side was used. At each apex of the tetrahedron is a miniature microphone. Light carton paper was used to support the microphone array, without causing significant sound absorption.

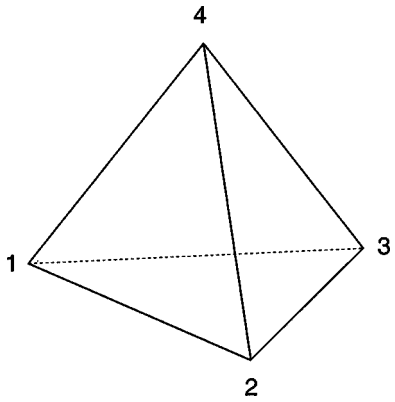


FIG. 7. Placement of the four microphones on the array.

III. METHOD

As outlined previously, efficient sound localization is possible after having extracted the necessary ITD and ILD measures. We have shown in Sec. II B that the minimum number of detectors required to obtain unambiguously a solution in 3D space is four and that it is unique in general. In order to fully solve the sound localization problem, two main steps need to be performed. Step 1, extract ITD and ILD measures from the acquired signal. Step 2, estimate the actual sound location using those feature measures.

A. Feature extraction

As discussed, the two features considered in this work are ITD and ILD. For successfully detecting ITD, we like to avoid narrow-band acoustic signal such as a tone of a single frequency, since such signals tend to be nearly periodic in time. ITD will have a severe ambiguity problem when it is computed from a periodic signal. In practice, we use a human voiced sentence as the sound source, since it is typically the case for our robot applications. A microphone signal from such human verbal sounds contains low frequencies as well as higher-order harmonics and thus has a relatively wide frequency band.

Using the appropriate hardware, the acoustic signal can be first converted to an electrical signal (by microphones) and then to a digital form (analog-to-digital converter board). The digitized data is a sequence of values representing the temporal waves of the sound, as picked by the respective detector for a determined period of time. A window of sufficient duration is used to define the searchable domain. Some preprocessing is applied to ensure satisfactory quality of the sound sample. For instance, the amplitude of the signal's envelope can vary over time. With speech this corresponds to, e.g., accents and pauses within and between words. These sound blanks contain little useful information and using them can degrade the quality of the estimates. In order to avoid this problem, the window is divided into smaller intervals in which the variation of the signal is evaluated (Fig. 8) This preprocessing selects only signals with high variance for feature extraction. A measure of the ''efficiency'' of the sample is returned by the procedure as the percentage of used subframes in the whole window.

The next phase involves the use of a cross-correlation

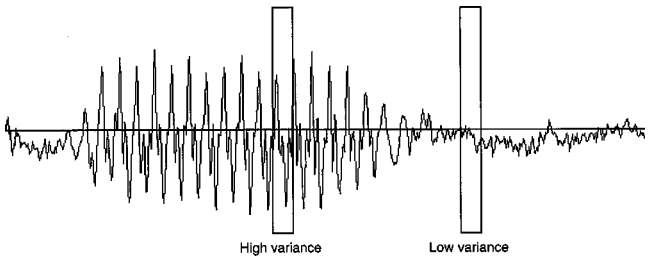


FIG. 8. Preselection according to signal variance.

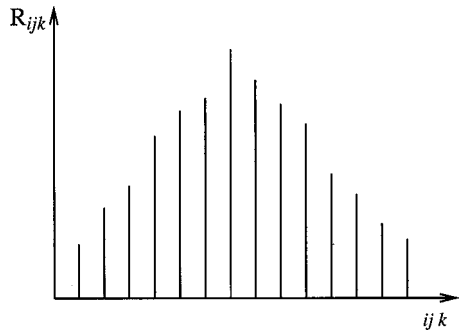


FIG. 9. Typical cross-correlation curve with good sharpness.

procedure to determine the shift between the sampled signals at each of the sensor couples. This gives a direct measure for the ITD.[14] We find the peak of the cross-correlation function varying across a range of possible time delays. Suppose that $\{x_{t_0}, x_{t_0+1}, \ldots, x_{t_0+N+n-1}\}$ is the digital signal from the first microphone channel and $\{y_{t_0}, y_{t_0+1}, \ldots, y_{t_0+N+n-1}\}$ from the other channel, where N is the length of the segments, in number of samples, used for correlation and n is half of the maximum possible time delay. We can define the normalized cross correlation at time t_0 with shift variable for the first microphone channel j and shift variable for the other channel k as

$$R_{j,k} = \frac{\sum_{i=0}^{N-1}(x_{i+j}-\overline{x_j})(y_{i+k}-\overline{y_k})}{\sqrt{\sum_{i=0}^{N-1}(x_{i+j}-\overline{x_j})^2}\sqrt{\sum_{i=0}^{N-1}(y_{i+k}-\overline{y_k})^2}}, \tag{3}$$

where $\overline{x_j} = \sum_{i=0}^{N-1} x_{i+j}/N$ and $\overline{y_k} = \sum_{i=0}^{N-1} y_{i+k}/N$. Using Eq. (3) we can find the maximum correlation by successively trying shifts in each direction: R_x to cover the case when the source is closer to the first sensor and R_y when the source is closer to the other sensor:

$$R_x = \max\{R_{j,k} | j = 0, 0 \leqslant k < n\}, \tag{4}$$

$$R_y = \max\{R_{j,k} | k = 0, 0 \leqslant j < n\}. \tag{5}$$

Then, $R_{\max}$ is the maximum cross-correlation:

$$R_{\max} = \max\{R_x, R_y\}. \tag{6}$$

The sign of the time-shift T, is determined by how R_x and R_y in (6) compare in magnitude. The absolute value of T is proportional to the value of j or k that maximizes the value in Eqs. (4) or (5), depending on the sign of T. More nearly precisely it is the product of that number and the digitization interval (the inverse of the sampling frequency). The value at the maximum is selected and is returned along with a parameter reflecting the sharpness of the correlation curve at the peak (Fig. 9). A combination of those two parameters is used as a quality estimate for the ITD (''score''). A high value of $R_{\max}$ is an indication of good similarity between the signals from both channels. However, a high value of $R_{\max}$ does not mean that the ITD is reliable. The second parameter, however, discriminates between a wide, flat correlation curve and a narrow, sharp one. The latter tends to give a more nearly accurate ITD.

We should note a very useful side effect, which is closely related to the precedence effect in human auditory perception.[4,5] In case there are reflections of the incoming sound from incident surfaces (echoes, reverberation of broad spectrum sounds[21]) a secondary maximum will appear in the correlation curve. However, with the presented approach, it will be ignored because the similarity of that signal will be significantly lower and thus it will correlate less well (lower and wider peak).

Using the obtained information for the ITD, it is possible to evaluate the ILD by computing an integral value of the signal energy from the shift-adjusted signal. To estimate the signal energy we can use the variation in time of the voltage output of each microphone $S(t)$. The value for microphone pair 1 and 2 is shown in Eq. (7):

$$\mathrm{ILD}_{12} = \frac{\int_{t_i}^{t_j} |S_1(t)|\, dt}{\int_{t_i}^{t_j} |S_2(t+T)|\, dt}, \tag{7}$$

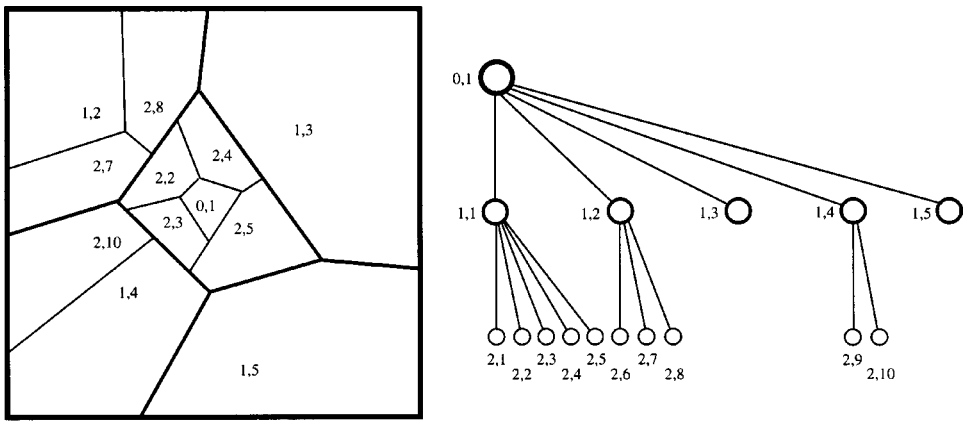
where $S_1(t)$ is a measure of the signal picked by microphone 1 and $S_2(t+T)$ by microphone 2, T is the previously determined time shift, and $t_j - t_i$ is the length of the sample window. We should note that the sample window is chosen such that it is much larger than the maximum possible time difference between any pair of microphones (500 ms vs 0.5 ms).

The estimates for ITD and ILD are considered reliable only if the efficiency and score of the sample are satisfactory, i.e., above a predefined threshold. Thus the described procedure extracts not only the needed features but also information about whether the feature can be used for localization or should be discarded as useless.

B. Tree-based learning

Once the ITD and ILD are extracted from the signal picked up by the detector array, the next step is to perform the actual sound source localization. The previously discussed disadvantages of the currently available methods can be avoided to a large extent by taking a learning-based approach. As stressed above in Sec. II B, these features uniquely define a solution and thus we have a direct correspondence between the extracted ITD and ILD values and the 3D coordinates of the sound source. Of course, in the presence of distortions, echoes and noise of various natures, all the ITD and ILD measures are not consistent and we like to use overdetermination to combat those effects that have not been modeled by our theoretical exposition.

In the current case the input feature space X is 12-dimensional: 6 for ITD and 6 for ILD (one for each combination of detector pairs). The output space Y is three-



(a) (b)

FIG. 10. A 2D illustration of a hierarchical Dirichlet partition and the corresponding recursive partition tree (RPT). (a) The partition, where the label indicates the center of a cell. The label of the child to which its parent's center belongs is not shown due to the lack of space. (b) The corresponding recursive partition tree.

dimensional: (y_a, y_e, y_r) representing azimuth angle, elevation angle, and radial distance, respectively. This polar representation in output space is used because we know from the above reliability analysis that the uncertainty regions for ITD and ILD are elongated in the radial direction or perpendicular to it. The polar representation can better isolate the errors in the resulting components. Thus the mapping to be computed is a mapping $f: X \mapsto Y$, where X is the 12-dimensional input space and Y is the three-dimensional output space.

We should note the extreme complexity of this actual mapping. The closed-form solution has not been found even in an ideal situation (e.g., no echo is considered). The existing methods have used an approximate closed-form solution or an iterative search procedure to minimize any nonlinear objective function. The former type suffers from the model error and the latter type either requires time for an exhaustive search or does not give a guarantee of global minima in the nonlinear objective function to be minimized (the practical situations may vary). Further, in a real situation, the model used should take into account the working environment, such as the material of the walls and objects in the environment as well as their acoustic properties. Such a complete model is extremely difficult to construct and even if it is constructed, the generality of the method is very limited since the model is only applicable to a particular known environment.

We use a learning based method by constructing a regression tree from a set of training samples $t=\{(x_i, y_i) | y_i = f(x_i), x_i \in X, y_i \in Y, i=1,2,...,n\}$. Although the system trained in one environment can be, in principle, only used in the same environment, our method can be used to train the system in virtually any environment. If the accuracy requirement is not very high, the system can be trained using a training set that uses training data collected in a variety of environments.

If the output space Y consists of a list of discrete class labels $\{l_1, l_2, \ldots, l_n\}$, the problem of constructing f is called a classification problem. If the output space Y is a numerical space, the problem is called a regression problem. Therefore, our problem here is a regression problem. Further, the number of samples n is typically large, which is an important issue for real time application. After input vector x is computed, the corresponding estimated output vector $y' \approx y = f(x)$ must be computed very quickly. Thus, it is not possible to search all the possible training pairs in L. The goal of our regression tree method is to organize the training samples by a tree data structure so that top $k>0$ best matched samples (x_i, y_i) in the training set L can be found quickly without having to examine all the training samples in L.

We present our SHOSLIF recursive partition tree (RPT) which we used to approximate the mapping f. Figure 10 illustrates how the training samples are organized by the RPT. The input space X is depicted as a 2D space in Fig. 10(a), where the location of each number pair i,j represents the x-part of a training sample (x,y). The number i corresponds to level in the tree shown in Fig. 10(b) and j denotes the index of the nodes at this level. The entire space X is represented by the root of the tree. The children of the root further divide the space of the root into smaller cells. The grandchildren of each node further divide the space of its parent into smaller cells. Such recursive division is carried on until the resulting cell contains only one training sample and the corresponding node is called a leaf node. In the experiment, we set the number of children to be two, and thus the tree used for the experiment was a binary tree in that each nonleaf node has exactly two children (a special case of that in Fig. 10).

How do we divide the cell of each node? We certainly do not want to divide the space in such a way that one node has all the samples in the parent and the other has no samples at all. Better, we want to divide the cell across the direction in which the samples x_i's spread wide. This is important since the input samples might lie in only a subspace X' of the input space X and the dimensionality of X' may be significantly smaller than that of X. Suppose $S = \{x_1, x_2, \ldots, x_m\}$ are the samples in a node A (A is the root of the tree to start with). We compute the first principle component vector $v \in X$ from the samples in S, which is the eigenvector of the sample covariance matrix of S corresponding to the largest eigenvalue.[22] In other words, along vector v, the samples in S have the greatest variance than any other direction. Then, we compute the projection a_i of x_i onto v. $a_i = \langle x_i, v \rangle$ where $\langle \cdot, \cdot \rangle$ denotes the vector inner product. The mean of a_i, $\bar{a} = \sum_{i=1}^{m} a_i / m$, is used to compute the position of the partition hyperplane along vector v. Every sample in S

is assigned to one of the two children in the following way: For each x_i, if its projection a_i is less than $\overline{a}$, x_i is assigned to the left child. Otherwise, x_i is assigned to the right child. In other words, a hyperplane whose normal is v and whose position is at the centroid of the projections of S along v, divides the samples in S to the two children. Such a division process for the root results in two children. The samples assigned to each child are further subdivided, recursively, until the resulting child has only one sample.

After the RPT has been constructed, it is used as follows. Given a measurement vector $x \in X$, we use the RPT to estimate the approximate output $y' \approx y = f(x)$ by querying the RPT. Starting from the root, decide to which child x belongs using the way in which the samples are assigned to the children. Only that child is further explored recursively. Finally, a leaf node will be reached which stores one sample (x_i, y_i). y_i can be used as the estimate for $f(x)$. It is worth noting that this tree retrieval process may not always give the best matched x_i for all possible x. Thus, instead of exploring one path down the tree, we explore $k>1$ parallel passes by keeping the best k nodes at each tree level for further exploration. Finally, k leaf nodes have been reached. y_i vectors of these top-k matched leaf nodes are used to give an interpolated output vector $y \in Y$, given output vector $x \in X$:

$$y = \frac{1}{\sum_{i=1}^{k} w_i} \sum_{i=1}^{k} w_i y_i, \tag{8}$$

where w_i is the weighting function and y_i is the output part of the ith nearest neighbor of x_i. The weights are determined in such a way that the nearest neighbor has the highest weight while others have less weight, depending on each distance from x:

$$w_i = \alpha^{-\|x - x_i\|/(\|x - x_0\| + \epsilon)}, \tag{9}$$

where x_0 is the nearest x_i from x, $\alpha > 1$ is a decreasing factor, and $\epsilon > 0$ is a small constant to prevent the denominator from going to zero. For example, when the distance $\|x - x_i\|$ is twice as large as that from the nearest neighbor, $\|x - x_0\|$, the corresponding weight is decreased by a factor α^{-1}. This way, the RPT can give a reasonably interpolated output from k near samples even if the nearest neighbor was not among the k leaf nodes provided by the RPT. The time complexity of computing the estimated output from the RPT is roughly the number of levels in the tree, which is roughly $O(k \log_2(n)) = O(k \log_2(n))$ where n is the number of samples in the entire training set T. For more detail about SHOSLIF RPT and its performance advantages over other general function approximators, such as the feedforward neural networks and the radial function networks, the reader is referred to Ref. 20.

IV. EXPERIMENTS

In order to test the methodology, an experimental setup was used to perform a number of tests. A set of four identical Lavalier microphones was placed at the tips of a solid tetrahedron with a 20 cm side (Fig. 7). The signal from the microphones was amplified by four modular microphone preamplifiers to bring the power of the signal level to a required range. It was then supplied to an analog-to-digital converter board mounted in a personal computer. The software was designed to visualize, train, recognize, and run various sound localization related tasks. Samples were taken from various points with known 3D coordinates, some were used for training and others for testing. The results were compared with linear search and the performance of SHOSLIF was evaluated.

FIG. 11. The experimental setup. The compact sensor array is shown at the top of the pole standing on the edge of the center table. The intersection points on the grid on the floor mark the control points in the work space.

A. Experimental environment

The dedicated hardware was built from off-the-shelf consumer and industrial quality items. All experiments were held in the Pattern Recognition and Image Processing laboratory in the Department of Computer Science at Michigan State University, which is hardly suitable for high precision acoustic experiments. As shown in Fig. 11, the test space was located in the middle of the laboratory, in between cubicles with computers and reflecting and absorbing surfaces of irregular shape. The number of devices producing weak to strong noise of different frequencies and levels was above 20. Often laboratory members would speak softly in the background while samples were being captured for training or retrieval. A room with better acoustic properties (e.g., anechoic chamber) is not consistent with our goal. The laboratory environment such as the one above was close to the one in which our actual sound localization device would be exposed to in our intended applications.

B. Experimental setup

At the training stage a continuous sound, originally produced by a human speaker uttering a short sentence, was reproduced using a hand-held tape recorder, from a set of previously defined locations (Fig. 12). Without significant loss of generality, the span of the training grid was set to an arbitrary section of 3 m by 3 m by 2.1 with the microphone array in the middle of one of the sides. The density is linear in Cartesian coordinates with a granularity of 0.3 m. However, only 237 of the defined 700 points were used for training. They were selected to simulate uniform angular density and ten samples were taken from each of those points. The approximate angular density of the training points was

FIG. 12. Training the system. The compact sensor array is shown at the upper left corner of the figure. A tape recorder plays the sentence at a predefined height right above a floor grid point. A weight is used to guarantee vertical displacement between the grid point and the tape recorder.

around 15 degrees. Thus the angular span of the training area was about 180 degrees in azimuth and a little less than that in elevation.

At the performance test stage, we used the same sound source. Other voices were also used to investigate the variation of accuracy. We were able to produce a similar accuracy using sentences of similar style of speaking. We have also observed that the quality of the sound directly influences the reliability and thus the accuracy of location estimate. The quality includes, e.g., ratio of sounds over pauses and the compactness of the sound source. In reality, the sound source is not exactly a point source.

The system can be operated to collect 0.5-s samples and provide location estimates based on each of them or to store them for further analysis. The test samples were obtained from 79 positions on the grid, but their locations were different from the ones used for training. They were used for computing estimation errors offline using another specially designed evaluation program. Estimates for the location of each of the test points were thus produced and recorded. They were compared to the known actual values of the 3D coordinates and the error was computed as the difference between the actual and estimated value for the angles, and the ratio of that difference to the actual value, for the radial distance.

The algorithm used to compute the coordinates of the sound source uses two parameters for fine tuning its performance. One is the relative weight of the two input arguments: ITD and ILD. Because of the greater magnitude of the ITD, the ILD was multiplied by a variable factor, called *scaling on ILD*, thus increasing its relative weight as needed. This allows us to estimate the relative significance of those two parameters on the accuracy of the final results. A low value of this parameter would mean neglecting the ILD (a value of zero means only ITD are used), while a very high value indicates a predominance of the ILD. Their relative weight is practically equal when the value of *scaling on ILD* is around 13. The other parameter is the weight coefficient α used in the interpolation of the retrieved nearest neighbors as defined in Eq. (9). A low value of α would indicate that all considered nearest neighbors are almost equally weighted when estimating the solution (for $\alpha=1$ we have averaging) while a big value of alpha emphasizes the role of the nearest neighbor.

It is known that ITD and ILD are frequency dependent, e.g., ITD uses predominantly the low frequencies, while higher frequencies are the only ones that can be used for estimating the ILD. A preliminary signal filtering can be employed to leave only the useful frequencies when determining each of those two parameters. The actual response of those two filters can be another subject for fine tuning. However, the real-time implementation requirements for this project impose serious limitations on the amount of preprocessing that can be performed and thus spectral analysis of the signal is not adopted.

C. Results

The results obtained with our experimental setup were used to study the above-mentioned relations. A number of plots are used to show actual values and some observed trends. Figure 13 shows how the relative weighing between ITD and ILD affects the accuracy of estimation of the azimuth, Fig. 15 for the elevation and Fig. 17 for the distance. The respective standard deviations are shown in Figs. 14, 16, and 18, respectively. The horizontal axes are the scaling on ILD—the coefficient by which ILD is multiplied when considered for estimating the nearest neighbors and the coefficient α [see Eqs. (8) and (9)]. The range on the axes was chosen equal for all the plots for compatibility. The direction of the axis for the scaling on ILD was inverted in Figs. 17 and 18 so the surface can face the viewer. The distance plot shows a trend of a descending error value but its further extent was not followed because the standard deviation is increasing for those same values, rendering the low error values unreliable. In these trials a number of KNN$=7$ (nearest neighbors) was used. The values of ILD are theoretically unbounded, hence it is impossible to get a correct number for the balance of relative weights of ITD and ILD, but an empirically estimated value of scaling on ILD of around 13, for which their weight is approximately equal, was found.

From these data plots we can see how the direction (angular) error is low when the relative importance of ITD is high (scaling on ILD is low). The minimum, however, is registered at a nonzero, but nearly zero, value of scaling on ILD as shown in Figs. 13 and 15. This means that ILD is useful for reducing angular error, but not much. The contribution of ILD in reducing the error in distance is, however, very significant, as shown in Fig. 17. This confirms our error analysis in Sec. II C. It becomes clear from those observed trends that when it is necessary to estimate both direction and distance to the sound source, both ITD and ILD should be taken into account. An appropriate amount of contribution

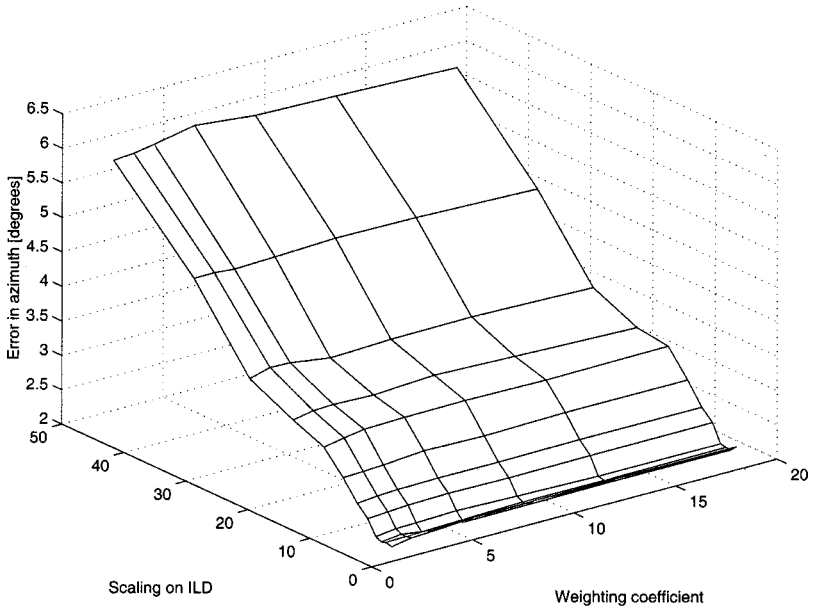


FIG. 13. Distribution of error values for azimuth.

from ILD could result in lower error in both direction and distance estimates, especially for distance.

The best precision measured for points located within the sector designated for training, but between the grid points used for training, was estimated at around ± 2.2 degrees in azimuth, ± 2.8 degrees in elevation, and $\pm 19\%$ in distance. The super-resolution is achieved by the KNN interpolation in the RPT. The relatively lower accuracy in distance is expected for such source-sensor placement. It should also be noted that the specific task the system was tested for—indicating the actual location of the sound—is very difficult for humans in such situations, as well, without the help of other sensing modalities such as vision.

A sample set of error values used to produce a single point in the average errors plots is presented in Table I.

D. RPT performance and speed

To verify the performance of the RPT and its speed, a deterministic linear search algorithm was implemented to find the nearest neighbor in input space from the set of training samples L. The results obtained in such a linear search method were used to compare the performance of the SHOSLIF procedure. The speed was confirmed to be considerably faster with SHOSLIF. As timed on the test PC, a single retrieval from the tree, with 2370 test samples, took 2.5 ms on average for SHOSLIF, versus 15 ms with the linear search (see Table II). The accuracy was comparable to that of the linear search. The timing for the preprocessing indicated an average of 230 ms which, although being considerably longer than the retrieval time, is still shorter than the signal scan time of 500 ms (single window). This situation indicates that for the current moderate number of training samples collected within the 3D work space tested, preprocessing takes far more time (230 ms) than the retrieval (2.5 ms). In other words, the RPT algorithm, as implemented, is twice as fast as needed for real-time application.

From the relative time difference between preprocessing and retrieval, one question is in order here. Is it true then that

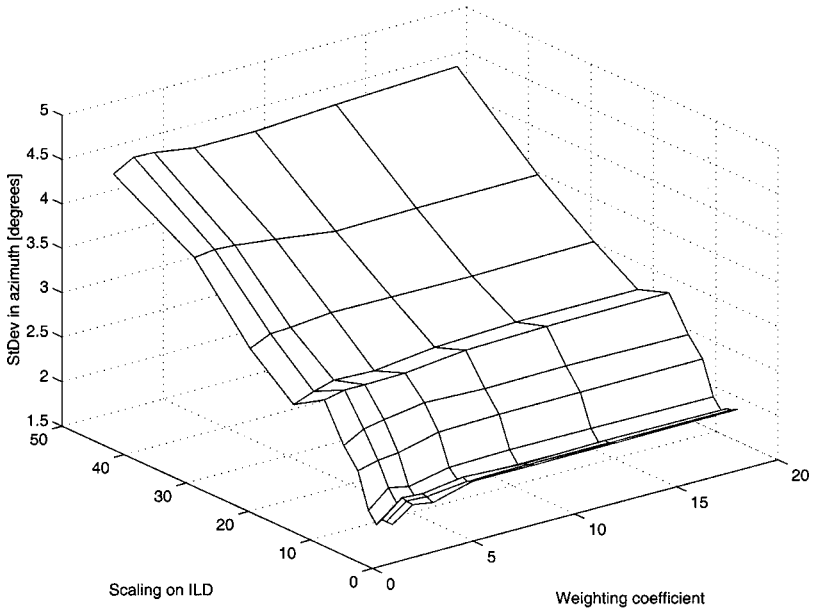


FIG. 14. Distribution of standard deviation for azimuth.

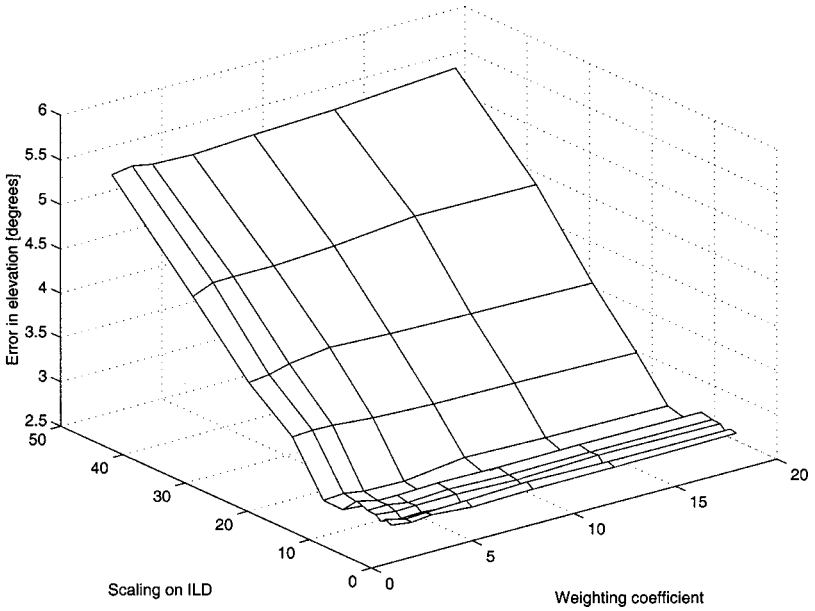


FIG. 15. Distribution of error values for elevation.

the time gain of SHOSLIF is not important? This is in fact not the case. Consider that the problem becomes larger, when the number of samples is increased greatly. Such a problem arises when, e.g., the work space to be covered is increased, the sample density is increased for better accuracy, or the number of environmental settings is increased for handling more settings. The preprocessing time is somewhat constant, but the retrieval time is not when the number of samples increases. SHOSLIF is able to handle a much larger number of samples without slowing down significantly, due to its logarithmic time complexity. But a linear search method cannot. In other words, SHOSLIF method can ''scale up'' to larger problems but a linear search cannot.

A graphical user interface has been developed for the program which allows the user to pass all necessary parameters to the program, to select the various options, as well as to view the waveform of the scanned signals.

E. Implementation restrictions

As mentioned before, the system performed well despite different unfavorable factors, like background noise, reflections, unreliable sound sources, etc. However, it should be noted that although no exact measurements have been performed, these and some other factors would influence its reliability and accuracy depending on their strength. In most experiments the acoustic noise was kept at a S/N ratio of around 20 dB (as estimated visually from the displayed waveform), but in real life situations the noise can be as strong as the signal or even stronger. Another problem would be multiple sound sources. In the case of signal reflection, the intensity of the reflected signal would be significantly weaker and thus it will be ignored by the preprocessing routine. However, with secondary sources, the intensity of the sources can be comparable and this might lead to jumping

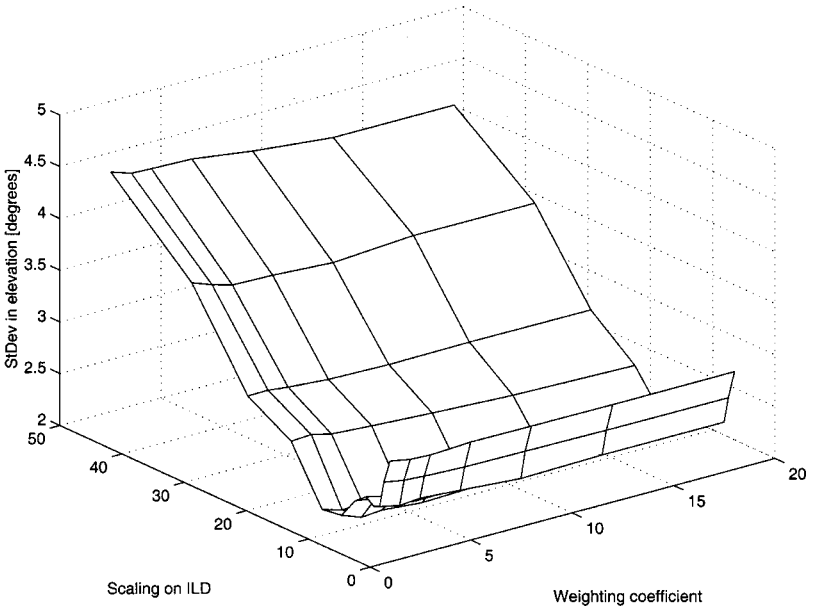


FIG. 16. Distribution of standard deviation for elevation.

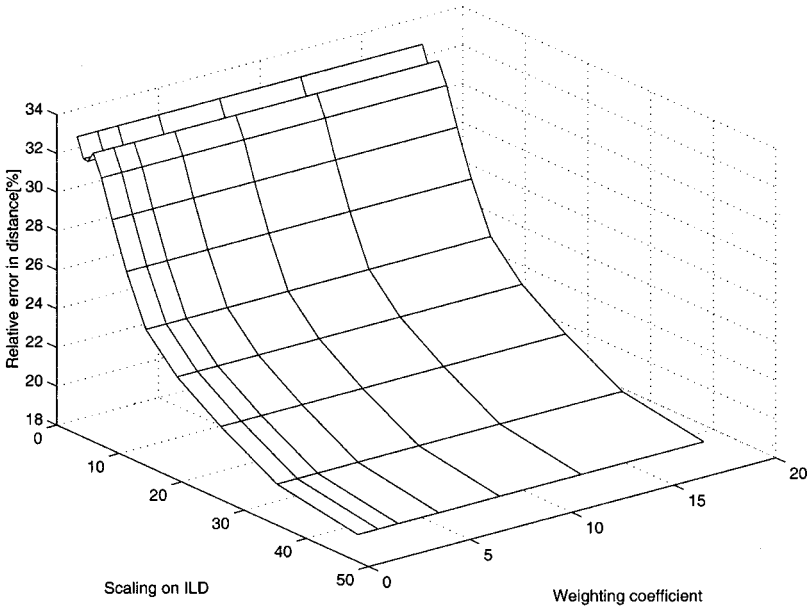


FIG. 17. Distribution of error values for distance. Note that the axis direction for ILD is inverted for better viewing.

between the two sources and even complete wash-out of the correlation curve and thus incorrect localization.

Most of the experiments were performed with a sound source steadily fixed in space. A moving source would present a challenge for the current implementation. With the current windowing approach, a source movement would be similar to having a source of a larger spatial size (aperture), which would produce a lower signal correlation. The performance with a shortened window has not been studied extensively at this point. In a similar way, an influence on the accuracy of detection was observed when varying the size of the aperture of the sound source. For instance, sounds produced with a wide open mouth would yield a higher error value. An accurate study of this relation needs to be performed in order to determine the correct way of compensating for the increase in source size.

One of the typical disadvantages that training presents is the difficulty for a learning system to perform in untrained environments, compared to the environment in which it was originally trained. The sensitivity to environmental changes has not been quantified yet.

V. CONCLUSIONS

We analyzed the uniqueness and reliability of 3D sound localization from ITD and ILD. The extraction of ITD and ILD is based on a fast and efficient algorithm that is capable of not only computing those parameters with satisfactory accuracy but also provides a useful means of evaluating the usability of the taken sample. The three-dimensional localization is performed by a learning technique. The applicability of the proposed implementation is more general than the majority of the currently available solutions in that various features can be used without the need to explicitly model the relationship between feature values and the 3D location es-

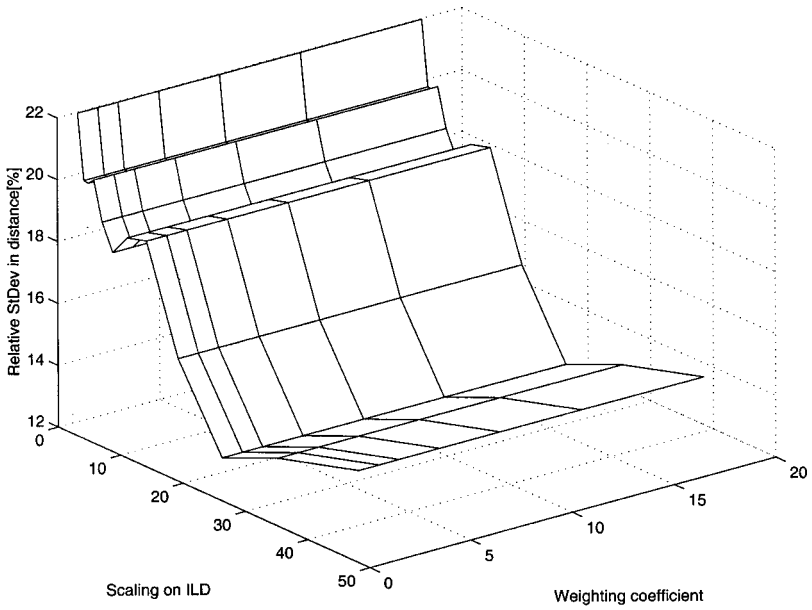


FIG. 18. Distribution of standard deviation for distance.

TABLE I. Error values for scaling on ILD=11 and weighting coefficient=1, A=azimuth [°], E=elevation [°], D=distance [%].

A	E	D	A	E	D	A	E	D
2.0	6.0	17.45	1.0	1.0	44.25	1.1	0.3	22.16
1.1	6.4	7.70	1.7	5.9	38.35	1.4	1.0	19.54
1.0	1.0	39.46	3.0	2.0	34.02	4.9	0.6	6.95
12.0	0.0	0.00	1.0	4.0	40.24	3.0	3.0	29.82
4.0	2.0	11.82	1.0	4.0	40.24	8.9	3.3	31.68
0.0	10.0	25.37	1.1	8.0	14.55	0.0	2.1	22.08
2.6	3.3	17.22	2.0	2.0	19.14	7.0	2.0	56.67
0.0	0.7	25.35	6.7	0.9	17.04	9.0	1.0	45.56
3.0	4.0	23.95	5.0	2.0	19.79	2.0	3.0	30.86
3.0	4.0	23.95	1.7	3.9	26.79	6.6	4.0	29.94
0.0	1.1	24.45	5.7	1.1	7.72	3.0	2.0	41.13
3.0	0.9	14.07	0.9	4.4	2.83	2.0	3.0	30.86
1.4	2.9	11.51	7.7	2.9	16.54	0.0	0.0	3.89
1.3	0.4	11.64	5.3	3.1	1.98	0.0	0.0	50.34
0.1	2.7	10.45	3.1	1.1	9.45	6.4	5.9	0.61
11.0	3.0	32.09	7.0	0.0	8.99	5.4	1.3	22.39
0.0	5.3	3.38	2.4	5.6	7.70	0.7	0.9	4.52
9.0	1.0	45.56	3.6	3.1	7.05	3.0	1.0	39.36
2.0	9.0	37.70	0.0	5.1	25.36	0.9	4.6	13.76
0.3	5.1	51.18	3.0	0.0	26.72	2.4	2.3	13.80
3.0	2.0	41.13	5.1	8.4	17.51	0.7	4.4	13.60
0.3	1.7	22.14	0.1	1.0	4.81	1.0	4.0	23.59
9.0	4.1	57.37	2.3	2.3	6.60	4.0	0.3	14.68
1.1	2.4	69.98	0.9	3.3	46.99	6.0	3.0	12.55
1.1	4.5	19.53	8.0	3.0	22.56	4.7	2.0	7.40
1.0	1.0	44.25	1.4	0.3	1.99	4.6	1.1	12.31

timates. The method needs to store a large number of samples (over-learning is avoided by SHOSLIF by only storing samples that are necessary). In order to achieve good accuracy the training density needs to be close to the expected resolution. This can lead to the need of taking samples from hundreds of three-dimensional locations, and, to ensure stability, several samples from each point need to be taken. However, the logarithmic retrieval enables the system to easily reach real-time response speed.

An originality of this work is in the versatility of its application domain. First the lack of spatial constraints allows for a wide range of applications. The use of a compact sensor array makes it suitable for mobile robots, embedded devices, and other human-machine interaction apparati. The simultaneous use of ITD and ILD as related attributes is another advantage because of their complementary character. It is made possible by the learning approach, also unique for this range of problems. We have shown that the use of ILD in addition to ITD does improve the accuracy of direction and distance estimates of the sound source, especially for the distance. The employed non-coplanar array with a minimal number of sensors is another distinctive feature of this work.

TABLE II. Comparative timings of various routines.

Processing	Linear search	SHOSLIF
230 ms	15 ms	2.5 ms

ACKNOWLEDGMENTS

The authors would like to thank Shaoyun Chen for making SHOSLIF-N code available for use by this project and having helped the project in many other ways.

[1] V. Capel, *Microphones in Action* (Fountain, Argus, Hertfordshire, England, 1978).
[2] H. A. Carr. *An Introduction to Space Perception* (Hafner, New York, 1966).
[3] C. J. MacCabe and D. J. Furlong, ''Virtual imaging capabilities of surround sound systems,'' J. Audio Eng. Soc. **42**(1/2), 38–48 (1994).
[4] W. M. Hartmann, ''Localization of a source of sound in a room,'' in *Proc. AES 8th International Conference*, 1990, pp. 27–32.
[5] W. M. Hartmann and B. Rakerd, ''Localization of sound in rooms IV: the Franssen effect,'' J. Acoust. Soc. Am. **86**, 1366–1373 (1989).
[6] F. L. Wightman and D. J. Kistler, ''The dominant role of low-frequency interaural time differences in sound localization,'' J. Acoust. Soc. Am. **91**, 1648–1661 (1992).
[7] W. A. Yost and G. Gourevitch, *Directional Hearing* (Springer-Verlag, New York, 1987).
[8] J. Blauert, ''Sound localization in the median plane,'' Acustica **22**, 205–213 (1969).
[9] M. S. Brandstein and H. F. Silverman, ''A practical methodology for speech source localization with microphone arrays,'' Comput. Speech Lang. **11**(2), 91–126 (1997).
[10] M. S. Brandstein, J. E. Adcock, and H. F. Silverman, ''A closed-form location estimator for use with room environment microphone arrays,'' IEEE Trans. Speech Audio Process. **5**(1), 45–50 (1997).
[11] M. S. Brandstein, J. E. Adcock, and H. F. Silverman, ''A practical time-delay estimator for localizing speech sources with a microphone array,'' Comput. Speech Lang. **9**, 153–169 (1995).
[12] D. V. Rabinkin *et al.*, ''A DSP implementation of source location using microphone arrays,'' J. Acoust. Soc. Am. **99**, 2503–2529(A) (1996).
[13] K. M. Martin, ''Estimating azimuth and elevation from interaural differences,'' in *1995 IEEE Mohonk Workshop on Applications of Signal Processing to Acoustics and Audio*, October 1995.
[14] U. Bub, M. Hunke, and A. Weibel, ''Knowing who to listen to in speech recognition: visually guided beamforming,'' in *Proceedings of the 1995 ICASSP*, Detroit, MI, 1995.
[15] M. S. Brandstein, ''A pitch based approach to time-delay estimation of reverberant speech,'' in *Proc. 1997 Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, 19–22 October 1997.
[16] Y. Chan, R. Hattin, and J. Plant, ''The least squares estimation of time delay and its use in signal detection,'' IEEE Trans. Acoust., Speech, Signal Process. **26**(3), 217–222 (1978).
[17] J. Ianiello, ''Time delay estimation via cross-correlation in the presence of large estimation errors,'' IEEE Trans. Acoust., Speech, Signal Process. **30**(6), 998–1003 (1982).
[18] C. Knapp and C. Carter, ''The generalized correlation method for estimation of time delay,'' IEEE Trans. Acoust., Speech, Signal Process. **24**(4), 320–327 (1976).
[19] S. L. Hobbs, ''Asymptotic statistics for location estimates of acoustic signals,'' J. Acoust. Soc. Am. **91**, 1538–1544 (1992).
[20] J. Weng and S. Chen, ''Vision-guided navigation using SHOSLIF,'' Neural Networks **11**, 1511–1529 (1998).
[21] B. Champagne, S. Bedard, and A. Stephenne, ''Performance of time-delay estimation in the presence of room reverberation,'' IEEE Trans. Speech Audio Process. **4**(2), 48–152 (1996).
[22] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. (Academic, New York, 1990).

Directivity factors for linear arrays of velocity sensors

Benjamin A. Cray[a] and Albert H. Nuttall
Naval Undersea Warfare Center Division, 1176 Howell Street, Newport, Rhode Island 02841

(Received 19 November 1999; revised 11 January 2001; accepted 23 March 2001)

Some of the features unique to beamforming a linear array of acoustic velocity sensors, which are not present with scalar-sensing elements (such as conventional pressure sensors), are described in this paper. Four types of sensors are considered here: a uniaxial motion sensor, which measures acoustic particle velocity along a single axis; a biaxial motion sensor measuring velocity in two orthogonal directions; a triaxial motion sensor that measures all three orthogonal components of the velocity vector; and a sensor, denoted as an acoustic vector sensor, that measures acoustic pressure as well as the complete velocity vector. Comparisons are made of the directivity index for each type of sensor and for linear arrays of sensors. It is shown that uniaxial velocity sensors can have a maximum directivity factor three times greater than an omnidirectional pressure sensor, a gain in directivity index of 4.8 dB. Not surprisingly, this directivity gain is highly dependent on signal arrival direction. Indeed, a uniaxial velocity sensor's directivity can be less than that of an omnidirectional pressure, indicative of a loss in signal level. These comparisons further indicate that a single vector sensor can provide 6 dB of directivity gain, four times the directivity of a pressure sensor. Line arrays of directional sensors can have a directivity index approximately 5 dB greater than that of an identical line array of pressure sensors for approximately all azimuthal array steerings. [DOI: 10.1121/1.1373706]

PACS numbers: 43.60.Qv, 43.30.Wi [JCB]

I. INTRODUCTION

Acoustic vector sensors are defined here to be sensors that measure the amplitude and phase of acoustic pressure and acoustic particle motion in a given direction at a collocated point. These sensors measure both a scalar (pressure) and vector quantity, i.e., acoustic particle motion (acceleration, velocity, or displacement). In this work only acoustic particle velocity will be considered. A single-axis, or uniaxis, acoustic motion sensor measures one component of the acoustic field vector, biaxial sensors measure motion in two orthogonal directions, while a triaxial sensor measures all three orthogonal components. These sensors can be configured into an array of elements, each element of which may be delayed, weighted, and summed, as commonly done with standard pressure hydrophones.

A group of researchers[1–13] have recently documented some aspects of beamforming arrays of vector sensors. D'Spain *et al.*[1–3] constructed and deployed a 16-element vertical line array that measured the three components of acoustic particle velocity as well as acoustic pressure. The low-frequency array DIFAR, as in DIrectional low-Frequency Analysis and Recording, consisted of pressure hydrophones with three orthogonally mounted geophones. The output from a geophone sensor is directly proportional to acoustic particle velocity. At-sea data collected from the array were beamformed conventionally (delayed and summed with uniform amplitude weighting) and adaptively, using a minimum variance Capon approach on all four quantities. Both single-element and whole array beamforming results were presented.

This paper extends the theoretical framework presented by D'Spain by quantifying, and comparing, the directivity gains of conventional pressure-sensing arrays to those of linear vector-sensing arrays. The directivity index (N_{di}) of an array is defined as a decibel measure of the improvement in the signal-to-noise ratio (SNR) that a beamformed array provides in an ideal isotropic noise field with a perfectly correlated plane-wave signal, relative to an omnidirectional array element located in an ideal reflection-free field, or simply a free-field environment.

Nehorai and Paldi[7] developed an analytical model to compare the direction-of-arrival (DOA) estimation performance of an array of vector sensors to that of an array that measures acoustic pressure only. The paper treated both single and multiple acoustic sources in an ideal free-field environment. Cramer–Rao lower bounds (CRLB) on DOA accuracy for the single-source single-vector sensor as well as multiple-source multivector sensors were given. The CRLB provides a universal measure of achievable unbiased estimation accuracy and is commonly used by signal processors to compare the merits of different high-resolution DOA methods. Hawkes and Nehorai[8,9] examined the influence of a plane boundary near a vector-sensing array had on DOA estimations, and, as done previously in the paper by Nehorai, expressed for the CRLB for various stand-off distances from ideal pressure-release and rigid boundaries.

Pesotskii and Smaryshev[10] considered the directional gain of point, line, and planar arrays consisting of monopole pressure sensors, and single-axis dipole sensors. Four types of array configurations were examined—a perfectly absorbing baffle, a compliant baffle, a rigid baffle, and a free-field environment. A comparison was then made for each type of sensor on each type of baffle. The study concluded that the combination of measuring acoustic pressure with a single

[a]Author to whom correspondence should be addressed.

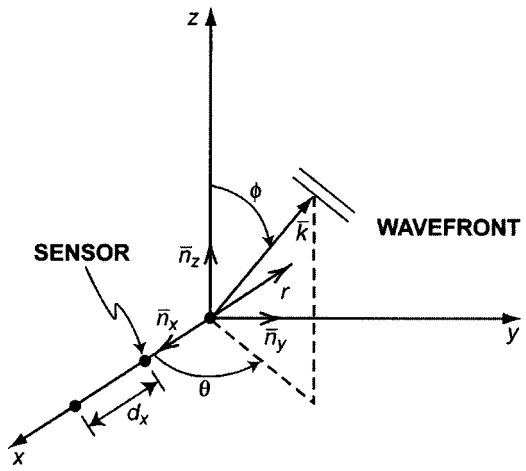


FIG. 1. Array coordinates and geometry for an equispaced line array of multiaxis velocity sensors along the *x*-axis.

component of the acoustic velocity increases the direction gain four-, three-, and two-fold for a point, line, and planar array in the free field, respectively. This is in agreement with a subset of the results presented here.

II. SINGLE SENSOR RESPONSE

In Cartesian (Fig. 1) coordinates, single-frequency, time-harmonic acoustic plane waves, which have wavefronts propagating in the direction of acoustic wave vector $-\mathbf{k}$, can be characterized by acoustic pressure or by acoustic particle velocity. Thus

$$p(\mathbf{r},t)=\mathrm{Re}\{P(\theta,\phi)\exp[i(\mathbf{k}\cdot\mathbf{r}+\omega t+\phi_p)]\}, \tag{1}$$

$$v_x(\mathbf{r},t)=\mathrm{Re}\{V_x(\theta,\phi)\exp[i(\mathbf{k}\cdot\mathbf{r}+\omega t+\phi_x)\}, \tag{2}$$

$$v_y(\mathbf{r},t)=\mathrm{Re}\{V_y(\theta,\phi)\exp[i(\mathbf{k}\cdot\mathbf{r}+\omega t+\phi_y)\}, \tag{3}$$

and

$$v_z(\mathbf{r},t)=\mathrm{Re}\{V_z(\theta,\phi)\exp[i(\mathbf{k}\cdot\mathbf{r}+\omega t+\phi_z)\}, \tag{4}$$

where $\|\mathbf{k}\|$ $(=\omega/c)$ is the acoustic wave number, ω is the circular frequency, c is the sound speed, $\mathbf{r}=[x,y,z]$ is the position vector, $P(\theta,\phi)$ is the amplitude of the sound wave pressure, and $V_x(\theta,\phi)$, $V_y(\theta,\phi)$, $V_z(\theta,\phi)$ are the amplitudes of the component velocities in the *x*, *y*, and *z* directions, respectively. The amplitudes are real-valued and dependent on azimuthal angle θ and polar angle ϕ. The complete acoustic velocity field may be written as

$$\mathbf{v}^{(3)}(\mathbf{r},t)=v_x(\mathbf{r},t)\mathbf{n}_x+v_y(\mathbf{r},t)\mathbf{n}_y+v_z(\mathbf{r},t)\mathbf{n}_z. \tag{5}$$

The vectors $\{\mathbf{n}_x,\mathbf{n}_y,\mathbf{n}_z\}$ are mutually perpendicular unit vectors aligned as shown in Fig. 1. The elements of a uniaxis vector-sensing array measure only one of the velocity vector components, such as v_x, v_y, or v_z, while the multiaxis arrays measure two or more of these orthogonal vectors. A superscript may be used to denote the dimension of the motion sensor.

For acoustic planewave propagation, the components of the collocated motion sensor are assumed to be in-phase, with the component phases set to zero. For brevity, the harmonic time-dependence $\exp(i\omega t)$ of the acoustic pressure and particle velocities will be suppressed henceforth.

The gradient of acoustic pressure and the particle velocities are related via Euler's linearized momentum equation.[11] For harmonic waves,

$$\nabla p(\mathbf{r})=-\rho_0\frac{\partial\mathbf{v}^{(3)}}{\partial t}$$

$$=-i\omega\rho_0(v_x(\mathbf{r})\mathbf{n}_x+v_y(\mathbf{r})\mathbf{n}_y+v_z(\mathbf{r})\mathbf{n}_z), \tag{6}$$

where the density of the fluid medium is denoted ρ_0.

Pressure-gradient hydrophones measure the quantity ∇p by measuring the phase difference between two spatially separated hydrophones. A well-defined dipole (or a cardioid acoustic response) can be obtained with the paired hydrophones. The technique introduces finite difference approximation errors due to the geometry of the paired hydrophones and not to the inherent sensitivity of either hydrophone.

Each orthogonal component measured by an ideal velocity sensor has an amplitude response proportional to the cosine of the conical angle Θ between the plane-wave incidence arrival direction and each component normal (the component's maximum response axis). Thus

$$V_j(\theta,\phi)=V\cos(\Theta_j),\quad j=x,y,z, \tag{7}$$

where direction cosines

$$\cos(\Theta_x)=\cos(\theta)\sin(\phi), \tag{8}$$

$$\cos(\Theta_y)=\sin(\theta)\sin(\phi), \tag{9}$$

$$\cos(\Theta_z)=\cos(\phi), \tag{10}$$

and V is the amplitude of the incident plane-wave acoustic velocity. The pressure is omnidirectional, that is, $P(\theta,\phi)=P$.

Euler's momentum equation relates the gradient of the pressure field to acoustic particle velocity. For single frequency harmonic plane waves, this relationship reduces[14] to the well-known definition for a fluid's characteristic impedance $|p(\mathbf{r},t)|/|v(\mathbf{r},t)|=\rho c$. Thus one may define an *equivalent* particle velocity amplitude, or normalized pressure, that is measured by the pressure sensor

$$v_p(\mathbf{r},t)=V\exp[i(\mathbf{k}\cdot\mathbf{r}+\omega t)], \tag{11}$$

where $V=-P/\rho c$.

Equation (5) represents the velocity vector, whereas Eq. (1), or equivalently, Eq. (11), is the scalar pressure field. To combine the scalar and vector quantities, a weighting vector (which may have complex-valued components) can be introduced: $\mathbf{w}=w_x\mathbf{n}_x+w_y\mathbf{n}_y+w_z\mathbf{n}_z$.

III. BEAMFORMING AN ARRAY OF VELOCITY SENSORS

A possible approach to beamforming an array of triaxial velocity sensors would be to delay, weight, and sum the velocity components from each sensor separately. Thus

$$\mathbf{v}_{\Sigma}^{(3)}(\theta,\phi)=V\cos(\Theta_x)\sum_{n=0}^{N-1} w_{xn}\exp[i(\mathbf{k}-\mathbf{k}_s)\cdot\mathbf{r}_n]\mathbf{n}_x + V\cos(\Theta_y)\sum_{n=0}^{N-1} w_{yn}\exp[i(\mathbf{k}-\mathbf{k}_s)\cdot\mathbf{r}_n]\mathbf{n}_y + V\cos(\Theta_z)\sum_{n=0}^{N-1} w_{zn}\exp[i(\mathbf{k}-\mathbf{k}_s)\cdot\mathbf{r}_n]\mathbf{n}_z, \tag{12}$$

where the amplitude shading coefficients w_{xn}, w_{yn}, and w_{zn} are arbitrary. The wave vector $\mathbf{k}$ corresponds to the acoustic plane wave arriving from any given direction, (θ,ϕ), and $\mathbf{k}_s$ is the wave vector which corresponds to the chosen steered, or look direction (θ_s,ϕ_s) of the array.

One may then be tempted to simply take the norm (or length of the weighted sum of all the vector components) of $\mathbf{v}_{\Sigma}^{(3)}$ in Eq. (12) as a means to obtain the beam response of the summed vector components. That is, the norm may be written as

$$\beta^{(3)}(\theta,\phi)=\sqrt{\left(\sum_{n=0}^{N-1} w_{xn}v_{xn}\right)^2+\left(\sum_{n=0}^{N-1} w_{yn}v_{yn}\right)^2+\left(\sum_{n=0}^{N-1} w_{zn}v_{zn}\right)^2}, \tag{13}$$

where $v_{xn}=V_x(\theta,\phi)\exp\{i(\mathbf{k}-\mathbf{k}_s)\cdot\mathbf{r}_n\}$, etc. This approach results, however, in taking the square-root of the sum of the squared vector velocity components; a nonlinear processing technique.

Another approach to array beamforming would be to steer *each* velocity sensor toward the source, and then obtain the squared magnitude of the sum of the sensors. That is, form the inner product of the summed velocities in Eq. (12), $\mathbf{v}_{\Sigma}^{(3)}$, with a steered triaxial velocity sensor of unit amplitude and then take the magnitude-squared of the product. Thus for the line array in Fig. 1, $(\mathbf{r}_n=[x_n,0.,0.])$, this approach yields

$$B^{(3)}(\theta,\phi)=\|\mathbf{v}_{\Sigma}^{(3)}(\theta,\phi)\cdot\mathbf{v}^{(3)}(\theta_s,\phi_s,\mathbf{0})\|^2 = g^{(3)}(\theta_s,\phi_s,\theta,\phi)^2 \times\left|\sum_{n=0}^{N-1} w_n\exp\left(i\frac{\omega}{c}x_n\alpha(\theta,\phi)\right)\right|^2, \tag{14}$$

where

$$g^{(3)}(\theta_s,\phi_s,\theta,\phi)=V(\sin\phi\sin\phi_s\cos(\theta-\theta_s)+\cos\phi\cos\phi_s),$$

and

$$\alpha(\theta,\phi)=\cos(\theta)\sin(\phi)-\cos(\theta_s)\sin(\phi_s).$$

The quantity $g^{(3)}$ reaches a minimum of V when $(\theta,\phi)=(\theta_s,\phi_s)$ and $(\theta,\phi)=(\theta_s+\pi,-\phi_s)$. Hence, without a corresponding measurement of pressure, the triaxial velocity sensor will have directional ambiguity.

This second approach to beamforming is limited in its capability. The restriction of weighting values to be essentially the direction cosines and constant from element to element disallows emphasis of some element outputs over others and it imposes the same relative weighting to the three velocity components.

A more general *linear* processing technique would form the inner product of the velocity at each element location with an arbitrary element weighting vector, $\mathbf{w}_n^{(3)}$, which now may vary at each element array location. The resulting scalar field may then be delayed and summed in a conventional manner. Thus in general,

$$B^{(3)}(\theta,\phi)=\left|\sum_{n=0}^{N-1}\mathbf{w}_n^{(3)}\cdot\mathbf{v}_n^{(3)}e^{i(\mathbf{k}-\mathbf{k}_s)\cdot\mathbf{r}_n}\right|^2, \tag{15}$$

and for the line array,

$$B_L^{(3)}(\theta,\phi)=\left|\sum_{n=0}^{N-1}(w_{xn}V_x+w_{yn}V_y+w_{zn}V_z)\times\exp\left(i\frac{\omega}{c}x_n\alpha(\theta,\phi)\right)\right|^2. \tag{16}$$

Hence, the result here is a linearly weighted combination of all three measured velocity components. That is, this approach yields the square of the sum of the weighted velocity components. A special case of the general weighting vector, $\mathbf{w}_n^{(3)}$, would be constant direction cosine weights.

The most general linear processing would be to augment Eq. (15) with a measurement of acoustic pressure. Processing all $4N$ measured quantities, i.e., pressure and three velocity components at each element location, yields

$$B^{(pv)}(\theta,\phi)=\left|\sum_{n=0}^{N-1}(\mathbf{w}_n^{(3)}\mathbf{v}_n^{(3)}+w_{pn}v_{pn})e^{i(\mathbf{k}-\mathbf{k}_s)\cdot\mathbf{r}_n}\right|^2. \tag{17}$$

There are now $4N$ degrees of freedom to manipulate, in order to maximize some measure of performance, such as the array output SNR. These $4N$ degrees of freedom are the N pressure weights $\{w_{pn}\}$ and the N velocity weights for each orthogonal direction, namely, $\{w_{xn}\}$, $\{w_{yn}\}$, $\{w_{zn}\}$ for $1\leqslant n\leqslant N$.

For reference, the power response for a line array of unit-amplitude scalar pressure sensors located along the x-axis is

$$B^{(p)}(\theta,\phi)=\left|\sum_{n=0}^{N-1} w_{pn}\,\exp\left(i\frac{\omega}{c}x_n\alpha(\theta,\phi)\right)\right|^2. \tag{18}$$

IV. ANALYTICAL EVALUATION OF DIRECTIVITY

Directivity is a fundamental measure of array performance; the directivity factor, *DF*, of an array can be expressed as[15]

$$DF=\frac{4\pi B(\theta_s,\phi_s)}{\int_0^{2\pi}\int_0^{\pi}B(\theta,\phi)\sin(\phi)d\phi\,d\theta}. \tag{19}$$

The directivity index is $10\log(DF)$.

Calculation of the DF typically requires the numerical evaluation of the above double integral. However, for certain array geometries and frequencies (e.g., for an equally spaced line array of pressure sensors), the integrals may be analytically evaluated exactly.

The maximum array response at steering angles (θ_s,ϕ_s) for a given sensor type can be obtained from Eqs. (15), (17), and (18). Hence, for unit-amplitude pressure, triaxial, and vector sensors, the maximum linear array responses may be written as

$$\begin{aligned}B^{(p)}(\theta_s,\phi_s)&=\left|\sum_{n=0}^{N-1} w_{pn}\right|^2,\\ B^{(3)}(\theta_s,\phi_s)&=\left|\sum_{n=0}^{N-1} a(\theta_s,\phi_s)w_{xn}+b(\theta_s,\phi_s)w_{yn}\right.\\ &\qquad\left.+c(\phi_s)w_{zn}\right|^2,\end{aligned} \tag{20}$$

and

$$B^{(pv)}(\theta_s,\phi_s)=\left|\sum_{n=0}^{N-1} w_{pn}+a(\theta_s,\phi_s)w_{xn}+b(\theta_s,\phi_s)w_{yn}+c(\phi_s)w_{zn}\right|^2, \tag{21}$$

where $a(\theta,\phi)=\cos(\theta)\sin(\phi)$, $b(\theta,\phi)=\sin(\theta)\sin(\phi)$, and $c(\phi)=\cos(\phi)$ for all θ,ϕ. These maximum array responses provide the numerator $[B(\theta_s,\phi_s)]$ for *DF* in Eq. (19).

The effort in obtaining expressions for the directivity factors is then in analytically evaluating the double integral given in Eq. (19), to be noted by *I*.

A. Directivity factors for a single element, $N=1$

It is straightforward to derive the directivity factor of a single omnidirectional pressure sensor. Namely,

$$I=\int_0^{2\pi}\int_0^{\pi}|w_{p0}|^2\,\sin(\phi)d\phi\,d\theta=4\pi(w_{p0})^2. \tag{22}$$

Hence, upon substituting Eqs. (22) and (20) into Eq. (19),

$$DF_{(p)}=\frac{4\pi(w_{p0}^2)}{4\pi(w_{p0}^2)}=1. \tag{23}$$

As expected, the directivity index is zero and the single pressure sensor is omnidirectional.

A similar calculation for a uniaxial sensor that measures the v_x component of particle velocity gives

$$\begin{aligned}I&=\int_0^{2\pi}\int_0^{\pi}(w_{x0}\cos(\theta)\sin(\phi))^2\,\sin(\phi)d\phi\,d\theta\\ &=\frac{4\pi}{3}w_{x0}^2.\end{aligned} \tag{24}$$

Upon substitution in Eq. (19),

$$DF_{(x)}=3\cos^2(\theta_s)\sin^2(\phi_s). \tag{25}$$

Likewise, for the velocity components v_y and v_z,

$$DF_{(y)}=3\sin^2(\theta_s)\sin^2(\phi_s), \tag{26}$$

and

$$DF_{(z)}=3\cos^2(\phi_s). \tag{27}$$

Equations (25), (26), and (27) indicate that a single velocity sensor may have a maximum directivity factor three times greater than an omnidirectional pressure sensor, that is, a gain in directivity index of 4.8 dB. Furthermore, directivity factors of velocity sensors are highly dependent on azimuth and elevation steering angles (θ_s,ϕ_s). At certain steering angles, the directivity factor of a uniaxial velocity sensor can be less than that of an omnidirectional pressure sensor.

The most general type of single array element would be a vector sensor. The maximum response of such an element is given as

$$\begin{aligned}B^{(pv)}(\theta_s,\phi_s)&=(w_{p0}+a(\theta_s,\phi_s)w_{x0}+b(\theta_s,\phi_s)w_{y0}\\ &\qquad+c(\phi_s)w_{z0})^2,\end{aligned} \tag{28}$$

and the two-fold integral becomes

$$\begin{aligned}I&=\int_0^{2\pi}\int_0^{\pi}B^{(pv)}(\theta,\phi)\sin(\phi)d\phi\,d\theta\\ &=\frac{4\pi}{3}(3w_{p0}^2+w_{x0}^2+w_{y0}^2+w_{z0}^2).\end{aligned} \tag{29}$$

Hence, the directivity factor of a single element that measures all three components of particle velocity and acoustic pressure is

$$DF_{(pv)}=\frac{B^{pv}(\theta_s,\phi_s)}{(w_{p0}^2)+\frac{1}{3}(w_{x0}^2+w_{y0}^2+w_{z0}^2)}. \tag{30}$$

The goal now becomes that of determining the weights that maximize the above ratio. Without loss in generality, one may let $w_{p0}=1$ and solve the following equations simultaneously:

$$\frac{\partial DF_{pv}}{\partial w_{x0}}=\frac{\partial DF_{pv}}{\partial w_{y0}}=\frac{\partial DF_{pv}}{\partial w_{z0}}=0. \tag{31}$$

This leads to two sets of optimal weights, one corresponding to a minimum DF_{pv} and the other, given below, for maximum directivity, i.e.,

$$w_{x0}=3a(\theta_s,\phi_s), \tag{32}$$

$$w_{y0}=3b(\theta_s,\phi_s), \tag{33}$$

$$w_{z0}=3c(\phi_s). \tag{34}$$

Substituting the optimal real weights in Eqs. (32), (33), and (34) into Eq. (30) yields

$$DF_{pv}=1+3(a(\theta_s,\phi_s)^2+b(\theta_s,\phi_s)^2+c(\phi_s)^2)=4, \tag{35}$$

since $a(\theta,\phi)^2+b(\theta,\phi)^2+c(\phi)^2=1$ for all θ,ϕ.

Therefore, the maximum directivity for a single sensor that measures all velocities as well as acoustic pressure is $DI=10\log(4)=6$ dB, and this holds for any steering angle. This gain is equivalent to the gain obtained from a four-element line array of pressure sensors (at the frequency for which there is half-wavelength separation between elements). The four-element line array would be longer than the one-element vector-sensing array, and would retain an ambiguous back lobe response. Of course, one could configure the four pressure elements, not as a linear array, but as a pressure-gradient sensing tetrahedron. This, then, would have the equivalent properties of the single-vector sensor.

B. Directivity factors for an *N*-element line array

To simplify the derivation here, the line array elements will now be taken to be *equispaced along the z-axis* at spacing $d_z=\lambda/2$. This array orientation removes the azimuthal angle dependence (θ) from the exponential term $e^{i(\bar{\mathbf{k}}-\bar{\mathbf{k}}_s)\cdot\bar{\mathbf{r}}_0}$ [given in Eq. (17)] and simplifies the subsequent integration for directivity. The integral given by Eq. (22) for pressure sensors may be integrated directly over azimuth angle θ, and assuming real-valued weights, we obtain

$$I=2\pi\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}w_{pn}w_{pm}\int_0^{\pi}\exp[i\pi(n-m)(\cos(\phi)-\cos(\phi_s))]\sin(\phi)d\phi. \tag{36}$$

Completing the ϕ-integration yields an exact expression for the *DF* of a line array of pressure sensors, i.e.,

$$DF_p=\frac{(\Sigma_{n=0}^{N-1}w_{pn})^2}{\Sigma_{n=0}^{N-1}w_{pn}^2}. \tag{37}$$

This ratio is a maximum for uniform weights and results[16] in an optimal directivity of $DI=10\log(N)$.

Similar analytical expressions for the *DF* of vector-sensing arrays are more difficult to obtain and require one to approximate the *DF* by assuming that there are a large number of array elements *N*. For example, consider an array of uniaxial velocity sensors, again aligned on the *z*-axis, that measure the v_y component of particle velocity. Then

$$I=\int_0^{2\pi}\int_0^{\pi}\sin^2(\theta)\sin^2(\phi)\left|\sum_{n=0}^{N-1}w_{yn}\exp[in\pi(\cos(\phi)-\cos(\phi_s))]\right|^2\sin(\phi)d\phi\,d\theta. \tag{38}$$

Letting $u=\cos(\phi)$ and $u_s=\cos(\phi_s)$ and integrating over θ yields

$$I=\pi\int_{-1}^{1}(1-u^2)\left|\sum_{n=0}^{N-1}w_{yn}e^{in\pi(u-u_s)}\right|^2du. \tag{39}$$

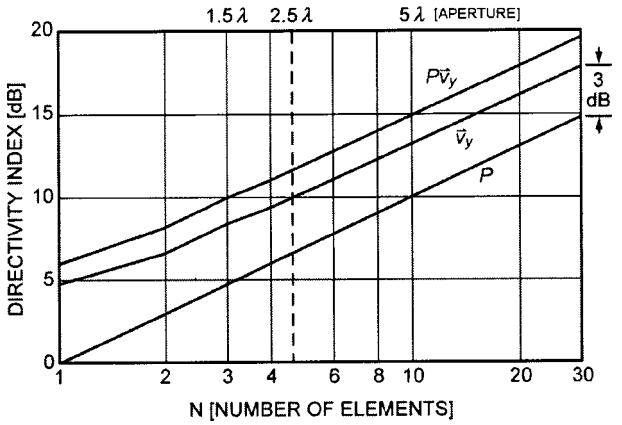


FIG. 2. Maximum directivity for an *N*-element equispaced line array at broadside ($\theta_s=90°$) utilizing optimal real weights.

For large *N* and for $u_s=\cos(\phi_s)$ not close to ±1, one may argue, as in the method of stationary phase, that for smooth weightings, the sum of the exponential terms in Eq. (39) essentially samples the remainder of the integrand at $u=u_s$. Therefore,

$$I\cong\pi(1-u_s^2)\int_{-1}^{1}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}w_{yn}w_{ym}e^{i(n-m)\pi(u-u_s)}\,du=2\pi\sin^2(\phi_s)\sum_{n=0}^{N-1}w_{yn}^2. \tag{40}$$

The *DF* then becomes, substituting Eq. (40) into Eq. (19),

$$DF_y=\frac{2\sin^2(\phi_s)(\Sigma_{n=0}^{N-1}w_{yn})^2}{\Sigma_{n=0}^{N-1}(w_{yn})^2}. \tag{41}$$

This agrees with the numerical calculations presented in Fig. 2, which will be described in more detail in Sec. V, and shows that, as the number *N* of array elements increase, and for $\phi_s=90°$, the directivity of a uniaxial velocity-sensing array approaches a gain that is 3 dB greater than that of a pressure-sensing array. Approximations for DF_x, DF_z, and even DF_{pv} can be obtained in a similar manner.

V. DERIVATION OF ARRAY GAIN

Array gain provides an SNR metric for an array's performance in various types of noise fields, directivity is limited to ideal isotropic noise. The formulation here will allow for the analysis of general noise fields, provided the array's covariance, denoted as *R*, is known or otherwise measured. Results, however, are presented for the specific case of spherically isotropic noise. To simplify the derivation of array gain, column matrices are introduced. Hence, the array's beam response is recast as

$$\beta^{(pv)}(\theta,\phi)=V\sum_{n=0}^{N-1}e_n^*(\theta_s,\phi_s)e_n(\theta,\phi)\times[w_{pn}+w_{xn}\sin\phi\cos\theta+w_{yn}\sin\phi\sin\theta+w_{zn}\cos\phi], \tag{42}$$

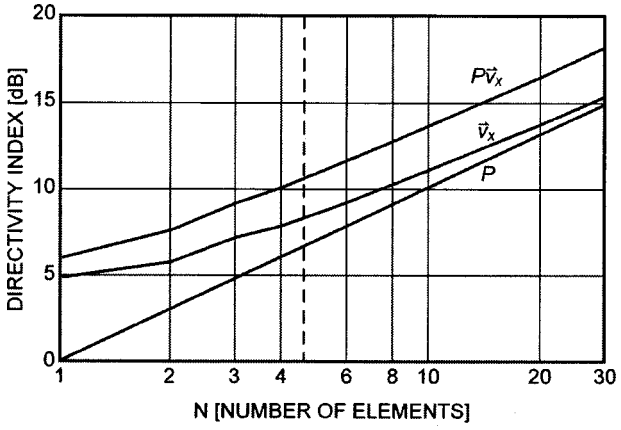


FIG. 3. Maximum directivity for an N-element equispaced line array at endfire ($\theta_s = 0°$) utilizing optimal real weights.

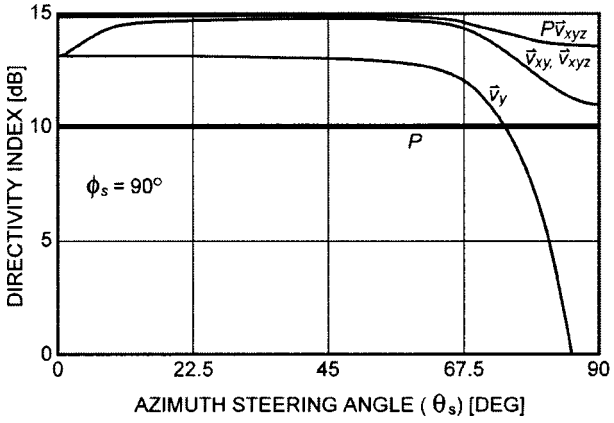


FIG. 4. Directivity versus azimuthal steering angle (θ_s) for a 10-element equispaced line array at elevation steering angle $\theta_s = 90°$.

where $e_n(\theta,\phi) = \exp[ik(x_n \sin(\phi)\cos(\theta) + y_n \sin(\theta)\sin(\theta) + z_n \cos(\phi))]$, and the $4N$ weights $\{w_{pn}\}$, $\{w_{xn}\}$, $\{w_{yn}\}$, $\{w_{zn}\}$, are defined to be real but otherwise arbitrary.

Two $4N\times 1$ real column matrices, matrices of weights and direction cosines, are defined as

$$W = [w_{p0} \ldots w_{pN-1} w_{x0} \ldots w_{xN-1} w_{y0} \ldots w_{yN-1} \times w_{z0} \ldots w_{zN-1}]^T, \tag{43}$$

$$D(\theta,\phi) = [1 \ldots 1 \cos(\Theta_x) \ldots \cos(\Theta_y) \ldots \cos(\Theta_z) \ldots].$$

Then, when the plane-wave signal arrival angle coincides with the steering angle, namely, $(\theta,\phi) = (\theta_s,\phi_s)$ the array signal output waveform is expressible as

$$Y_s = \mathbf{V} W^T D(\theta_s,\phi_s). \tag{44}$$

The array signal output power is then

$$B_{s0}^{(pv)} = |Y_s|^2 = |\mathbf{V}|^2 [W^T D(\theta_s,\phi_s)]^2. \tag{45}$$

At the same time, the array output noise waveform can be expressed as

$$Y_0 = W^T V_0, \tag{46}$$

with the $4N\times 1$ column matrix

$$V_0 = \begin{bmatrix} \beta^{(p)} \\ \beta^{(x)} \\ \beta^{(y)} \\ \beta^{(z)} \end{bmatrix}, \tag{47}$$

and the four $N\times 1$ column matrices

$$\begin{aligned} \beta^{(p)} &\equiv [e_0^*(\theta_s,\phi_s) v_{p0} \ldots e_{N-1}^*(\theta_s,\phi_s) v_{pN-1}]^T, \\ \beta^{(x)} &\equiv [e_0^*(\theta_s,\phi_s) v_{x0} \ldots e_{N-1}^*(\theta_s,\phi_s) v_{xN-1}]^T, \\ \beta^{(y)} &\equiv [e_0^*(\theta_s,\phi_s) v_{y0} \ldots e_{N-1}^*(\theta_s,\phi_s) v_{yN-1}]^T, \\ \beta^{(z)} &\equiv [e_0^*(\theta_s,\phi_s) v_{z0} \ldots e_{N-1}^*(\theta_s,\phi_s) v_{zN-1}]^T. \end{aligned} \tag{48}$$

This leads to the array output average noise power in the form

$$B_0^{(pv)} = |\overline{Y_0}|^2 = W^T \overline{V_0 V_0^H} W = W^T R W, \tag{49}$$

where R is a Hermitian $4N\times 4N$ covariance matrix of the totality of $4N$ received noise waveforms. Because all the weights in Eq. (43) are real, Eq. (49) can be written as

$$B_0^{(pv)} = W^T R_r W, \tag{50}$$

where R_r is the real part of the covariance matrix R.

We define the array gain as the quotient of the output SNR to the input SNR. There follows

$$AG = \frac{[W^T D(\theta_s,\phi_s)]^2}{W^T R_r W}, \tag{51}$$

where we have normalized the input noise pressure power at unity.

The $4N$ optimum real weights that maximize this array gain are

$$W_0 = \alpha R_r^{-1} D(\theta_s,\phi_s), \quad \alpha \text{ arbitrary}, \tag{52}$$

while the corresponding optimum array gain is

$$AG_0 = \max AG = D^T(\theta_s,\phi_s) R_r^{-1} D(\theta_s,\theta_s). \tag{53}$$

This relation was used in the following section to obtain the numerical values for maximum directivity (optimum array gain in an ideal isotropic noise field) in Figs. 2–6.

VI. DIRECTIVITY CALCULATIONS

Figure 2 compares the maximum directivity index of an N-element pressure sensing line array to velocity-sensing line arrays. The elements of the arrays are equispaced at half-wavelength spacing and are placed along the x-axis in the free field. The arrays have been steered to broadside ($\theta_s = \phi_s = 90°$) in Fig. 2 and endfire ($\theta_s = 0°$, $\phi_s = 90°$) in Fig. 3.

For a plane-wave signal arriving at broadside, that is, along the y-axis, there would be no particle velocity in the x- or z-directions. Hence, for this array steering, there would be no additional directivity gains from measuring the velocities v_x or v_z. Similarly, for plane waves arriving at endfire, only the x-component of acoustic particle velocity is nonzero.

Since the directivity calculations are based on a half-wavelength equispaced line, an $N = 10$-element pressure-sensing array has an acoustic aperture 5λ (a physical aperture of 4.5λ) for a directivity of 10 dB. Notice, in Fig. 2, that the array that measures *both* pressure and velocity provides 10 dB of directivity for an acoustic aperture of only 1.5λ, and

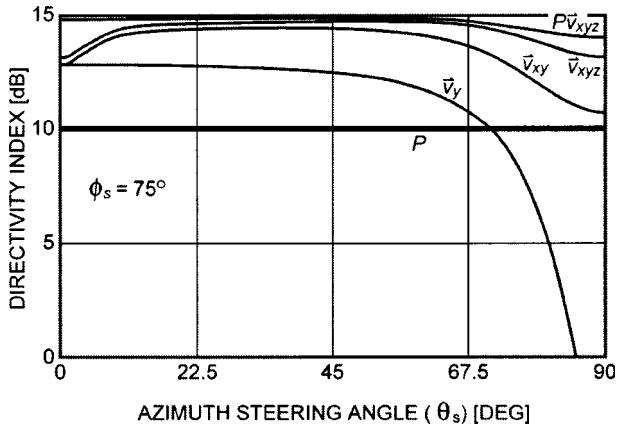


FIG. 5. Directivity versus azimuthal steering angle (θ_s) for a 10-element equispaced line array at elevation steering angle $\theta_s = 75°$.

the total number of channels, namely, 6, is still less than the number of channels of the pressure-sensing array.

The difference between the directivity gains shown in these figures is due to the orientation of the velocity sensors (v_x and v_y) with respect to the incident plane-wave signal. In Fig. 2, the difference in directivity between pressure sensing and velocity sensing (v_y-component only) remains essentially constant with increasing array aperture. This is because the azimuthal (that is, in the θ-plane shown in Fig. 1) beamwidth narrows, with increasing aperture, identically for the pressure and velocity sensors. Hence, the gain shown is due to a constant difference between elevation angle (the ϕ-plane in Fig. 1) beamwidths. The uniaxial velocity sensor has a cosine element response, whereas the pressure sensor is assumed omnidirectional. In elevation, the directivity of the velocity sensor is twice that of the pressure sensor. Hence, there is a constant difference of 3 dB.

For endfire steering, the difference between the directivity index of the uniaxial (v_x) velocity-sensing array and the pressure-sensing array diminishes with increasing array aperture. Again, the endfire beamwidth narrows identically for both the pressure- and velocity-sensing arrays with increasing aperture. However, since the arrays are steered to endfire, the effect of the broad angular difference between the cosine and omnidirectional element response becomes negligible in comparison to the endfire beamwidth. That is, with increasing aperture, the directivity is determined entirely by the array beamwidth and is minimally influenced by the individual element response.

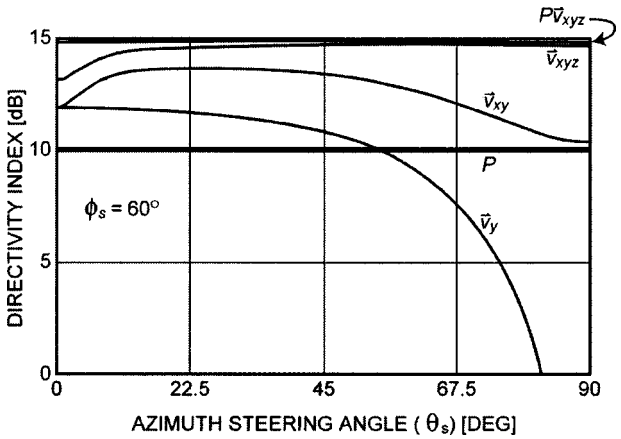


FIG. 6. Directivity versus azimuthal steering angle (θ_s) for a 10-element equispaced line array at elevation steering angle $\theta_s = 60°$.

The figures also show the directivity of an array that measures both acoustic pressure and the *y*-component of particle velocity. Including the pressure increases the directivity for both array steerings. Notice that, for a *single* element ($N=1$) and for either array steerings, the uniaxial velocity sensor provides a 4.8-dB gain over the omnidirectional pressure-sensing element. Measurement of both pressure and velocity yields a 6-dB gain.

In Fig. 4, a comparison is made between the variation of directivity with azimuth for a pressure-sensing and vector-sensing line array at a fixed frequency ($f=c/2d_x$). A 10-element line array, again aligned along the *x*-axis and in the free field, is assumed and the elevation angle is fixed at broadside. In Fig. 5, the elevation angle is fixed at 75°, whereas in Fig. 6, the elevation angle is 60°. In the free field, the optimal directivity of an array of omnidirectional pressure sensors is proportional to the number of array elements, *N*, for all array steerings. Hence, the constant directivity index of 10 dB for the pressure-sensing array is observed. Notice in Fig. 4 that the array of uniaxial velocity sensors, denoted v_y, provides 3-dB gain over pressure sensors for most azimuthal steering angles. This gain is due to the differences in the cosine and omnidirectional element response patterns. However, at oblique elevation angles, this gain diminishes in proportion to the cosine of the elevation angle.

Biaxial and triaxial sensors further improve directivity, particularly for incident angles far from broadside where the gains can be very large. For example, in Fig. 5, the directivity index of the uniaxial array is 5 dB at 77° azimuth. On the other hand, the triaxial array directivity is 14 dB. Notice that the maximum gain is achieved when the pressure and all three components of particle velocity are measured and that this optimal gain is essentially constant for all array steerings. However, measuring each component of acoustic velocity can be difficult. Additional self-noise mechanisms exist with velocity sensors.

VII. CONCLUSIONS

Presented in this study are formulations for the beam response and directivity of arrays of uniaxial, biaxial, triaxial velocity sensors, and vector sensors; sensors that measure pressure and all components of acoustic velocity at a collocated point. To simplify the development, each orthogonal component of a multiaxis velocity sensor was assumed to be collocated and the orientation of the components was unvarying from sensor to sensor. However, the amplitude shading coefficients for each component of particle velocity was allowed to be arbitrary.

Figures 4, 5, and 6 indicate that a 10-element line array of vector sensors, with elements separated by one-half an acoustic wavelength, has a directivity index approximately 5 dB greater than that of an identical line array of pressure sensors for approximately all azimuthal array steerings. This implies that vector-sensing arrays, albeit with a larger num-

ber of output channels, may be physically smaller than pressure-sensing arrays, *without* a loss in directivity.

By definition, the uniaxial velocity sensor is more directional than an omnidirectional pressure-sensing element, and it is this directionality that provides the constant 3-dB increase in the array directivity index shown in Fig. 2. This gain then, for broadside array steering only, is due simply to the assumed differences in the pressure and uniaxial element responses.

It should be noted that pressure sensors on or near acoustic baffles often have an element response pattern approximated well by a cosine raised to a fractional power near unity. For these configurations, the directivity index for a linear pressure-sensing array and any type of linear velocity-sensing array *steered to broadside*, would be approximately equivalent. However, vector sensors may be able to provide additional gains (on an ideal anechoic baffle), as suggested in Figs. 3, 4, and 5 for other array steerings. Pressure-sensing directivity remains constant, whereas the multiaxis velocity sensor gains vary with array steering and these gains are due to measuring the additional vector components of acoustic particle velocity.

[1] G. L. D'Spain, W. S. Hodgkiss, G. L. Edmonds, J. C. Nickles, F. H. Fisher, and R. A. Harriss, ''Initial Analysis of the Data from the Vertical DIFAR Array,'' *Proceedings of Mastering the Oceans Through Technology*, (OCEANS 92), Newport, Rhode Island, October 26–29, 1992.

[2] V. A. Shchurov, ''Modern State and Prospects for Use of Underwater Acoustic Intensity Measurements,'' Scientific Information Services Preprint, Pacific Oceanological Institute, Far Eastern Branch of Russian Academy of Sciences, 1998.

[3] G. L. D'Spain, W. S. Hodgkiss, and G. L. Edmonds, ''The simultaneous measurement of infrasonic acoustic particle velocity and acoustic pressure in the ocean by freely drifting swallow floats,'' IEEE J. Ocean Eng. **16**, 195–207 (1991).

[4] K. T. Wong and D. Zoltowski, ''Closed-form underwater acoustic direction-finding with arbitrarily spaced vector hydrophones at unknown locations,'' IEEE J. Ocean Eng. **22**, 566–575 (1997).

[5] J. C. Nickles, G. L. Edmonds, R. A. Harris, F. H. Fisher, W. S. Hodgkiss, J. Giles, and G. L. D'Spain, ''A Vertical Array of Directional Acoustic Sensors,'' IEEE Oceans '92 Conference Proceedings, pp. 340–345, 1992.

[6] P. Gerstoft and J. T. Goh, ''Performance evaluation of horizontal and vertical vector sensor arrays in shallow water environments,'' Proceedings 16th International Conference on Acoustics, pp. 1643–1644, 1998.

[7] A. Nehorai and E. Paldi, ''Acoustic vector-sensor array processing,'' IEEE Trans. Signal Process. **42**, 2481–2491 (1994).

[8] M. Hawkes and A. Nehorai, ''Surface-Mounted Acoustic Vector-Sensor Array Processing,'' *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP96)*, Atlanta, GA, May 1996.

[9] M. Hawkes and A. Nehorai, ''Acoustic vector-sensor beamforming and capon direction estimation,'' IEEE Trans. Signal Process. **46**, 2291–2304 (1998).

[10] A. V. Pesotskii and M. D. Smaryshev, ''Comparative evaluation of the efficiency of receiving arrays consisting of combination monopole-dipole receivers in a free field and near a plane barrier,'' Sov. Phys. Acoust. **35**, 289–291 (1989).

[11] L. Camp, *Underwater Acoustics* (Wiley-Interscience, New York, 1970).

[12] G. L. D'Spain, F. H. Fisher, W. S. Hodgkiss, J. C. Nickles, G. L. Edmonds, and R. A. Harriss, ''The Vertical DIFAR Array,'' *Proceedings of the Deployable Surveillance Workshop*, April 1992.

[13] G. L. D'Spain, W. S. Hodgkiss, and G. L. Edmonds, ''Energetics of the deep ocean's infrasonic sound field,'' J. Acoust. Soc. Am. **89**, 1134–1158 (1990).

[14] M. C. Junger and D. Feit, *Sound, Structures, and Their Interaction* (MIT Press, Cambridge, MA, 1986).

[15] A. H. Nuttall and B. A. Cray, ''Efficient Calculation of Directivity Indices for Certain Three-Dimensional Arrays,'' NUWC-NPT Technical Report 11,129, Naval Undersea Warfare Center Detachment, New London, CT, 26 July 1996.

[16] A. H. Nuttall and B. A. Cray, ''Approximations to Directivity for Linear, Planar, and Volumetric Apertures and Arrays,'' NUWC-NPT Technical Report 10,798, Naval Undersea Warfare Center Division, Newport, RI, 25 July 1997.

Intensity-invariance of fine time structure in basilar-membrane click responses: Implications for cochlear mechanics

Christopher A. Shera[a)]
Eaton-Peabody Laboratory of Auditory Physiology, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, Massachusetts 02114 and Department of Otology and Laryngology, Harvard Medical School, Boston, Massachusetts 02115

(Received 31 January 2001; accepted for publication 16 April 2001)

Basilar-membrane and auditory-nerve responses to impulsive acoustic stimuli, whether measured directly in response to clicks or obtained indirectly using cross- or reverse-correlation and/or Fourier analysis, manifest a striking symmetry: near-invariance with stimulus intensity of the fine time structure of the response over almost the entire dynamic range of hearing. This paper explores the origin and implications of this symmetry for cochlear mechanics. Intensity-invariance is investigated by applying the EQ-NL theorem [de Boer, Aud. Neurosci. **3**, 377–388 (1997)] to define a family of linear cochlear models in which the strength of the active force generators is controlled by a real-valued, intensity-dependent parameter, γ (with $0 \leqslant \gamma \leqslant 1$). The invariance of fine time structure is conjectured to imply that as γ is varied the poles of the admittance of the cochlear partition remain within relatively narrow bands of the complex plane oriented perpendicular to the real frequency axis. Physically, the conjecture implies that the local resonant frequencies of the cochlear partition are nearly independent of intensity. Cochlear-model responses, computed by extending the model obtained by solution of the inverse problem in squirrel monkey at low sound levels [Zweig, J. Acoust. Soc. Am. **89**, 1229–1254 (1991)] with three different forms of the intensity dependence of the partition admittance, support the conjecture. Intensity-invariance of cochlear resonant frequencies is shown to be consistent with the well-known ''half-octave shift,'' describing the shift with intensity in the peak (or best) frequency of the basilar-membrane frequency response. Shifts in best frequency do not arise locally, via changes in the underlying resonant frequencies of the partition, but globally through the intensity dependence of the driving pressure. Near-invariance of fine time structure places strong constraints on the mechanical effects of force generation by outer hair cells. In particular, the symmetry requires that the feedback forces generated by outer hair cells (OHCs) not significantly affect the natural resonant frequencies of the cochlear partition. These results contradict many, if not most, cochlear models, in which OHC forces produce significant changes in the reactance and resonant frequencies of the partition. [DOI: 10.1121/1.1378349]

PACS numbers: 43.64.Bt, 43.64.Kc, 43.66.Ba [LHC]

I. INTRODUCTION

Responses of the basilar membrane and auditory nerve to acoustic clicks reveal a striking symmetry: near-invariance with stimulus intensity of the fine time structure of the response over almost the entire dynamic range of hearing. In measurements of basilar-membrane motion the symmetry appears as a near-invariance of the zero crossings of the mechanical waveform (e.g., Robles *et al.*, 1976; Ruggero *et al.*, 1992; de Boer and Nuttall, 1997; Recio *et al.*, 1998; Recio and Rhode, 2000). In the auditory nerve, the invariance is manifest at low and moderate sound levels in the approximate level independence of the latency to the peaks of both standard poststimulus-time (PST) and recovered-probability compound PST histograms (e.g., Kiang *et al.*, 1965; Goblick and Pfeiffer, 1969; Lin and Guinan, 2000).[1] Intensity-invariance of fine temporal detail is remarkably robust to the method of measurement: The symmetry appears in (1) direct measurements of mechanical and neural ''impulse responses'' obtained using acoustic clicks; (2) indirect estimates obtained by cross- or reverse-correlation using wideband noise stimuli (e.g., Carney and Yin, 1988; Carney *et al.*, 1999; de Boer and Nuttall, 1997, 2000); and (3) ''synthetic'' time-domain waveforms obtained by applying inverse Fourier analysis to frequency-domain transfer functions measured with pure tones (e.g., Recio and Rhode, 2000).

Figure 1 illustrates the near intensity-invariance of fine time structure using recent measurements of basilar-membrane (BM) click responses in chinchilla (Recio and Rhode, 2000). Although the envelopes of the response waveforms shift systematically with stimulus intensity over the 70 dB range represented in the figure, the timings of the peaks, valleys, and zero crossings remain almost unchanged. The problem of understanding the origin of this symmetry has been nicely highlighted by de Boer and Nuttall (2000). Seeking to identify necessary and sufficient conditions for obtaining the symmetry, they applied the ''EQ-NL theorem'' (de Boer, 1997) to study the intensity dependence of basilar-membrane motion in the guinea pig. Although they derived a

a)Electronic mail: shera@epl.meei.harvard.edu

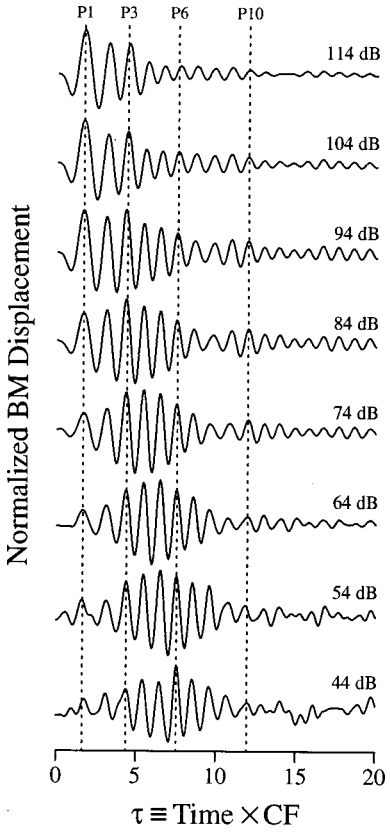


FIG. 1. Near-invariance of fine time structure in basilar-membrane click responses. The figure shows normalized BM responses to clicks from Recio and Rhode (2000, Fig. 2, chinchilla CB21). The horizontal axis measures time after the onset of umbo vibration in periods of the CF (14.5 kHz). Displacement waveforms are normalized to unit amplitude with peak click intensities indicated on the right in dB pSPL. The vertical dotted lines mark selected peaks in the response waveform (P1, P3, P6, and P10). Although the waveform envelopes vary systematically with click intensity (e.g., the envelope maximum shifts in time from about eight periods after the onset of middle-ear vibration at 44 dB pSPL to about two periods at 114 dB), the underlying fine time structure remains nearly invariant (e.g., the times of occurrence of the waveform peaks, valleys, and zero crossings generally vary by substantially less than a quarter period over the same intensity range). Adapted from Recio and Rhode (2000).

cochlear model—based on basilar-membrane impedance functions obtained from inverse analysis of measured cross-correlation functions—that exhibits the near-invariance of the timing of the mechanical impulse response, the origin and implications of the symmetry for cochlear mechanics have remained elusive.

This paper takes up the problem, adopting the modeling framework introduced in an earlier study of the frequency modulations (or ''glides'') evident in impulse responses of the basilar membrane and auditory nerve (Shera, 2001). Glides, which represent a change over time in the instantaneous frequency of oscillation of the response waveform, are nearly independent of stimulus intensity (e.g., de Boer and Nuttall, 1997; Recio *et al.*, 1998; Carney *et al.*, 1999; Recio and Rhode, 2000) and even maintain their general form post mortem (e.g., Recio *et al.*, 1998). The intensity-invariance of glides follows from the invariance of fine time structure we explore here. In an earlier paper (Shera, 2001), we demonstrated basic properties of glides, emphasizing their scaling behavior and their relation to the group delay of the frequency-domain transfer function. Here, we examine the implications for cochlear mechanics—and, in particular, for the mechanisms underlying cochlear amplification and dynamic-range compression—of the near intensity-invariance of the timing of the oscillations of mechanical and neural impulse responses. It might seem unlikely that one can learn much about the mechanisms of cochlear amplification from a phenomenon that appears to be largely independent of that amplification. That independence, however, is the key property: Whatever the outer hair cells are doing, they are doing it in a way that leaves the fine time structure of the impulse response invariant. Requiring this invariance turns out to place strong constraints on the mechanical effects of force generation by outer hair cells.

II. INVARIANCE OF BM RESONANT FREQUENCIES

A. Modeling framework

We adopt a simple modeling framework based on the classical point-impedance model of the cochlea. We assume that at sound intensities in the low-level linear regime near threshold, the velocity $V_{\rm BM}(x,f)$ of the basilar membrane at position x due to sinusoidal stimulation at frequency f can be written as the product of two factors:

$$V_{\rm BM}(x,f)=Y_{\rm BM}(x,f)P(x,f). \tag{1}$$

The first term, $Y_{\rm BM}(x,f)$, represents the admittance of the cochlear partition, and the second term, $P(x,f)$, represents the driving pressure difference across it. In the time domain, Eq. (1) becomes a convolution:

$$v_{\rm BM}(x,t)=y_{\rm BM}(x,t)*p(x,t), \tag{2}$$

where lower- and upper-case quantities (e.g., $v_{\rm BM}$ and $V_{\rm BM}$) are related by Fourier transformation. As noted in an earlier paper on glides (Shera, 2001), the admittance $y_{\rm BM}(x,t)$ and driving pressure $p(x,t)$ differ profoundly in character. The admittance term $y_{\rm BM}(x,t)$ characterizes the response of an isolated section of the cochlear partition to an impulsive force and depends only on the *local* properties of the partition at position x. The pressure term $p(x,t)$, by contrast, represents the driving force applied *in situ* and is therefore *global*; since stimuli are usually applied in the ear canal (or, effectively, at the stapes when responses are normalized by stapes motion) and must propagate to the measurement location, $p(x,t)$ depends not only on the form of the stimulus, but also on the mechanics of the entire cochlea, including boundary conditions at the stapes and helicotrema.

We simplify the discussion by assuming the approximate local scaling symmetry (Zweig, 1976; Siebert, 1968; Sondhi, 1978) manifest by basilar-membrane transfer functions (Rhode, 1971; Gummer *et al.*, 1987) and neural tuning curves (e.g., Kiang and Moxon, 1974; Liberman, 1978). Local scaling symmetry implies that rather than depending on position and frequency independently, mechanical transfer functions and tuning curves in fact depend on the two variables f and x primarily in the dimensionless combination $\beta(x,f)\equiv f/f_{\rm CF}(x)$, where $f_{\rm CF}(x)$ is the CF at location x (i.e., the cochlear position-frequency map).[2] In the time domain, scaling implies that corresponding basilar-membrane and

neural impulse responses depend on t and x through the dimensionless combination $\tau(x,t) \equiv t f_{CF}(x)$ representing time measured in periods of the characteristic frequency. Data supporting this time-domain manifestation of scaling were presented in the earlier paper on glides (Shera, 2001). When rewritten in terms of the scaling variables β and τ, Eqs. (1) and (2) become

$$V_{BM}(\beta) = Y_{BM}(\beta) P(\beta)$$

and (3)

$$v_{BM}(\tau) = y_{BM}(\tau) * p(\tau).$$

B. Parametrizing the intensity dependence

The model impedance $Z_{BM}(\beta) \equiv 1/Y_{BM}(\beta)$ characterizes the motion of the cochlear partition at sound levels in the low-level linear regime near threshold. Preparatory to modeling the intensity dependence, we follow others (e.g., Neely, 1983; Zweig, 1990; de Boer and Nuttall, 2000) and write Z_{BM} in the form

$$Z_{BM}(\beta) = Z_p(\beta) + Z_a(\beta), \tag{4}$$

representing the sum of a ''passive'' and an ''active'' component. The impedance $Z_a(\beta)$ characterizes the local effect of force generation by outer hair cells (OHCs); the impedance $Z_p(\beta)$ characterizes the mechanics of the partition obtained when those force generators have been disabled. This ''two-component'' form of the impedance is consistent with the experimental findings of de Boer and Nuttall, as reflected in their solutions to the inverse problem in guinea pig (de Boer and Nuttall, 2000). Note that the impedances are assumed to scale and are therefore written as functions of the scaling variable $\beta(x,f) = f/f_{CF}(x)$.

In the spirit of the EQ-NL theorem (de Boer, 1997) and its application in guinea pig (de Boer and Nuttall, 2000), we then model intensity dependence by defining a family of linear models in which the effective strength of the active force generators is parametrized by the factor γ:

$$Z_{BM}(\beta;\gamma) = Z_p(\beta) + \gamma Z_a(\beta), \tag{5}$$

where the real parameter γ satisfies $0 \leqslant \gamma \leqslant 1$ and depends on the amplitude of local basilar-membrane displacement, and thus, indirectly, on stimulus intensity. To indicate this dependence we write $\gamma = \gamma(I/I_0)$, where I is the intensity and I_0 a reference that sets the scale. In the low-level linear limit near threshold ($I \ll I_0$), γ is approximately 1, independent of I; at high intensities ($I \gg I_0$), γ approaches 0. At intermediate intensities, γ is presumed to vary monotonically between 1 and 0; its precise value at any given intensity depends on the form of the nonlinearity associated with the active force generators (e.g., the form of the saturating displacement–voltage transduction function of the OHC stereocilia).[3]

Each value of γ yields a corresponding linear model, $\mathcal{M}(\gamma)$. According to the EQ-NL theorem, the linear model $\mathcal{M}(\gamma)$ has the same input–output cross-correlation function as a nonlinear model in which (a) the low-level linear limit is described by $\mathcal{M}(1)$ [i.e., by Eq. (5) with $\gamma = 1$] and (b) the effect of increasing stimulus intensity is partially to saturate the active force generators (e.g., the OHCs), reducing their effective strength by the factor γ. The model $\mathcal{M}(0)$ thus describes the high-level linear limit (or postmortem condition) in which the active force generators have been entirely disabled. Application of the EQ-NL theorem to cochlear responses requires that the input–output cross-correlation functions be measured with flat-spectrum wideband noise (so that the same value of γ characterizes OHC saturation throughout the cochlea) and subsequently normalized by stapes velocity. The OHC transduction nonlinearity is assumed to be memoryless and instantaneous. For a full discussion, see de Boer (1997).

By applying the EQ-NL theorem, we have replaced the analysis of a single nonlinear model (difficult) with the analysis of a large number of linear models, one for each noise intensity (easier). Although the net result is a substantial simplification of the analysis, the substitution is valid only for quantities, such as cross-correlation functions, measured with wideband noise stimuli (de Boer, 1997). Note, however, that the phenomena we explore here—the intensity-invariance of fine time structure in mechanical and neural impulse responses—is robust to the measurement technique and is seen in derived impulse responses measured with noise stimuli (e.g., de Boer and Nuttall, 1997, 2000; Carney *et al.*, 1999).

C. Admittance poles of a simple oscillator

To probe the origin of the near-invariance of fine time structure in cochlear responses, we first consider a simpler example: the impulse response of an harmonic oscillator and its relation to the poles of the admittance in the complex plane. The admittance, Y_p, of a simple harmonic oscillator—such as the passive resonator later assumed to characterize the cochlear partition at high sound intensities—has the form

$$Y_p(f) \propto \frac{i f f_p}{f_p^2 - f^2 + i \delta_p f f_p}, \tag{6}$$

where f_p is the resonant frequency in the limit of zero damping and δ_p is the dimensionless damping constant. Introducing the normalized frequency $\beta = f/f_{CF}$ for future convenience, we now rewrite Eq. (6) to express Y_p in terms of the locations of its poles. Equation (6) becomes

$$Y_p(\zeta) \propto \frac{i \zeta \nu}{(\zeta - \zeta_\times)(\zeta - \zeta_\times^*)}, \tag{7}$$

where the normalized pole locations are denoted $\zeta_\times$ and $-\zeta_\times^*$, and the variable ζ represents the complex extension of the real variable β, defined so that ζ equals β along the real axis ($\beta = \mathrm{Re}\{\zeta\}$). The superscripted asterisk ($*$) denotes complex conjugation. The positive-frequency pole, at location $\zeta_\times$, has real and imaginary parts $\zeta_\times \equiv \beta_\times + i\alpha_\times$, where the constants $\alpha_\times = \nu \delta_p/2$ and $\beta_\times = \nu\sqrt{1 - (\delta_p/2)^2}$, with $\nu \equiv f_p/f_{CF} = |\zeta_\times|$.

Since the impulse response of the oscillator has the form $\sin(2\pi\beta_\times\tau) e^{-2\pi\alpha_\times\tau}$, where $\tau \equiv t f_{CF}$ is normalized time, we see that the real part of the pole location ($\beta_\times$) corresponds to the normalized natural frequency of oscillation and the imaginary part ($\alpha_\times$) to the decay constant of the envelope.

This correspondence implies that by varying $\alpha_\times$ while holding $\beta_\times$ fixed one generates a family of oscillators[4] whose impulse responses differ in their envelopes but maintain identical underlying frequencies of oscillation. In other words, moving the poles along lines perpendicular to the real frequency axis changes the envelope of the impulse response while preserving the fine time structure of the waveform.[5]

D. Conjecture

Analysis of the harmonic oscillator indicates that intensity-invariance of the impulse response timing corresponds to movement of the admittance poles along lines nearly perpendicular to the real frequency axis. Applying these ideas to the motion of the basilar membrane, we recall from Eq. (3) that the velocity impulse response, $v_{\mathrm{BM}}(\tau;\gamma)$, is the convolution of $p(\tau;\gamma)$ and $y_{\mathrm{BM}}(\tau;\gamma)$. Since $y_{\mathrm{BM}}(\tau;\gamma)$ depends on the analytic structure of the mechanical admittance, $Y_{\mathrm{BM}}(\beta;\gamma)$, in the complex plane, our results from the oscillator example would carry over immediately if the driving force producing the motion $v_{\mathrm{BM}}(\tau;\gamma)$ were a single impulse applied locally, as it is for the oscillator. In the cochlea, however, the driving force consists of the traveling pressure wave, $p(\tau;\gamma)$, whose dispersive character introduces additional time and frequency dependence (e.g., Shera, 2001). We note, however, that the traveling pressure wave is not independent of the mechanics of the partition; indeed, the pressure $p(\tau;\gamma)$ depends intimately on the spatial variation of the admittance, and the intensity dependence of $p(\tau;\gamma)$ is ultimately determined by that of $Y_{\mathrm{BM}}(\beta;\gamma)$. We therefore conjecture that our conclusions from the oscillator apply also to cochlear mechanics. In particular, we suggest that the near-invariance of fine time structure in BM impulse responses implies that the poles of the effective BM admittance remain within relatively narrow bands of the complex plane oriented perpendicular to the real frequency axis as the parameter γ (i.e., stimulus intensity) is varied. If our conjecture is correct, the natural resonant frequencies of the cochlear partition, defined by the real part of the admittance pole locations, must be nearly independent of intensity.

III. TESTING THE CONJECTURE

We now explore this conjecture using a simple model of cochlear mechanics. The model defined by Eq. (5) requires specification of two impedances. For later convenience—and because it corresponds with the procedure for estimating these impedances experimentally (e.g., Zweig, 1990; de Boer and Nuttall, 2000)—we take the two impedances to be those obtained (1) in the low-level linear limit [i.e., $Z_{\mathrm{BM}}(\beta;1)$] and (2) with the active mechanisms disabled [i.e., $Z_{\mathrm{BM}}(\beta;0)$]. Note that we can write $Z_{\mathrm{p}}(\beta)=Z_{\mathrm{BM}}(\beta;0)$ and $Z_{\mathrm{a}}(\beta)=Z_{\mathrm{BM}}(\beta;1)-Z_{\mathrm{p}}(\beta)$. Thus,

$$Z_{\mathrm{BM}}(\beta;\gamma)=Z_{\mathrm{p}}(\beta)+\gamma[Z_{\mathrm{BM}}(\beta)-Z_{\mathrm{p}}(\beta)], \tag{8}$$

where we define $Z_{\mathrm{BM}}(\beta)\equiv Z_{\mathrm{BM}}(\beta;1)$ as a notational shorthand. In the following sections we discuss our model forms for the low- and high-level impedances $Z_{\mathrm{BM}}(\beta)$ and $Z_{\mathrm{p}}(\beta)$ appearing in Eq. (8).

A. The impedance of the cochlear partition, $Z_{\mathrm{BM}}(\beta)$

To characterize the response of the cochlear partition in the low-level linear regime ($\gamma=1$), we adopt a variant of the model impedance obtained by solution of the inverse problem in the squirrel monkey (Zweig, 1991). We use this model both for its convenient analytic form and because it constitutes perhaps the simplest system that displays many of the qualitative features of the real cochlea. In the model, the BM impedance scales and has the form of an harmonic oscillator, with a net *negative* damping, stabilized by a feedback force proportional to the oscillator displacement at an earlier time. The model admittance has the form (Zweig, 1991)[6]

$$Y_{\mathrm{BM}}(\beta)\propto\frac{i\beta}{1-\beta^2+i\delta\beta+\rho e^{-2\pi i\mu\beta}}, \tag{9}$$

where the dimensionless parameter δ represents the net damping (with $\delta<0$); and the dimensionless parameters ρ and μ characterize, respectively, the strength and the time delay (in periods of the local resonant frequency) of the stabilizing feedback force.

The parameter values found by Zweig (1991) imply that the admittance $Y_{\mathrm{BM}}(\beta)$ has, among an infinite series of poles in the complex β (or ζ) plane, two closely spaced poles just above the real frequency axis near $\beta=1$.[7] By making both the impedance magnitude and its derivative small at frequencies near CF, the two closely spaced poles in $Y_{\mathrm{BM}}(\beta)$ help create the tall, broad peak of the transfer function (Zweig, 1990). The model variant used here has the same functional form as the original, but differs somewhat in its parameter values. By using slightly different parameter values, we can make the two closely spaced poles coincident without significant effect on the corresponding transfer function (Shera, 1992; Zweig and Shera, 1995). For ease of analysis, we use this simpler ''double-pole'' form of the admittance. The model parameter values are thus determined by specifying that the two poles principally responsible for the peak in the admittance near $\beta=1$ coincide at a given distance from the real frequency axis.[8] The parameter that sets this distance and the additional parameter, N, representing the approximate number of wavelengths of the traveling wave on the basilar membrane (Zweig *et al.*, 1976; Zweig, 1991)[9] were chosen in order to produce a BM velocity impulse response that peaks after about ten periods of the characteristic frequency (in rough agreement with data at low sound–pressure levels from guinea pig and chinchilla).

B. The passive impedance, $Z_{\mathrm{p}}(\beta)$

We take the impedance that characterizes the passive ($\gamma=0$) system, $Z_{\mathrm{p}}(\beta)$, to be a simple harmonic oscillator with positive damping.[10] This form is consistent with recent attempts to fit experimental data obtained from the basal turns of the cochlea in passive preparations (e.g., Mammano and Nobili, 1993; Brass, 2000). Initially, we take the resonant frequency characterizing the passive system (f_{p}) to have a value approximately equal to the local CF [i.e., $\nu=f_{\mathrm{p}}(x)/f_{\mathrm{CF}}(x)\approx 1$]. We explore the implications of this choice of resonant frequency in Sec. III D. To reflect the

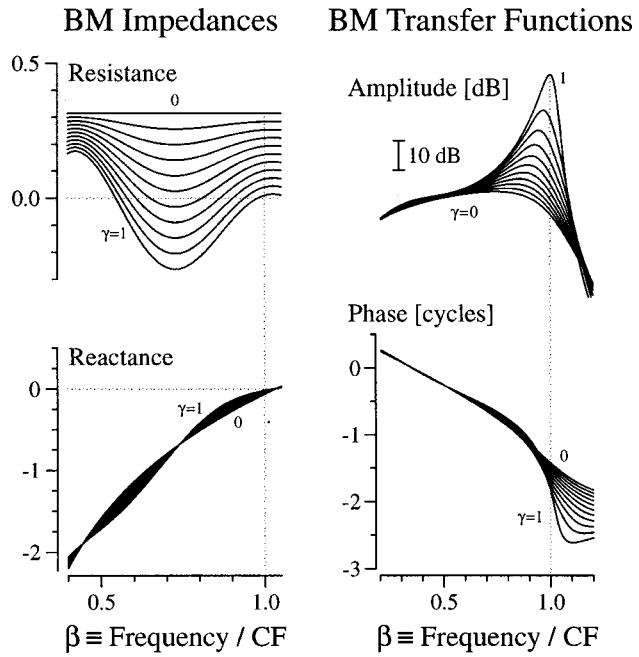


FIG. 2. Intensity dependence of BM impedances and velocity transfer functions for model $\mathcal{M}^{\approx}$. The left-hand panel shows the real (top) and imaginary parts (bottom) of the BM impedance $Z_{\mathrm{BM}}(\beta;\gamma)$ for eleven values of γ spanning the range 0–1 in steps of 0.1. Units are defined so that the proportionality constant in Eq. (9) for $Y_{\mathrm{BM}}(\beta)$ equals unity. The right-hand panel shows the amplitudes (top) and phases (bottom) of corresponding BM transfer functions, $T(\beta;\gamma)$. The dimensionless frequency variable $\beta \equiv f/f_{\mathrm{CF}}(x)$ increases along the abscissa. The vertical dotted lines locate the peak of the transfer function in the low-level linear limit ($\gamma=1$).

approximate equality of f_{p} and f_{CF}, we denote the model $\mathcal{M}^{\approx}$.[11] With functional forms now specified for both Z_{p} and Z_{BM}, the ''active impedance,'' Z_{a}, can be obtained by subtraction. Equation (8) allows us then to compute $Z_{\mathrm{BM}}(\beta;\gamma)$ as a function of γ by interpolating between these extremes.

C. Results for model $\mathcal{M}^{\approx}$

1. Transfer functions and impedances

Figure 2 shows the impedances $Z_{\mathrm{BM}}(\beta;\gamma)$ and corresponding BM transfer functions $T(\beta;\gamma)$ for values of γ spanning the full range [0,1]. The model transfer functions and impedances—both their form and their variation with intensity—bear a strong qualitative resemblance to those measured experimentally or obtained using the inverse method (cf. Figs. 2–4 of de Boer and Nuttall, 2000). For example, at the lowest effective intensity ($\gamma=1$), the real part of $Z_{\mathrm{BM}}(\beta;\gamma)$ is negative over an extended region of β just basal to the response peak at $\beta=1$. (In this description, we have used scaling to regard the figure as illustrating model impedances and transfer functions versus cochlear location at fixed frequency.[12]) The traveling wave is amplified as it propagates through the region of negative damping. At smaller values of γ (i.e., at higher intensities), $\mathrm{Re}\{Z_{\mathrm{BM}}(\beta;\gamma)\}$ increases towards 0, the region of amplification narrows, and the total gain decreases (as measured, for example, by the height of the transfer-function peak). As intensity increases further, the region of power amplification rapidly shrinks to zero and disappears (in this case, at $\gamma\approx 0.55$). Although the real part of the impedance then becomes everywhere positive, the influence of the active force generators in shaping the impedance remains evident at even the highest intensities. These changes in $\mathrm{Re}\{Z_{\mathrm{BM}}(\beta;\gamma)\}$ are accompanied by corresponding, although less dramatic, changes in the reactive component of the impedance. At all intensities the reactive component remains negative, resembling a stiffness, throughout the region of the transfer-function peak ($\beta<1.03$).[13] Note that the changes in $\mathrm{Im}\{Z_{\mathrm{BM}}(\beta;\gamma)\}$, as with those in $\mathrm{Re}\{Z_{\mathrm{BM}}(\beta;\gamma)\}$, depend strongly on location and frequency. Near the peak the cochlear amplifier acts to reduce the effective stiffness, and this reduction diminishes at higher intensities; thus, changes in $\mathrm{Im}\{Z_{\mathrm{BM}}(\beta;\gamma)\}$ are such that the effective stiffness, like the effective damping, increases with intensity.

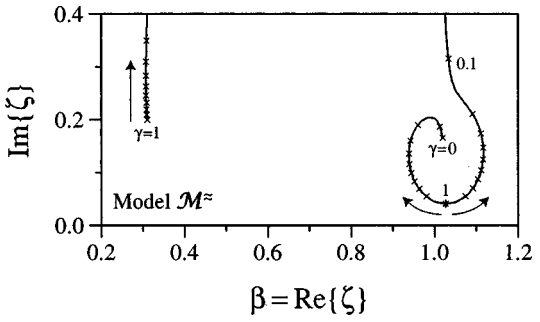


FIG. 3. Admittance pole trajectories for model $\mathcal{M}^{\approx}$. The figure shows trajectories in the ζ plane of the principal positive-frequency poles of $Y_{\mathrm{BM}}(\zeta;\gamma)$ for model $\mathcal{M}^{\approx}(\gamma)$ as a function of γ. At $\gamma=1$, $Y_{\mathrm{BM}}(\zeta;\gamma)$ has a double pole (*) at $\zeta_{*}\approx 1.03+i0.04$. As γ decreases the two poles separate slightly and move out along the solid lines in the directions shown by the arrows. Markers ($\times$) measure off equal intervals of γ spanning the range 0–1 in steps of 0.1. A third pole near $\mathrm{Re}\{\zeta\}=0.3$ contributes to shaping the ''tail'' of the transfer function. In the limit $\gamma\to 0$, all but one of the poles move off towards infinity, and the admittance $Y_{\mathrm{BM}}(\zeta;\gamma)$ becomes a passive harmonic oscillator, $Y_{\mathrm{p}}(\zeta)$, characterized by a single pole at $\zeta\approx 1.03+i0.17$.

2. Admittance pole trajectories

Figure 3 shows the locations of the principal, positive-frequency poles of the model-$\mathcal{M}^{\approx}$ BM admittance, $Y_{\mathrm{BM}}(\zeta;\gamma)$, as a function of γ. [By principal poles we mean those whose projection onto the real frequency axis falls near or below CF (i.e., $|\beta_{\times}|\lesssim 1$, where $\beta_{\times}=\mathrm{Re}\{\zeta_{\times}\}$). The remaining poles, located substantially above CF,[14] are less important in shaping the peak of the transfer function.] The lines trace out the trajectories of the principal poles as γ is decreased from 1 towards 0. For $\gamma=1$ (i.e., in the low-level linear limit), two poles coincide close to the real axis near $\mathrm{Re}\{\zeta\}=1.03$. Coincident poles are shown with an asterisk (*). A third pole near $\mathrm{Re}\{\zeta\}=0.3$ contributes to shaping the ''tail'' of the transfer function. In addition to these three principal poles (and their counterparts in the negative-frequency half plane), $Y_{\mathrm{BM}}(\zeta;\gamma)$ has an infinite string of poles at higher values of $\mathrm{Re}\{\zeta\}$ (Zweig, 1991). Note that the two coincident poles occur slightly above CF at a value $\beta_{*}=\mathrm{Re}\{\zeta_{*}\}\approx 1.03$ greater than 1. This value of β locates the normalized natural ''resonant frequency'' of an isolated section of the cochlear partition and corresponds closely to the point in Fig. 2 where the imaginary part of $Z_{\mathrm{BM}}(\beta;\gamma)$ crosses the zero line.

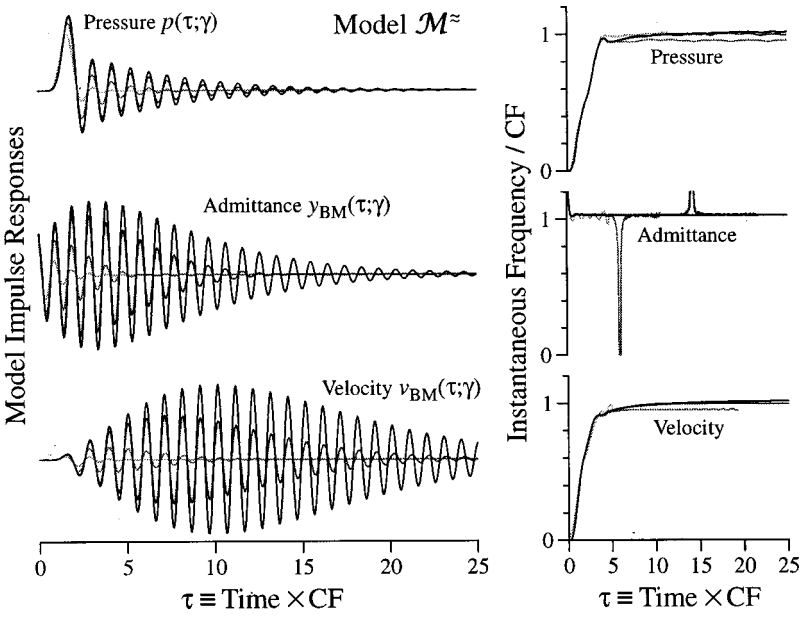


FIG. 4. Impulse responses and their instantaneous-frequency trajectories for several values of γ in model $\mathcal{M}^{\approx}$. The figure shows impulse responses (left) and corresponding instantaneous-frequency (IF) trajectories (right) for model pressure, admittance, and velocity responses (top to bottom, respectively). Responses are shown as a function of the dimensionless time variable $\tau \equiv t f_{CF}(x)$ for $\gamma = \{1, 0.95, 0.7, 0\}$, with darker lines indicating larger values of γ (i.e., lower intensities). IF trajectories were computed as described elsewhere (Shera, 2001). The sharp notches apparent in the IF trajectory for $y_{BM}(\tau;\gamma)$ (e.g., near $\tau \approx 14$ for $\gamma = 0.95$ or $\tau \approx 6$ for $\gamma = 0.7$) reflect transient phase reversals in the corresponding impulse response that result from beating between contributions from the two poles in the admittance near $f \approx f_{CF}$ (cf. Fig. 3). Note how the fine time structure for all three responses (i.e., pressure, admittance, and velocity) remains approximately independent of γ (i.e., of intensity). Indeed, the corresponding IF trajectories can be difficult to distinguish in the plot; their asymptotic values generally differ by less than 5% over the entire range of γ.

How do the poles move with intensity? As γ decreases from 1, the double pole splits apart, and the poles move out along the solid lines in the directions shown by the arrows. The symbols ($\times$) mark off equal intervals of γ along each trajectory. As γ approaches 0, all but one of the poles (and its counterpart in the negative-frequency half-plane) move off towards infinity. In the limit $\gamma \to 0$, the mechanical admittance $Y_{BM}(\zeta;\gamma)$ takes the form of a passive harmonic oscillator (Y_p) characterized by a single pole near $\zeta = 1.03 + 0.17i$. Recall from Sec. III B that the parameters of the oscillator $Y_p(\beta)$ characterizing the passive system at high sound levels ($\gamma = 0$) were chosen so that its single pole would lie almost directly above the double pole characterizing the combined (i.e., ''active+passive'') system at low levels ($\gamma = 1$). In other words, the natural ''resonant frequencies'' of the two systems, given by the real parts of the pole locations, are nearly identical. As a consequence of the approximate alignment of pole locations at the two extremes ($\gamma = 1$ and $\gamma = 0$)—and the corresponding invariance of natural resonant frequencies—the poles at intermediate values of γ also remain fairly close to the line $\mathrm{Re}\{\zeta\} = 1.03$. Although the poles separate slightly, they do so almost symmetrically about the line $\mathrm{Re}\{\zeta\} = 1.03$, so that their mean frequency stays nearly constant.

3. Impulse responses

If our conjecture is correct, we expect the fine time structure of the model impulse responses to be approximately independent of intensity (since the poles of the admittance are confined to a relatively narrow strip of the complex plane). Figure 4 demonstrates that model $\mathcal{M}^{\approx}$ does indeed capture this symmetry of the data. The figure shows model pressure, admittance,[15] and velocity impulse responses, along with corresponding instantaneous-frequency (IF) trajectories, computed at several values of γ. As discussed above, the responses at different γ can be interpreted as derived impulse responses (input–output cross-correlation functions) obtained from a single nonlinear model at different intensities (and subsequently normalized by the input at the stapes). As γ decreases (i.e., as intensity increases), the response envelopes shrink in size (the nonlinearity is compressive) and peak at progressively earlier times. Note, however, that despite these changes in the envelope of the response, the fine time structure remains almost independent of γ. For example, the asymptotic values of the IF trajectories differ by less than 5% over the entire range of γ. Frequency-domain analogues of these intensity effects—namely a strong reduction in peak amplitude accompanied by relatively small but systematic changes in phase below CF (e.g., Rhode and Recio, 2000)—can be seen in the model BM velocity transfer functions $T(\beta;\gamma)$ shown in Fig. 2.

D. Other models of intensity dependence

In model $\mathcal{M}^{\approx}$ described above the approximate invariance with intensity of the fine time structure of the impulse response is a direct consequence of the roughly vertical alignment of admittance-pole locations about the line $\mathrm{Re}\{\zeta\} = 1.03$. We illustrate this point by considering two other heuristic models of the intensity dependence of the BM admittance.

1. Model $\mathcal{M}^{<}$

In model $\mathcal{M}^{<}$, the resonant frequency of the passive admittance $Y_p(\beta)$ obtained in the limit $\gamma \to 0$ is taken to be roughly one-half octave below CF [i.e., $\nu = f_p(x)/f_{CF}(x) \approx \sqrt{2}/2$]. As a consequence of this half-octave downwards shift in resonant frequency, the admittance-pole trajectories are no longer confined to the vicinity of the line $\mathrm{Re}\{\zeta\} = 1.03$. As illustrated in Fig. 5(b), the pole closest to the real axis moves off to lower frequencies as γ decreases; at $\gamma = 0$, the pole converges on the passive pole at $\zeta \approx 0.77 + 0.17i$. Figure 6(a) shows that this variation in resonant frequency destroys the near-invariance of fine time structure (cf. model $\mathcal{M}^{\approx}$ in Fig. 4). As expected, the asymptotic frequencies of corresponding IF trajectories (not shown) vary systematically with γ, decreasing by roughly half an octave as γ approaches 0. In

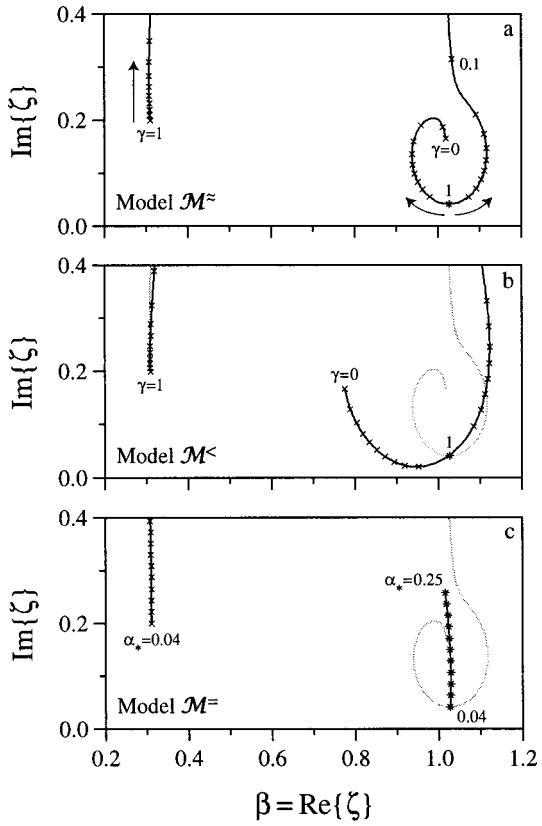


FIG. 5. Admittance-pole trajectories in the complex plane. The three panels [(a), (b), and (c)] show trajectories in the ζ plane of the principal positive-frequency poles of the BM admittance for three different model forms ($\mathcal{M}^{\approx}$, $\mathcal{M}^{<}$, and $\mathcal{M}^{=}$) of the intensity dependence. Panel (a), reproduced from Fig. 3, shows the pole trajectories for model $\mathcal{M}^{\approx}$ as a function of γ. At $\gamma=1$, $Y_{\mathrm{BM}}(\zeta;\gamma)$ has a double pole (*) at $\zeta_*=1.03+i0.04$. As γ decreases the two poles separate slightly and move out along the solid lines in the directions shown by the arrows. Markers (×) measure off ten equal intervals of γ spanning the range 0–1. In the limit $\gamma\to 0$, all but one of the poles move off towards infinity, and the admittance $Y_{\mathrm{BM}}(\zeta;\gamma)$ becomes a passive harmonic oscillator, $Y_{\mathrm{p}}(\zeta)$. Panel (b) shows the pole trajectories for model $\mathcal{M}^{<}$ as a function of γ. As in panel (a), markers (×) measure off ten equal intervals of γ spanning the range 0–1. In model $\mathcal{M}^{<}$, the resonant frequency of the passive admittance $Y_{\mathrm{p}}(\beta)$ obtained in the limit $\gamma\to 0$ is roughly one-half octave below CF. For comparison, the gray lines show the trajectories from model $\mathcal{M}^{\approx}$. Panel (c) shows the pole trajectories for model $\mathcal{M}^{=}$ as a function of α_*, the imaginary part of the double pole of the admittance. Markers (×) measure off ten equal intervals of α_* spanning the range 0.04–0.25. In model $\mathcal{M}^{=}$, the double poles of the admittance $Y_{\mathrm{BM}}(\zeta;\alpha_*)$ move along a curve nearly perpendicular to the real frequency axis. Note that all three models are identical in the low-level linear limit (i.e., for $\gamma=1$ in models $\mathcal{M}^{\approx}$ and $\mathcal{M}^{<}$ and for $\alpha_*=0.04$ in model $\mathcal{M}^{=}$).

the frequency domain, the model BM transfer functions [see Fig. 7(b)] manifest unrealistically large shifts in peak frequency as well as considerable changes in phase below CF.

2. Model $\mathcal{M}^{=}$

Our analysis predicts exact invariance of fine time structure when the poles of the admittance move along vertical lines. The pole trajectories of model $\mathcal{M}^{=}$, illustrated in Fig. 5(c), approximate this ideal, limiting case. To achieve near-perfect alignment of the pole positions in model $\mathcal{M}^{=}$, we began, as before, with the double-pole form of the BM admittance described in Sec. III A. But rather than simulating intensity dependence by varying γ in the two-component form of the impedance, as we did in models $\mathcal{M}^{\approx}$ and $\mathcal{M}^{<}$, we varied the imaginary part of the double pole directly (see Note 8). By varying the imaginary part of the double pole (denoted $\alpha_*\equiv\mathrm{Im}\{\zeta_*\}$, where the subscripted asterisk symbolizes a double pole), we constrain the principal poles of the model admittance, $Y_{\mathrm{BM}}(\zeta;\alpha_*)$, to move along a nearly vertical line in the complex plane (i.e., along a line $\beta_*\approx$constant).[16] As a consequence of this difference in model structure, the model $\mathcal{M}^{=}$ impedances, $Z_{\mathrm{BM}}(\beta;\alpha_*)$, are only approximately of the two-component form given by Eq. (5). Figure 6(b) shows that near-vertical alignment of admittance pole locations yields near-perfect invariance of fine time structure. The corresponding frequency-domain transfer functions [see Fig. 7(c)] are, on this scale, almost indistinguishable from those of model $\mathcal{M}^{\approx}$.

E. Recapitulation

Figure 8 summarizes the intensity dependence of the fine time structure in each of the three models (see also Table I). Models $\mathcal{M}^{\approx}$ and $\mathcal{M}^{=}$ manifest the near-invariance of fine time structure seen in measured responses, and in these models the poles of the BM admittance move along trajectories roughly perpendicular to the real frequency axis. We suggest that this result applies to the real cochlea: The approximate invariance of the fine time structure of the impulse response implies that the poles of the effective BM admittance remain within relatively narrow bands of the complex plane oriented perpendicular to the real frequency axis as the stimulus intensity is varied. We expect our conjecture to apply so long as the driving pressure force $p(\tau;\gamma)$ inherits its intensity dependence through the admittance, as it does in simple models. Physically, our conjecture implies that the natural resonant frequencies of the cochlear partition are nearly independent of intensity. Put yet another way, the feedback forces generated by the outer hair cells (or, more generally, by the ''cochlear amplifier'') do not significantly change the natural resonant frequencies of the cochlear partition.

IV. CONSISTENCY WITH THE ''HALF-OCTAVE SHIFT''

Is our conjecture that the resonant frequencies of the cochlear partition are nearly independent of intensity contradicted by the well-known intensity dependence of the peak (or best) frequency of the BM transfer function, which shifts to lower frequencies at higher intensities? Reference to the responses of models $\mathcal{M}^{\approx}$ and $\mathcal{M}^{=}$ in Fig. 7 demonstrates that the answer is ''No.'' Note, for example, that although the poles of the model-$\mathcal{M}^{=}$ admittance move nearly vertically—and the resonant frequencies are therefore essentially independent of intensity—the best frequency (BF) of the transfer function shifts systematically with level from a peak at $f/f_{\mathrm{CF}}=1$ (for $\alpha_*=0.04$) to a peak roughly one-half octave lower (at $\alpha_*=0.25$). Models $\mathcal{M}^{\approx}$ and $\mathcal{M}^{=}$ therefore reproduce the ''half-octave shift'' in best frequency without any corresponding change in the underlying resonant frequencies of the system.

If the half-octave shift does not reflect a change in the local resonant frequency of the cochlear partition (e.g., due to a change in stiffness), what then is the mechanism that

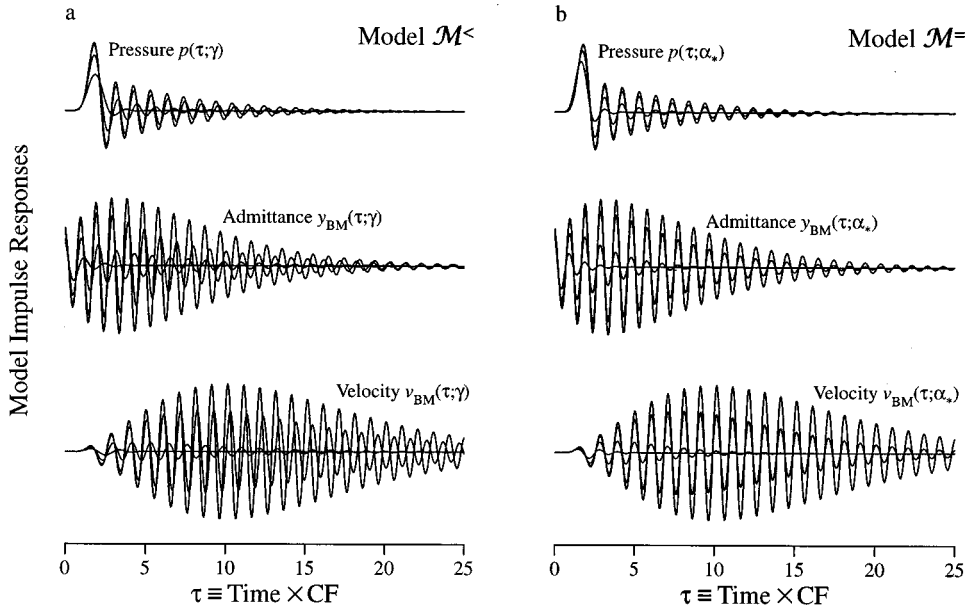


FIG. 6. Impulse responses for two different model forms ($\mathcal{M}^<$ and $\mathcal{M}^=$) of the intensity dependence of the effective BM admittance. Panel (a) shows model $\mathcal{M}^<$ impulse responses for several values of γ. As in Fig. 4, the figure shows impulse responses for model pressure, admittance, and velocity responses (top to bottom, respectively). Responses are computed for $\gamma=\{1,0.95,0.7,0\}$ and are normalized by input at the stapes so that response amplitudes decrease at smaller values of γ (i.e., at higher intensities). In model $\mathcal{M}^<$, the resonant frequency of the passive admittance $Y_p(\beta)$ obtained in the limit $\gamma\rightarrow 0$ is roughly one-half octave below CF. Note how the fine time structure of the response varies strongly with γ (i.e., with intensity). Panel (b) shows model $\mathcal{M}^=$ impulse responses for several values of α_*. Responses are computed for $\alpha_*=\{0.04,0.05,0.1,0.25\}$ and normalized by input at the stapes so that response amplitudes decrease at larger values of α_* (i.e., at higher intensities). In model $\mathcal{M}^=$, the double poles of the admittance $Y_{BM}(\zeta;\alpha_*)$ move along a curve nearly perpendicular to the real frequency axis. As a result, the fine time structure is essentially independent of intensity.

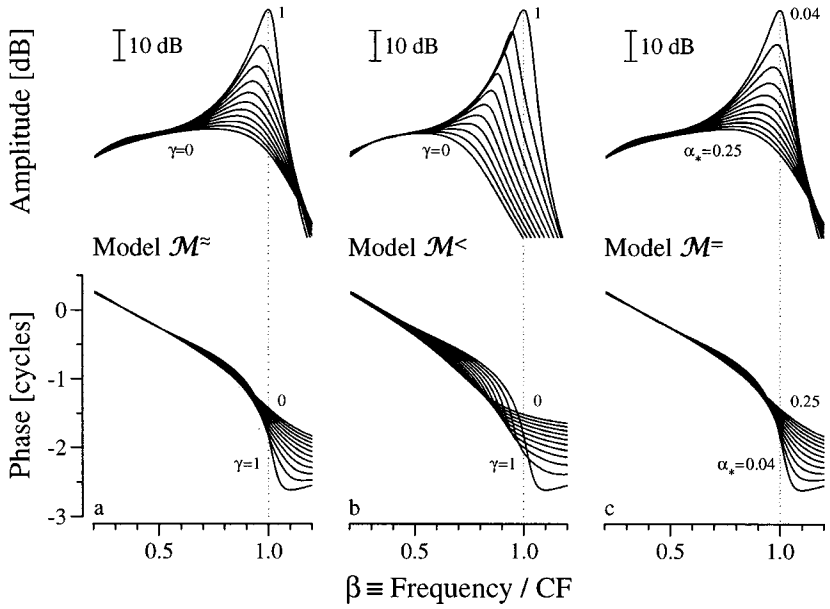


FIG. 7. Intensity dependence of model BM velocity transfer functions. The three panels [(a), (b), and (c)] show the amplitude (top) and phase (bottom) of the BM velocity transfer function (BM/stapes) for three different model forms ($\mathcal{M}^\approx$, $\mathcal{M}^<$, and $\mathcal{M}^=$) of the intensity dependence of the BM admittance. Panel (a) shows transfer functions for model $\mathcal{M}^\approx$, defined by the admittance-pole trajectories in Fig. 5(a). Reproduced from Fig. 2, the model-$\mathcal{M}^\approx$ transfer functions are shown for values of γ corresponding to the markers ($\times$) in Fig. 5(a) (i.e., for eleven values spanning the range 0–1 in steps of 0.1). In model $\mathcal{M}^\approx$, the resonant frequency of the passive admittance $Y_p(\beta)$ is equal to CF. Panel (b) shows transfer functions for model $\mathcal{M}^<$ at the values of γ given by the markers on the admittance-pole trajectories of Fig. 5(b). In model $\mathcal{M}^<$, the resonant frequency of the passive admittance Y_p is roughly one-half octave below CF. Panel (c) shows transfer functions for model $\mathcal{M}^=$ at the values of α_* given by the markers on the admittance-pole trajectories of Fig. 5(c). In model $\mathcal{M}^=$, the double poles of the admittance $Y_{BM}(\zeta;\alpha_*)$ move along a curve nearly perpendicular to the real frequency axis. Note that all three models are identical in the low-level linear limit.

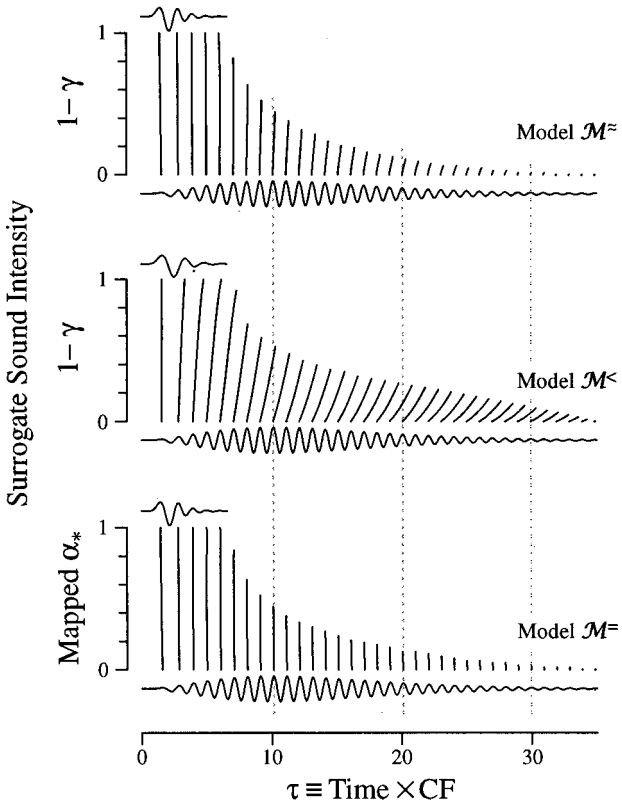


FIG. 8. Variation of fine time structure in impulse responses for three different model forms ($\mathcal{M}^{\approx}$, $\mathcal{M}^{<}$, and $\mathcal{M}^{=}$) of the intensity dependence of the effective BM admittance. In the top panel, the nearly vertical lines represent trajectories, traced out as the parameter γ is varied over the interval [0,1], marking the times of occurrence of corresponding peaks in the model-$\mathcal{M}^{\approx}$ BM velocity impulse response. So that intensity increases from top to bottom along the axis, the surrogate value $1-\gamma$ is shown along the ordinate. Time is plotted along the abscissa in units of periods of CF. Normalized response waveforms, corresponding to $\gamma=0$ and $\gamma=1$, respectively, appear at the top and bottom of each panel. At each surrogate intensity, trajectories are plotted over a total time interval equal to three times the energy-weighted average group delay of the response (truncated to the nearest whole period). The middle panel shows trajectories for model $\mathcal{M}^{<}$ in the same format as the top panel. The bottom panel shows trajectories for model $\mathcal{M}^{=}$. The ordinate shows values of α_* in the range [0.04,0.25] remapped linearly onto the interval [0,1], using the equation $\hat{\alpha}_*=[\alpha_* - \min(\alpha_*)]/[\max(\alpha_*)-\min(\alpha_*)]$, for consistency in the display. The three vertical dotted lines spanning the figure mark selected peaks in the model-$\mathcal{M}^{=}$ waveform.

creates the shift in best frequency? The answer—as with the origin of the glide (Shera, 2001)—is to be found not in the admittance, but in the driving pressure; not locally at the point of measurement, but globally in the spatial variation of geometry and mechanics that underlies the cochlear map. To see this, consider model $\mathcal{M}^{\approx}$ and recall that the BM velocity impulse response, $v_{\mathrm{BM}}(\tau;\gamma)$, is the convolution of $p(\tau;\gamma)$ and $y_{\mathrm{BM}}(\tau;\gamma)$ [Eq. (3)]. Although the best frequency of the admittance spectrum is nearly independent of intensity, the same is not true of the pressure. As intensity increases, the amplification of the traveling pressure wave is reduced. As a consequence, the pressure impulse response $p(\tau;\gamma)$ decays from its maximum amplitude more quickly and its ''center of energy'' moves to earlier times (see Fig. 4). Because of traveling-wave dispersion (Shera, 2001), however, $p(\tau;\gamma)$ at early times is dominated by the glide, which, in the base of the cochlea, consists primarily of frequencies below CF. Thus, as the duration of the driving pressure $p(\tau;\gamma)$ shortens with increasing intensity, the period of the glide becomes an ever-increasing fraction of the total duration of the response. As a result, the motion of the membrane at the measurement location becomes more and more dominated by driving frequencies lower than CF. The peak of the velocity spectrum (i.e., the best frequency) therefore shifts to lower frequencies.[17] Thus, as with the glide (Shera, 2001), the shift in BF arises not through the *local* properties of the cochlear partition—the resonant frequencies of Y_{BM} change neither with time nor with intensity—but through the *global* properties of the driving pressure.

TABLE I. Summary of the three different model forms of the intensity dependence of the effective BM admittance discussed in this paper.

Summary of models			
Model	Description	Formula	Invariance?
$\mathcal{M}^{\approx}$	Resonant frequency of Y_{p} approximately equal to CF	$f_{\mathrm{p}}(x)\approx f_{\mathrm{CF}}(x)$	Very good
$\mathcal{M}^{<}$	Resonant frequency of Y_{p} approximately one-half octave below CF	$f_{\mathrm{p}}(x)\approx f_{\mathrm{CF}}(x)/\sqrt{2}$	Poor
$\mathcal{M}^{=}$	Poles of the BM admittance move along lines nearly perpendicular to the real frequency axis	$\beta_*\approx$ constant	Excellent

A. Complementary shifts in best frequency and bandwidth

In addition to the half-octave shift, models $\mathcal{M}^{\approx}$ and $\mathcal{M}^{=}$ capture another important characteristic of the intensity dependence of BM transfer functions. As illustrated in Fig. 7, the best frequency changes relatively little over the first 30 dB reduction in peak amplitude; the bulk of the frequency shift occurs at the highest intensities over a relatively small

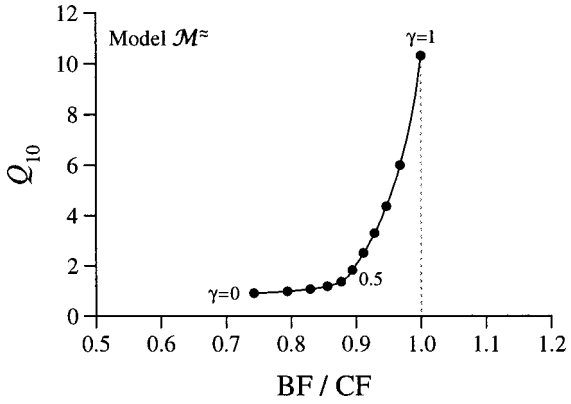


FIG. 9. Intensity dependence of transfer-function bandwidth and best frequency in model $\mathcal{M}^{\approx}$. The figure plots the Q_{10} vs the normalized best frequency (BF/CF) of the model transfer-function $T(\beta;\gamma)$ with γ as parameter. Q_{10} is defined as the ratio $\mathrm{BF}/\Delta f_{10}$, where Δf_{10} is the transfer-function bandwidth 10 dB below the peak. The dots mark the eleven values of γ corresponding to the markers on the admittance-pole trajectories in Fig. 3; they span the range 0–1 in steps of 0.1. The dotted line marks the value BF/CF=1 obtained in the low-level linear limit near threshold. Note that changes in BF and Q_{10} occur over complementary intensity ranges: Most of the change in BF occurs for $0\lesssim\gamma\lesssim 0.5$; most of the change in Q_{10} for $0.5\lesssim\gamma\lesssim 1$.

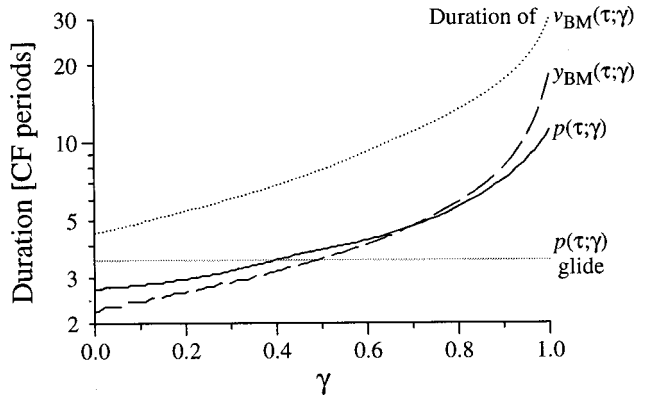


FIG. 10. Impulse-response durations in model $\mathcal{M}^{\approx}$. The figure shows the durations of the impulse responses $p(\tau;\gamma)$ (solid line), $y_{BM}(\tau;\gamma)$ (dashed line), and $v_{BM}(\tau;\gamma)$ (dotted line) as a function of γ. Response duration is defined here as the time (in periods of the CF) at which the envelope of the response decays to 10% of its peak value. As expected from the convolution in Eq. (2), the duration of $v_{BM}(\tau;\gamma)$ is approximately equal to the sum of the durations of $p(\tau;\gamma)$ and $y_{BM}(\tau;\gamma)$. The gray line shows the duration of the glide in the pressure $p(\tau,\gamma)$, defined as the time at which the instantaneous frequency reaches 90% of CF. Its value is independent of γ. Note that the durations of $p(\tau;\gamma)$ and $y_{BM}(\tau;\gamma)$ become comparable to the duration of the pressure glide at $\gamma\approx 0.5$, corresponding roughly with the value of γ at the bend in the curve of Fig. 9.

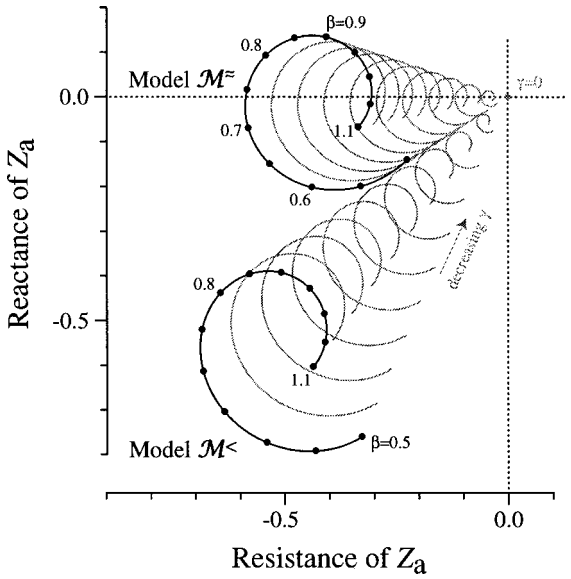


FIG. 11. Polar plot of the impedance $Z_a(\beta)$ for models $\mathcal{M}^{\approx}$ and $\mathcal{M}^{<}$. The two black curves trace out the real (resistive) and imaginary (reactive) parts of $Z_a(\beta)$ as β varies over the interval [0.5,1.1] containing the peak of the transfer function. Dots on the two curves indicate equal intervals (0.05) of β. The gray lines show how the effective impedance $\gamma Z_a(\beta)$ changes as γ decreases to 0 in steps of 0.1. Dotted lines mark the positions of the real and imaginary axes.

part of the total dynamic range. The curve shown in Fig. 9 quantifies this nonuniform shift in BF for model $\mathcal{M}^{\approx}$. Figure 9 also illustrates how changes in the best frequency and bandwidth of the response occur over complementary parts of the intensity range, in agreement with experimental data (e.g., Møller, 1977). Note, in addition, that the slope of the curve indicates that the bandwidth of the transfer function is a strong function of γ at values of γ close to 1. Even small reductions in the effective strength of the cochlear amplifier (e.g., due to surgical trauma) can therefore produce relatively large changes in the bandwidth (and group delay) of the response to threshold-level sounds.

The nonuniform shift in best frequency with intensity can be understood from the conceptual model used to explain the half-octave shift. The model suggests that shifts in the BF of $v_{BM}(\tau;\gamma)$ remain relatively small so long as the impulse responses $p(\tau;\gamma)$ and $y_{BM}(\tau;\gamma)$ last longer than the duration of the pressure glide. Figure 10 compares these durations as a function of γ for model $\mathcal{M}^{\approx}$. Although quantitative details depend on precisely how one defines the duration of the response, model results are qualitatively consistent with the conceptual analysis. At values $\gamma\gtrsim 0.5$, both $p(\tau;\gamma)$ and $y_{BM}(\tau;\gamma)$ last longer than the duration of the glide, and the BF therefore changes relatively little with intensity (cf. Fig. 9). At values $\gamma\lesssim 0.5$, however, the durations of both $p(\tau;\gamma)$ and $y_{BM}(\tau;\gamma)$ become comparable to or less than the length of the glide, and intensity-related shifts in BF become larger.

V. CONSTRAINTS ON THE MECHANISMS OF COCHLEAR AMPLIFICATION

In Sec. III B we represented the total basilar-membrane impedance $Z_{BM}(\beta;\gamma)$ in the form

$$Z_{BM}(\beta;\gamma)=Z_p(\beta)+\gamma Z_a(\beta), \tag{10}$$

representing the sum of a passive and an active component (Z_p and Z_a, respectively), where the coefficient of Z_a varies with intensity. For purposes of explication we have regarded the two models $\mathcal{M}^{\approx}$ and $\mathcal{M}^{<}$—which are, by construction, identical in the low-level linear limit ($\gamma=1$)—as differing in the form of the underlying passive impedances, Z_p. In particular, the passive impedances in these two models were taken to differ in the locations of their resonant frequencies relative to CF. To explore the implications of our results for cochlear biophysics, we now take a complementary view and ask: What constraint does the intensity-invariance of BM resonant frequencies place on the mechanical effects of force generation by outer hair cells, as characterized by the active impedance Z_a?[18]

A. Two-component form of $Z_a(\beta)$

We begin by examining the form of $Z_a(\beta)$ in models $\mathcal{M}^{\approx}$ and $\mathcal{M}^{<}$. For impedances of the two-component form (10), the active impedance $Z_a(\beta)$ can be obtained by simple subtraction of the impedances $Z_{BM}(\beta)$ and $Z_p(\beta)$ characterizing the low- and high-level linear limits, respectively: $Z_a(\beta)=Z_{BM}(\beta)-Z_p(\beta)$. Figure 11 shows a polar plot of the impedance $Z_a(\beta)$ for the models $\mathcal{M}^{\approx}$ and $\mathcal{M}^{<}$. For reference, Fig. 12 shows corresponding values of $Z_{BM}(\beta;\gamma)$ and their variations with γ. The solid curves in Fig. 11 trace out the real and imaginary parts of the model impedances $Z_a(\beta)$ as β varies over the interval [0.5,1.1] containing the peak of the transfer function. We focus on this region because outside the peak region the BM impedance obtained by solution of the inverse problem is less reliable (and its precise form less important in determining the shape of the transfer function).

In both models, $Z_a(\beta)$ resembles a spiral arc, offset

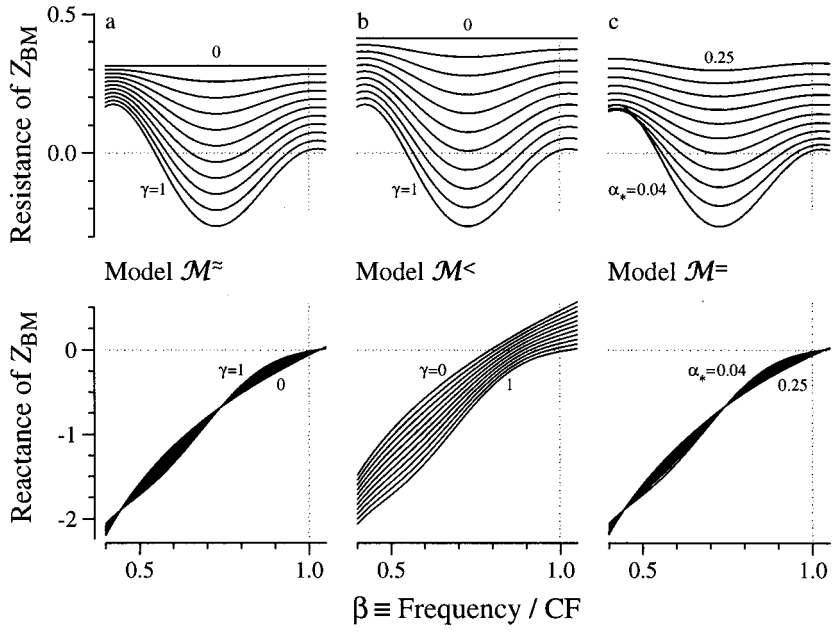


FIG. 12. Intensity dependence of model BM impedance functions. The three panels [(a), (b), and (c)] show the real (top) and imaginary parts (bottom) of $Z_{BM}(\beta;\gamma)$ for the three different model forms ($\mathcal{M}^{\approx}$, $\mathcal{M}^{<}$, and $\mathcal{M}^{=}$) of its intensity dependence defined by the admittance-pole trajectories in Fig. 5. Panel (a) shows impedances for model $\mathcal{M}^{\approx}$, reproduced from Fig. 2, at values of γ corresponding to the markers ($\times$) in Fig. 5(a). Panel (b) shows impedances for model $\mathcal{M}^{<}$ at the values of γ corresponding to the markers in Fig. 5(b). Panel (c) shows impedances for model $\mathcal{M}^{=}$ at the values of α_* corresponding to the markers in Fig. 5(c). Units are defined so that the proportionality constant in Eq. (9) for $Y_{BM}(\beta)$ equals unity. The impedances of all three models are identical in the low-level linear limit.

from the origin and traced out clockwise at nearly constant ''angular velocity'' as β increases uniformly. As γ decreases from 1, the spiral arcs contract towards the origin. Figure 11 suggests that the impedance $Z_a(\beta)$ can be approximated as the sum of two components:[19]

$$Z_a(\beta) \approx Z_a^{\text{constant}} + Z_a^{\text{spiral}}(\beta), \tag{11}$$

where Z_a^{constant} is a frequency-independent component that locates the center of the spiral and $Z_a^{\text{spiral}}(\beta)$ traces out the spiral by circling about Z_a^{constant}. The two models $\mathcal{M}^{\approx}$ and $\mathcal{M}^{<}$ differ primarily in the form of Z_a^{constant}. In model $\mathcal{M}^{\approx}$, the impedance Z_a^{constant} is negative real (a negative resistance). In model $\mathcal{M}^{<}$, however, Z_a^{constant} is complex, implying that it affects both the resistance and the reactance of the partition. These impedance changes, and their dependence on γ, are evident in the plots of $Z_{BM}(\beta;\gamma)$ shown in Fig. 12.

The constant impedance change effected by Z_a^{constant} is modulated with frequency by $Z_a^{\text{spiral}}(\beta)$, and these modulations appear in $Z_{BM}(\beta;\gamma)$. For example, the real part of $Z_{BM}(\beta;\gamma)$ manifests a bowl-shaped minimum centered roughly one-half octave below CF (see Fig. 12). The depth of the bowl, but not its ''axis of symmetry,'' varies with γ, reaching its furthest negative excursion at $\gamma=1$ (i.e., at low intensities). As discussed in Sec. III C, these variations with intensity are similar to those seen in impedances estimated using the inverse method (de Boer and Nuttal, 2000).

The bowl-shaped form of $\text{Re}\{Z_{BM}(\beta;\gamma)\}$ evident in both models is created by the oscillation in $\text{Re}\{Z_a^{\text{spiral}}(\beta)\}$, which reaches a minimum near $\beta\approx 0.74$ corresponding to the bottom of the bowl. The impedance $Z_a^{\text{spiral}}(\beta)$ also modulates the reactance, creating frequency oscillations in $\text{Im}\{Z_{BM}(\beta;\gamma)\}$ that appear roughly 90° out of phase with the modulations in the resistance. As discussed below in Sec. V B, oscillations in the resistance and reactance that appear 90° out of phase with one another are expected from causality, which requires that the real and imaginary parts of $Z_a(\beta)$ be Hilbert transforms of one another. In both models $\text{Im}\{Z_a^{\text{spiral}}(\beta)\}$ increases the effective stiffness of the partition in parts of the ''tail'' of the transfer function ($\beta \lesssim 0.74$) while decreasing the stiffness throughout most of the peak region ($0.74 \lesssim \beta \lesssim 1.03$).

Significantly, $\text{Im}\{Z_a^{\text{spiral}}(\beta)\}$ passes through 0, so that $Z_a^{\text{spiral}}(\beta)$ is nearly real, just above CF at $\beta\approx 1.03$. In model $\mathcal{M}^{\approx}$, where Z_a^{constant} is also nearly real, the zero crossing of $\text{Im}\{Z_a^{\text{spiral}}(\beta)\}$ implies that the impedance $Z_a(\beta)$ leaves the reactive component of the total partition impedance near this value of β essentially unchanged at all intensities. Reference to Fig. 12(a) shows that in model $\mathcal{M}^{\approx}$ the value $\beta\approx 1.03$ is the value where the reactive component of $Z_p(\beta)$ vanishes. Since $\text{Im}\{Z_p(\beta)\}$ and $\text{Im}\{Z_a(\beta)\}$ both vanish at the same value of β, their sum, $\text{Im}\{Z_{BM}(\beta;\gamma)\} = \text{Im}\{Z_p(\beta)\} + \gamma\text{Im}\{Z_a(\beta)\}$, also vanishes at this point and does so independent of γ. Note, however, that the vanishing of the reactance [zero crossing of $\text{Im}\{Z_{BM}(\beta;\gamma)\}$] locates the approximate natural resonant frequency of the partition [i.e., the projection of the nearby pole of the admittance $Y_{BM}(\beta;\gamma)$ along the real frequency axis].[20] Our analysis of the impedance $Z_a(\beta)$ has therefore brought us full circle: In model $\mathcal{M}^{\approx}$ we conclude that the natural resonant frequencies of the partition must be nearly independent of intensity. In model $\mathcal{M}^{<}$, by contrast, Z_a^{constant} is large and complex so that $\text{Im}\{Z_a(\beta)\}$ is always negative; in this model, therefore, $Z_a(\beta)$ modifies the reactive component of the impedance at all frequencies and intensities [see Fig. 12(b)]. As a consequence, the resonant frequencies of the partition depend on γ.

Our analysis has thus identified the constraint that intensity invariance of the resonant frequencies places on the mechanical effect of force generation by OHCs, as summarized in the impedance $Z_a(\beta)$: While generally affecting a substantial reduction in the effective damping, the OHCs must not significantly change the reactance of the passive partition at frequencies in a neighborhood about its natural resonant frequency.[21]

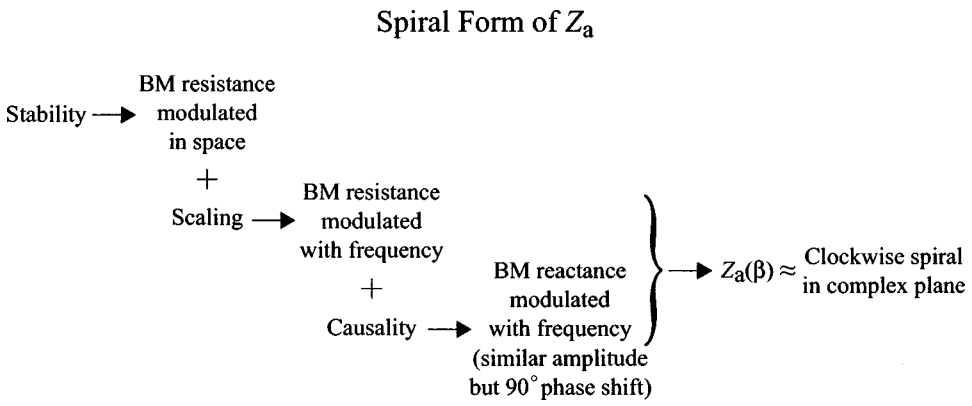


FIG. 13. Summary of the argument deducing the qualitative form of $Z_a(\beta)$ from general physical principles (stability, local scaling, and causality). The argument implies that $Z_a(\beta)$ has an approximately spiral form in the complex plane. The near-invariance of the fine time structure in the impulse response locates this spiral near the real (i.e., resistive) axis.

B. General argument for the spiral form of $Z_a(\beta)$

Illustrated above for two specific models (i.e., $\mathcal{M}^{\approx}$ and $\mathcal{M}^{<}$), the approximately spiral form of the impedance $Z_a(\beta)$ can be deduced from general principles. Figure 13 summarizes the argument. For the cochlea to remain stable, the damping of the partition cannot everywhere be negative. Stability requires that energy added to the traveling wave in one region be absorbed in another; the amplifier must therefore create negative damping over only a finite region of the cochlea (e.g., just basal to the peak of the traveling wave). Thus, the effective damping must be modulated in space. According to local scaling symmetry, however, modulation in space (at fixed frequency) requires a corresponding modulation in frequency (at fixed position). But in the frequency domain, causality implies that the real and imaginary parts of an impedance are not independent; rather, they are Hilbert transforms of one another (e.g., Bode, 1945; Papoulis, 1977). Thus, frequency modulations in the damping (real part) are necessarily accompanied by frequency modulations of similar amplitude in the reactance (imaginary part).[22] Since the Hilbert transformer acts like a 90° phase shifter (e.g., the Hilbert transform of a cosine modulation is a $-$sine), the frequency modulations (oscillations) in the damping and reactance are roughly 90° out of phase with one another. In Fig. 12(a), for example, regions of local decrease in the damping (e.g., near $\beta \approx 0.6$) correspond to local minima in the reactance oscillation, local minima in the damping (e.g., the bottom of the bowl near $\beta \approx 0.74$) correspond to regions of local increase in the reactance, and so on. As a consequence of these coupled modulations in resistance and reactance, $Z_a(\beta)$ must have an approximately spiral form traced out clockwise with increasing β (cf. Fig. 11).

This general argument yields only the approximate *shape* of the $Z_a(\beta)$ trajectory in the complex plane; it does not, of course, determine the radius of the spiral, the way the radius changes with β, nor the rate at which the spiral is traversed. Neither does the argument locate the absolute position of the spiral (i.e., the value of Z_a^{constant}) in the complex plane. As illustrated above in Sec. V A, the location of the spiral—straddling the real-frequency axis—is set by the requirement of near-invariance of fine time structure. If the resonant frequencies of the partition are to remain invariant, the impedance $Z_a(\beta)$ must leave the partition reactance unchanged at frequencies near CF.

Guided by simple models based on the inverse solution in the low-level linear regime (Zweig, 1991), we have used general principles (stability, local scaling, and causality) and the intensity-invariance of the fine time structure of the impulse response to deduce the qualitative form of $Z_a(\beta)$ representing the collective action of the OHCs. At frequencies about CF, the impedance $Z_a(\beta)$ must (1) be roughly spiral in form, (2) rotate clockwise with increasing frequency about a center with a negative real part, and (3) intersect (or at least approach) the real (resistive) axis near the natural resonant frequency of the passive partition.

C. Implications for the origin of negative damping: The fast-time-delayed stiffness model

Among the most biophysically plausible models so far proposed for the origin of negative damping has been the fast-time-delayed stiffness model (Neely, 1983; Zweig, 1990, 1991). This model is based on the observation that a negatively damped oscillator can be created from one with positive damping by the addition of a feedback force proportional to the oscillator's displacement at a previous time (i.e., a time-delayed stiffness). The impedance of a time-delayed stiffness can be written in the form

$$Z_f = \frac{K_f}{i\omega} e^{-i\omega\tau_f}, \tag{12}$$

where $K_f > 0$ is the amplitude and $\tau_f > 0$ the time delay of the feedback force. Represented in vector form, a pure stiffness (such as $K_f/i\omega$) points along the negative imaginary axis. The time delay has the effect of rotating this stiffness vector clockwise through the angle $\omega\tau_f$. For $0 < \omega\tau_f < \pi$ the time delay rotates the impedance vector into the negative real half plane, giving the impedance Z_f a negative real part. In this configuration, the feedback force can therefore reduce the damping of the system to which it is coupled.

An attractive feature of the model is that when the time delay is small compared to the period of the driving frequency, so that

$$0 < |\omega\tau_f| \ll 1, \tag{13}$$

the resistive component of Z_f is negative over a wide range of frequencies (specifically, for all $f < 1/2\tau_f$). Thus, if the feedback force is sufficiently strong and delayed by a nonzero time small compared to the oscillator's period, it can create negative damping over a broad range of characteristic

frequencies without recourse to additional tuning mechanisms (such as having the time delay vary strongly with position in the cochlea). Transduction delays on the order of a few microseconds have been suggested (Neely, 1983; Zweig, 1990); we therefore follow Zweig and label the impedance Z_f and its parameters with the subscript ''f'' for ''fast.'' Note, however, that in addition to reducing the resistance, Z_f also generally changes the reactive component of the system (e.g., its stiffness). Of course, whether Z_f produces a significant change in *either* the net resistance or reactance depends on the system to which the impedance is coupled. In the paragraphs that follow we argue that in the context of cochlear mechanics (as currently understood) the fast-time-delayed stiffness model cannot provide sufficient force to counteract the damping without also producing significant effects on the stiffness.

Zweig (1991) analyzed the fast-time-delayed stiffness model in detail; we begin by recapitulating key elements of that analysis. When the time delay is sufficiently fast to satisfy Eq. (13), the complex exponential in Eq. (12) can be expanded in powers of its argument:

$$e^{-i\omega\tau_f} = 1 - i\omega\tau_f - \tfrac{1}{2}\omega^2\tau_f^2 + \ldots . \tag{14}$$

Keeping the first three terms yields

$$Z_f = K_f/i\omega + R_f + i\omega M_f, \tag{15}$$

where

$$R_f \equiv -K_f\tau_f \quad \text{and} \quad M_f \equiv \tfrac{1}{2}K_f\tau_f^2. \tag{16}$$

The impedance of the fast-acting feedback force thus has mass and stiffness terms, together with a net negative resistance. When added to a passive harmonic oscillator (with impedance Z_p), the feedback force therefore both reduces the effective damping and modifies the natural resonant frequency. The combined system, $Z_{pf} = Z_p + Z_f$, has a net damping (Zweig, 1991)[23]

$$\delta_{pf} \approx \frac{\delta_p - \rho_f\psi_f}{\sqrt{1+\rho_f}}, \tag{17}$$

where $\delta_p > 0$ is the passive damping, $\rho_f \equiv K_f/K_p$ is the feedback strength relative to the passive stiffness, and $\psi_f \equiv \omega_p\tau_f$. The parameters K_p and $\omega_p \equiv 2\pi f_p$ are, respectively, the stiffness and resonant angular frequency of the original, passive oscillator. Similarly, the ratio of resonant frequencies becomes

$$f_{pf}/f_p \approx \sqrt{1+\rho_f}. \tag{18}$$

Consider now the constraints imposed by the near-invariance of the zero crossings of the impulse response. Figure 4 suggests that consistency with the data requires that the fractional change in resonant frequency, $|f_{pf} - f_p|/f_p$, due to the feedback force be small, say no more than roughly 10%. According to Eq. (18), this requirement imposes an upper bound on the strength of the feedback force: $\rho_f \lesssim 0.2$. Now to create a net negative damping, the feedback force must be strong enough that $\delta_{pf} < 0$. Analysis of the impedance $Z_{BM}(\beta)$ obtained using the inverse method suggests the rough estimate $\delta_{pf} \sim -\delta_p$ (Zweig, 1991). According to Eq. (17), this requires $\rho_f\psi_f \sim 2\delta_p$.[24] Recent attempts to fit experimental data obtained from the basal turns of the cochlea in passive preparations suggest that $\delta_p \gtrsim 0.1$ (e.g., Mammano and Nobili, 1993; Brass, 2000). Combining the equations yields the inequality

$$\psi_f \equiv \omega_p\tau_f \sim 2\delta_p/\rho_f \gtrsim 2(0.1)/0.2 \approx 1, \tag{19}$$

a constraint inconsistent with Eq. (13) and the assumption that the time delay is small compared to the oscillator period. In other words, the fast-time-delayed stiffness model cannot provide sufficient force to counteract the damping without shifting the resonant frequency of the system beyond the limits allowed by the data. In effect, the fast-time-delayed stiffness model therefore yields an impedance similar to the Z_a^{constant} of model $\mathcal{M}^<$, which fails to reproduce the near-invariance of fine time structure characteristic of the data.

The argument presented here does not, of course, rule out all time-delayed stiffness models, but only the simplest, in which the time delay is small compared to the period of the characteristic frequency. More elaborate models for negative damping—e.g., those that invoke additional tuning mechanisms, such as having the time delay depend strongly on position in the cochlea—remain viable. For example, if the time delay in Eq. (12) were to vary with position inversely with CF, so that $\tau_f(x)\omega_{CF}(x) \approx \pi/4$, then for frequencies $f \approx f_{CF}$ the impedance Z_f would contribute a nearly pure negative-resistance component and could presumably be arranged to produce only minor changes in the resonant frequency of the system.[25]

VI. SUMMARY AND DISCUSSION

Basilar-membrane and auditory-nerve responses to impulsive acoustic stimuli—whether measured directly in response to clicks or obtained indirectly using cross- or reverse-correlation and/or Fourier analysis—manifest a striking symmetry. A symmetry is something that stays the same while something else changes. In this case, the thing that changes is the intensity of the stimulus; the thing that stays the same is the phase of the oscillations in the response waveform (e.g., Kiang *et al.*, 1965; Goblick and Pfeiffer, 1969; Robles *et al.*, 1976; Carney and Yin, 1988; Ruggero *et al.*, 1992; de Boer and Nuttall, 1997; Recio *et al.*, 1998; Carney *et al.*, 1999; Lin and Guinan, 2000; de Boer and Nuttall, 2000; Recio and Rhode, 2000). In this paper, we have explored the origin and implications of this symmetry for cochlear mechanics. Applying the EQ-NL theorem (de Boer, 1997), we defined a family of linear cochlear models in which the strength of the active force generators is controlled by an intensity-dependent parameter, γ. We conjectured that invariance of fine time structure implies that as γ is varied the poles of the BM admittance remain within relatively narrow bands of the complex plane oriented perpendicular to the real frequency axis. Cochlear-model responses, computed by extending the model obtained by solution of the inverse problem in squirrel monkey at low sound levels (Zweig, 1991) with three different forms of the intensity dependence of the partition admittance, support the conjecture.

The models we employ here, which summarize the mechanics of the organ of Corti using equivalent point impedances and include only one-dimensional treatments of the full three-dimensional motion of the cochlear fluids, provide highly simplified representations of cochlear mechanics. By simplifying, however, we hope to ''eliminate the unnecessary so that the necessary may speak.''[26] Our intention, in other words, is not to exhibit models necessarily realistic in every detail, but rather to identify and explicate basic principles of cochlear function in the most transparent manner possible. Although obtained here using one-dimensional point-impedance models, our conclusions nevertheless apply in more realistic geometries as well. For example, one-dimensional models appear to capture, both qualitatively and semiquantitatively, the essential physics that gives rise to traveling-wave dispersion and glides (Shera, 2001). In addition, the impedances of the cochlear partition obtained as solutions to the inverse problem in long-wave, short-wave, and three-dimensional models are all in remarkable qualitative agreement (e.g., Zweig, 1991; de Boer, 1995a, b; de Boer and Nuttall, 1999). This general agreement among solutions to the inverse problem supports Zweig's (1991) conclusion that to reproduce the data, ''it is more important, in the hierarchy of approximations, to approximate . . . the impedance of the organ of Corti accurately than to work with the correct number of spatial dimensions.'' The success of our simple model, achieved despite Kolston's (2000) claim that ''three-dimensional fluid behavior should be regarded as a bare minimum in any quantitative description of cochlear mechanics,''[27] corroborates Zweig's remarks.

Physically, our conjecture implies that the local resonant frequencies of the cochlear partition are nearly independent of intensity. We demonstrate that this intensity *independence* of resonant frequencies is consistent with the well-known intensity *dependence* of the peak frequency of the BM transfer function, which shifts to lower frequencies at higher intensities (producing, at high intensities, the so-called ''half-octave shift''). We propose that, as with the glide (Shera, 2001), *the shift in best frequency arises globally*, through the intensity dependence of the dominant frequency of the driving pressure, *rather than locally*, through shifts in the local resonant frequencies of the partition. Our proposal thus resolves the long-standing paradox presented by measurements of mechanical click responses, which exhibit two seemingly contradictory features: On the other hand, the responses manifest the half-octave shift in best frequency with intensity; on the other, they exhibit near intensity-invariance of fine time structure.

Near-invariance of fine time structure requires that the feedback forces generated by the outer hair cells not significantly affect the natural resonant frequencies of the cochlear partition, which appear to vary by no more than roughly 10% over the full dynamic range of hearing. This requirement places strong constraints on the biophysical action of the cochlear amplifier or, more generally, on the mechanisms of cochlear dynamic-range compression, as characterized by the impedance $Z_a(\beta)$. In particular, we argue that the intensity invariance of fine time structure—combined with general principles, such as stability, local scaling, and causality—requires that in the region near CF the impedance $Z_a(\beta)$ must (1) be roughly spiral in form, (2) rotate clockwise with increasing frequency about a center with a negative real part, and (3) closely approach the real axis near the natural resonant frequency of the passive partition. This requirement appears inconsistent with models in which the OHCs modify the tonic stiffness of the cochlear partition, thereby effecting substantial changes in its resonant frequency (Allen, 1990, 1997). The requirement also suggests that tonic changes in OHC stiffness, whether mediated by somatic motor proteins (He and Dallos, 1999, 2000) or via the ciliary bundle (e.g., Howard and Hudspeth, 1988), have a relatively small effect on the total stiffness of the partition, at least for near-best-frequency stimuli in the basal turns of the cochlea. In this respect, our conclusions are consistent with current measurements, which suggest that the axial stiffness and the OHC is considerably smaller than the stiffness of the basilar membrane (Russell and Schauz, 1995; He and Dallos, 1999).

Our conclusions thus contradict many, if not most, cochlear models. Although most cochlear models are not manifestly nonlinear, they usually specify the equivalents of what we call $Z_{BM}(\beta)$ and $Z_p(\beta)$ (i.e., the impedances with and without contributions from force generation by OHCs). Since intensity variations appear to interpolate smoothly between these two extremes (de Boer and Nuttall, 2000), the qualitative behavior of a model's implicit intensity dependence can often be inferred from the relation between these two impedances. Our results indicate that to reproduce the invariance of the fine time structure of the impulse response, the resonant frequencies of $Z_{BM}(\beta)$ and $Z_p(\beta)$ need to be nearly identical (i.e., within roughly 10% of one another). However, plots of the effect of the cochlear amplifier on the BM admittance (Hubbard and Mountain, 1996) indicate that many cochlear models (e.g., Mountain *et al.*, 1983; Kolston *et al.*, 1990; Geisler, 1991; Hubbard, 1993) fail to satisfy this constraint, indicating that such models cannot reproduce the approximate invariance of response timing, as assessed either by varying intensity or by disabling the active mechanisms. Furthermore, our results rule out what is perhaps the most biophysically plausible mechanism so far proposed for the origin of negative damping, namely the fast-time-delayed stiffness model (Neely, 1983; Zweig, 1990). Although current cochlear models reproduce, to varying degrees, the form of empirical transfer functions measured in sensitive preparations near threshold, the problem they evidently leave unsolved is understanding the biophysical basis of an active feedback force that is strong enough to reverse the sign of the partition damping while leaving its resonant frequencies nearly unchanged.

ACKNOWLEDGMENTS

The author gratefully acknowledges many helpful discussions with and/or comments from Jont Allen, Egbert de Boer, Paul Fahey, John Guinan, Stephen Neely, William Peake, Robert Withnell, and George Zweig. This work was supported by Grant No. R01 DC03687 from the NIDCD, National Institutes of Health.

[1]Intensity-invariance of fine time structure, an excellent approximation at low and moderate sound-pressure levels (SPLs), breaks down in auditory-nerve responses at the highest sound levels. The data of Lin and Guinan (2000), for example, show clear evidence for phase reversals and other ''anomalies'' at click levels of 90 dB pSPL (peak-equivalent SPL) and above. Although species and methodological issues complicate the comparison, click responses measured on the basilar membrane show little evidence of comparable features, maintaining near-invariance of their zero crossings even at levels exceeding 115 pSPL (e.g., Recio and Rhode, 2000).

[2]In this paper, the term ''best frequency'' (BF) is used to locate the maximum of the BM frequency response, which may vary with intensity. The ''characteristic frequency'' (CF) is defined as the best frequency measured in the low-level linear limit. By definition, the CF is therefore independent of intensity.

[3]For an example illustrating the calculation of $\gamma(I/I_0)$ for a particular form of the transduction nonlinearity, see Appendix B of de Boer and Nuttall (2000).

[4]Given the normalized pole location $\zeta_\times \equiv \beta_\times + i\alpha_\times$, one can find the corresponding undamped resonant frequency and damping constant of the oscillator from the relations $f_p = f_{CF}|\zeta_\times|$ and $\delta_p = 2\alpha_\times/|\zeta_\times|$. Note that fixing the natural resonant frequency of the oscillator by moving its poles along lines of constant $\beta_\times$ requires changing both δ_p and f_p. Fractional changes in f_p, however, are generally small.

[5]Movement of the admittance poles along lines perpendicular to the real frequency axis yields exact invariance of fine time structure for the displacement response of the oscillator. For the velocity response, however, the invariance is only approximate. To see this, note that the velocity response of the oscillator has the form $v(\tau) \propto \cos(2\pi\beta_\times\tau + \phi)e^{-2\pi\alpha_\times\tau}$, where $\sin(\phi) = r/\sqrt{1+r^2}$ with $r \equiv \alpha_\times/\beta_\times$. The phase shift ϕ—and thus the fine time structure of the waveform (e.g., the position of its zero crossings)—therefore depends on $\alpha_\times$. Note, however, that this dependence on $\alpha_\times$ is weak (i.e., $\phi \ll 2\pi$ for the values $r \ll 1$ characteristic of tuned oscillators).

[6]In this paper the scaling variable β is defined as the model-independent ratio $f/f_{CF}(x)$, where $f_{CF}(x)$ is the characteristic frequency defined by the peak of the transfer function (see Note 2). Note, however, that in the model of cochlear mechanics defined by Eq. (9) (Zweig 1991), β refers to the ratio $f/f_r(x)$, where $f_r(x)$ is the undamped resonant frequency of the oscillator (i.e., the resonant frequency in the limit when the damping, δ, and stabilizing feedback force, ρ, are both negligible). The parameter values given in Note 8 imply that $f_{CF}(x)$ and $f_r(x)$ are everywhere proportional, with $f_r/f_{CF} \approx 1.03$. We have maintained the distinction between these two frequencies in all model calculations, but, for clarity of exposition, have ignored this small difference in the main text.

[7]Note that each of the infinite number of poles has a positive imaginary part. Despite creating a region of negative damping, the model is therefore stable at all frequencies (energy created at one location is absorbed at another). Equation (145) of Zweig (1991) gives an explicit expression for Y_{BM} in terms of its poles and their residues.

[8]This note describes the procedure used to determine the parameter values in the model admittance given in Eq. (9). The admittance $Y_{BM}(\beta)$ is first obtained as a function of ζ (or complex β) by analytic continuation into the complex frequency plane. [Recall from Note 6 that in this context the normalized frequency β is defined as the ratio $f/f_r(x)$.] Three constraining equations are then used to determine the three model parameters $\{\delta, \rho, \mu\}$. We specify (1) the imaginary part of one of the two closely spaced poles of $Y_{BM}(\zeta)$ and then require that (2) the real and (3) the imaginary parts of the second pole coincide with those of the first. More precisely, given $\alpha_* \equiv \mathrm{Im}\{\zeta_*\} > 0$, where ζ_* denotes the double pole of $Y_{BM}(\zeta)$, one determines the three parameters $\{\delta, \rho, \mu\}$; the real part of the double pole, $\beta_* \equiv \mathrm{Re}\{\zeta_*\}$; and the auxiliary variable a by solving the system of five simultaneous equations

$$\alpha_* = \delta/2 + a;$$

$$2\pi a\mu = 1;$$

$$a/\beta_* = \tan[2\pi(n/2 + \tfrac{3}{4}) - \beta_*/a];$$

$$\beta_*^2 = 1 - (\delta/2)^2 - a^2;$$

and

$$\rho = 2a[1 - (\delta/2)^2]^{1/2} e^{-\alpha_*/a}.$$

The solution for the auxiliary variable a is

$$a = (\alpha_* + \sqrt{\alpha_*^2 + c(1-\alpha_*^2)})/c,$$

where $c = 2 + x^{-2}$ and x is the solution to the equation

$$x^{-1} + \tan^{-1}(x) = 2\pi(n/2 + \tfrac{3}{4}),$$

obtained numerically. The values of $\{\delta, \rho, \mu\}$ can then be obtained by direct substitution. Evaluating the equations using the value $\alpha_* = \mathrm{Im}\{\zeta_*\} = 0.04$ adopted in the text (and taking $n=2$ for $\mu \approx 1\frac{3}{4}$) yields $\{\delta, \rho, \mu\} = \{-0.1024, 0.1175, 1.7450\}$. Finally, the proportionality constant in Eq. (9) for $Y_{BM}(\beta)$ was set equal to 1.

[9]The model parameter N, which determines the approximate number of wavelengths of the traveling wave on the basilar membrane in response to sinusoidal stimulation (Zweig *et al.*, 1976; Zweig, 1991), was given the value $N = 2.5$.

[10]The passive admittance thus has the form (6), with a proportionality constant of 1. In model $\mathcal{M}^{\approx}$, the damping constant δ_p was given the value $\delta_p = 0.32$; in model $\mathcal{M}^{<}$, introduced in Sec. III C 1, a slightly higher damping ($\delta_p = 0.42$) was needed to maintain model stability.

[11]Although we write ''model $\mathcal{M}^{\approx}$'' using the singular, $\mathcal{M}^{\approx}$ actually denotes an entire *family* of models, one for each value of γ.

[12]Recall that scaling relates properties of the mechanical transfer function to those of the traveling wave. In particular, mechanical transfer functions $T[f/f_{CF}(x)]$ measured as a function of f at fixed x also describe the traveling displacement wave as a function of x at fixed f. At fixed position, T is the transfer function; at fixed frequency, the traveling wave.

[13]Although qualitative agreement with the findings of de Boer and Nuttall (2000) remains strong, note that in the one-dimensional model used here the imaginary part of the partition impedance goes through zero at a value of β (≈ 1.03) closer to the location of the transfer-function peak than indicated by solutions to the inverse method obtained using two- and three-dimensional models (de Boer and Nuttall, 2000).

[14]Several of these poles located ''above CF'' are illustrated in Fig. 3 of Chap. VIII of Shera (1992) and in Fig. 5 of Zweig and Shera (1995).

[15]The admittance $y_{BM}(\tau;\gamma)$ represents the velocity response to a pressure impulse applied locally; it therefore jumps discontinuously to a nonzero value at $\tau = 0$ (see, e.g., Note 16 of Shera, 2001).

[16]To see that the principal poles of the model-$\mathcal{M}^{=}$ admittance are arrayed along a curve nearly perpendicular to the real-frequency axis, note that the equations in Note 8 yield $\beta_*^2 = 1 - \alpha_*^2 + \delta/2\pi\mu$, where $\zeta_* \equiv \beta_* + i\alpha_*$ denotes the location of the double pole of $Y_{BM}(\zeta)$ in the complex β plane. Since $\mu \approx 1\frac{3}{4}$, the quantity $|\delta|/2\pi\mu$ is typically much less than 1. Thus, $\beta_* \approx 1$ for values $\alpha_* \ll 1$. In other words, for $\alpha_* \ll 1$ the double pole lies approximately along the vertical line $\beta_* = 1$ at a distance α_* from the real axis. The small but systematic deviations from the vertical predicted by this analysis are evident in the trajectory shown in Fig. 5(c). [Recall that the line $\beta_* = 1$, for $\beta = f/f_r$, corresponds to the line $\beta_* = 1.03$, for $\beta = f/f_{CF}$ (see Note 6).]

[17]This explanation for the half-octave shift has been proposed independently by Carney (1999), who noted that intensity-dependent shifts in the temporal envelope of BM and auditory-nerve click responses, when combined with the intensity-independent frequency glide, can produce changes in the best frequency of the response.

[18]Although it characterizes the mechanical effects of local force generation by OHCs, the impedance Z_a should not be regarded as characterizing the ''cochlear amplifier.'' Cochlear amplification of traveling waves depends on both the active and passive mechanics and their interaction with the surrounding fluids over a fairly broad region of the cochlea.

[19]The suggestion that $Z_a(\beta)$ be represented as the sum of two components with this same qualitative form has been made earlier by Zweig (1990, 1991). In an effort to provide a biophysical basis for negative damping, Zweig suggested that $Z_a(\beta)$ be written as the sum of a fast- and a slow-acting time-delayed stiffness. In that model, the fast-time-delayed stiffness (delay much smaller than a period) provides negative damping and the slow-time-delayed stiffness (delay approximately $1\frac{3}{4}$ periods) provides the necessary frequency modulation by stabilizing the resulting oscillator. We discuss the fast-time-delayed stiffness model, and show that it yields an impedance Z_a^{constant} similar to that of model $\mathcal{M}^{<}$, in Sec. III C.

[20]For a passive oscillator of the form (7), the zero crossing of the reactance

$(\mathrm{Im}\{Z_p(\beta_0)\}=0)$ occurs at the value $\beta_0=|\zeta_\times|$, corresponding to the undamped resonant frequency of the oscillator (f_p).

[21]Reactance changes due to OHCs at other frequencies are not precluded: Indeed, as shown in Sec. III B, they are required by stability, local scaling, and causality.

[22]Note that the constraints of causality apply even though the real part of the impedance $Z_a(\beta)$ is negative.

[23]In deriving Eqs. (17) and (18) we have used the inequality $\psi_f=\omega_p\tau_f\ll 1$, an approximation equivalent to neglecting the mass term in Eq. (15).

[24]We have used the inequality $\rho_f\lesssim 0.2$ to approximate $\sqrt{1+\rho_f}$ as unity.

[25]This model for the origin of negative damping might be called the ''tuned-time-delayed stiffness model.''

[26]The quotation is from Hans Hoffmann (quoted in Efron and Tibshirani, 1993).

[27]Kolston may be referring here to recent suggestions from M theory that the universe we inhabit actually comprises ten spatial dimensions (see, e.g., Greene, 1999).

Allen, J. B. (**1990**). ''Modeling the noise damaged cochlea,'' in *Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer, Berlin), pp. 324–331.

Allen, J. B. (**1997**). ''OHCs shift the excitation pattern via BM tension,'' in *Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. L. Hecht-Poinar (World Scientific, Singapore), pp. 167–175.

Bode, H. (**1945**). *Network Analysis and Feedback Amplifier Design* (Van Nostrand Reinhold, Princeton).

Brass, D. (**2000**). ''A macromechanical model of the guinea pig cochlea with realistic parameters,'' J. Acoust. Soc. Am. **107**, 894–907.

Carney, L. H. (**1999**). ''Temporal response properties of neurons in the auditory pathway,'' Curr. Opin. Neurobiol. **9**, 442–446.

Carney, L. H., McDuffy, M. J., and Shekhter, I. (**1999**). ''Frequency glides in the impulse responses of auditory-nerve fibers,'' J. Acoust. Soc. Am. **105**, 2384–2391.

Carney, L. H., and Yin, T. C. T. (**1988**). ''Temporal coding of resonances by low-frequency auditory nerve fibers: Single fiber responses and a population model,'' J. Neurophysiol. **60**, 1653–1677.

de Boer, E. (**1995a**). ''The inverse problem solved for a three-dimensional model of the cochlea. I. Analysis,'' J. Acoust. Soc. Am. **98**, 896–903.

de Boer, E. (**1995b**). ''The inverse problem solved for a three-dimensional model of the cochlea. II. Application to experimental data sets,'' J. Acoust. Soc. Am. **98**, 904–910.

de Boer, E. (**1997**). ''Connecting frequency selectivity and nonlinearly for models of the cochlea,'' Aud. Neurosci. **3**, 377–388.

de Boer, E., and Nuttall, A. L. (**1997**). ''The mechanical waveform of the basilar membrane. I. Frequency modulations ('glides') in impulse responses and cross-correlation functions,'' J. Acoust. Soc. Am. **101**, 3583–3592.

de Boer, E., and Nuttall, A. L. (**1999**). ''The inverse problem solved for a three-dimensional model of the cochlea. III. Brushing up the solution method,'' J. Acoust. Soc. Am. **105**, 3410–3420.

de Boer, E., and Nuttall, A. L. (**2000**). ''The mechanical waveform of the basilar membrane. III. Intensity effects,'' J. Acoust. Soc. Am. **107**, 1497–1507.

Efron, B., and Tibshirani, R. J. (**1993**). *An Introduction to the Bootstrap* (Chapman and Hall, New York).

Geisler, C. D. (**1991**). ''A cochlear model using feedback from motile outer hair cells,'' Hear. Res. **54**, 105–117.

Goblick, T. J., and Pfeiffer, R. R. (**1969**). ''Time-domain measurements of cochlear nonlinearities using combination click stimuli,'' J. Acoust. Soc. Am. **46**, 924–938.

Greene, B. (**1999**). *The Elegant Universe—Superstrings, Hidden Dimensions, and the Quest for the Ultimate Theory* (Norton, New York).

Gummer, A. W., Smolders, J. W. T., and Klinke, R. (**1987**). ''Basilar membrane motion in the pigeon measured with the Mössbauer technique,'' Hear. Res. **29**, 63–92.

He, D. Z. Z., and Dallos, P. (**1999**). ''Somatic stiffness of cochlear outer hair cells is voltage dependent,'' Proc. Natl. Acad. Sci. U.S.A. **96**, 8223–8228.

He, D. Z. Z., and Dallos, P. (**2000**). ''Properties of voltage-dependent somatic stiffness of cochlear outer hair cells,'' J. Assoc. Res. Otolaryngol. **1**, 46–813.

Howard, J., and Hudspeth, A. J. (**1988**). ''Compliance of the hair bundle associated with gating of mechanoelectrical transduction channels in the bullfrog's saccular hair cell,'' Neuron **1**, 189–199.

Hubbard, A. E. (**1993**). ''A traveling-wave amplifier model of the cochlea,'' Science **259**, 68–71.

Hubbard, A. E., and Mountain, D. C. (**1996**). ''Models of the cochlea,'' in *Auditory Computation*, edited by H. L. Hawkins, T. A. McMullen, A. N. Popper, and R. R. Fay (Springer, New York), pp. 62–120.

Kiang, N. Y. S., and Moxon, E. C. (**1974**). ''Tails of tuning curves of auditory-nerve fibers,'' J. Acoust. Soc. Am. **55**, 620–630.

Kiang, N. Y. S., Watanabe, T., Thomas, E. C., and Clark, L. F. (**1965**). *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (MIT, Cambridge, MA).

Kolston, P. J. (**2000**). ''The importance of phase data and model dimensionality to cochlear mechanics,'' Hear. Res. **145**, 25–36.

Kolston, P. J., Viergever, M. A., de Boer, E., and Smoorenburg, G. F. (**1990**). ''What type of force does the cochlear amplifier produce?'' J. Acoust. Soc. Am. **88**, 1794–1801.

Liberman, M. C. (**1978**). ''Auditory-nerve response from cats raised in a low-noise chamber,'' J. Acoust. Soc. Am. **63**, 442–455.

Lin, T., and Guinan, J. J. (**2000**). ''Auditory-nerve-fiber responses to high-level clicks: Interference patterns indicate that excitation is due to the combination of multiple devices,'' J. Acoust. Soc. Am. **107**, 2615–2630.

Mammano, F., and Nobili, R. (**1993**). ''Biophysics of the cochlea: Linear approximation,'' J. Acoust. Soc. Am. **93**, 3320–3332.

Moller, A. R. (**1977**). ''Frequency selectivity of single auditory-nerve fibers in response to broadband noise stimuli,'' J. Acoust. Soc. Am. **62**, 135–142.

Mountain, D. C., Hubbard, A. E., and McMullen, T. A. (**1983**). ''Electromechanical processes in the cochlea,'' in *Mechanics of Hearing*, edited by E. de Boer and M. A. Viergever (Martinus Nijhoff, The Hague), pp. 119–126.

Neely, S. T. (**1983**). ''The cochlear amplifier,'' in *Mechanics of Hearing*, edited by E. Boer and M. A. Viergever (Martinus Nijhoff, The Hague), pp. 111–118.

Papoulis, A. (**1977**). *Signal Analysis* (McGraw-Hill, New York).

Recio, A., and Rhode, W. S. (**2000**). ''Basilar membrane responses to broadband stimuli,'' J. Acoust. Soc. Am. **108**, 2281–2298.

Recio, A., Rich, N. C., Narayan, S. S., and Ruggero, M. A. (**1998**). ''Basilar-membrane responses to clicks at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **103**, 1972–1989.

Rhode, W. S. (**1971**). ''Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique,'' J. Acoust. Soc. Am. **49**, 1218–1231.

Rhode, W. S., and Recio, A. (**2000**). ''Study of mechanical motions in the basal region of the chinchilla cochlea,'' J. Acoust. Soc. Am. **107**, 3317–3332.

Robles, L., Rhode, W. S., and Geisler, C. D. (**1976**). ''Transient response of the basilar membrane measured in squirrel monkeys using the Mössbauer effect,'' J. Acoust. Soc. Am. **59**, 926–939.

Ruggero, M. A., Rich, N. C., and Recio, A. (**1992**). ''Basilar membrane responses to clicks,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, K. Horner, and L. Demany (Pergamon, Oxford), pp. 85–91.

Russell, I., and Schauz, C. (**1995**). ''Salicylate ototoxicity: Effects on the stiffness and electromotility of outer hair cells isolated from the guinea pig cochlea,'' Aud. Neurosci. **1**, 309–319.

Shera, C. A. (**1992**). ''Listening to the Ear,'' Ph.D. thesis, California Institute of Technology.

Shera, C. A. (**2001**). ''Frequency glides in click responses of the basilar membrane and auditory nerve: Their scaling behavior and origin in traveling-wave dispersion,'' J. Acoust. Soc. Am. **109**, 2023–2034.

Siebert, W. M. (**1968**). ''Stimulus transformations in the peripheral auditory system,'' in *Recognizing Patterns*, edited by P. A. Kolers and M. Eden (MIT, Cambridge), pp. 104–133.

Sondhi, M. M. (**1978**). ''Method for computing motion in a two-dimensional cochlear model,'' J. Acoust. Soc. Am. **63**, 1468–1477.

Zweig, G. (**1976**). ''Basilar membrane motion,'' in *Cold Spring Harbor Symposia on Quantitative Biology, Volume XL, 1975* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 619–633.

Zweig, G. (**1990**). ''The impedance of the organ of Corti,'' in *Mechanics*

and Biophysics of Hearing, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer, Berlin), pp. 362–369.
Zweig, G. (**1991**). ''Finding the impedance of the organ of Corti,'' J. Acoust. Soc. Am. **89**, 1229–1254.
Zweig, G., Lipes, R., and Pierce, J. R. (**1976**). ''The cochlear compromise,'' J. Acoust. Soc. Am. **59**, 975–982.
Zweig, G., and Shera, C. A. (**1995**). ''The origin of periodicity in the spectrum of evoked otoacoustic emissions,'' J. Acoust. Soc. Am. **98**, 2018–2047.

Intracochlear pressure measurements related to cochlear tuning

Elizabeth S. Olson
Physics Department, Princeton University, Princeton, New Jersey 08544

(Received 14 November 2000; accepted for publication 7 March 2001)

Pressure in turn one of the scala tympani (s.t.) was measured close to the basilar membrane (b.m.) and at additional positions as the pressure sensor approached and/or withdrew from the b.m. The s.t. pressure measured within about 100 μm of the b.m. varied rapidly in space at frequencies around the region's best frequency. Very close to the b.m. the s.t. pressure was tuned and scaled nonlinearly with sound level. The scala vestibuli (s.v.) pressure was measured at one position close to the stapes within seconds of the s.t. pressure and served primarily as a reference pressure. The driving pressure across the organ of Corti and the b.m. velocity were derived from the pressure data. Both were tuned and nonlinear. Therefore, their ratio, the specific acoustic impedance of the organ of Corti complex, was relatively untuned, and only subtly nonlinear. The impedance was inspected specifically for negative resistance (amplification) and resonance. Both were detected in some instances; taken as a whole, the current results constrain the possibilities for these qualities. [DOI: 10.1121/1.1369098]

PACS numbers: 43.64.Kc [LHC]

I. INTRODUCTION

Basilar membrane (b.m.) motion is tuned and nonlinear (Rhode, 1971). Probing the mechanical basis for b.m. tuning and nonlinearity was the major objective of this work. The experimental strategy was to find the basilar membrane's motion and local driving pressure over a wide range of frequencies including the best frequency of the observation point. The driving pressure (the pressure difference across the organ of Corti complex, defined to include the organ of Corti and the basilar and tectorial membranes) was estimated according to cochlear-mechanical theory by combining intracochlear pressure measurements in the scala tympani (s.t.) close to the b.m. with measurements of the scalar vestibuli (s.v.) pressure near the stapes. The b.m. velocity was estimated from measurements of the s.t. pressure gradient near the b.m. The primary observation was that the driving pressure was tuned and nonlinear to nearly the same degree as b.m. motion. This observation speaks for the global nature of tuning in the mammalian cochlea, which sets it apart from hearing organs in which local tuning mechanisms, e.g., electrical resonances in turtle hair cells (Crawford and Fettiplace, 1981), mechanical resonances of the stereociliary bundles of alligator lizard hair cells (Freeman and Weiss, 1990) have been observed or inferred to be dominant. It is notable that in previous results from the extreme basal region the driving pressure appeared to be tuned substantially less sharply than b.m. motion (Olson, 1998). This difference between the extreme base and turn one is discussed at the end of Sec. VI B, impedance results.

The specific mechanical impedance of the organ of Corti complex (OCC) is equal to driving pressure divided by b.m. velocity. The impedance was found and inspected specifically for negative resistance and a spring-mass resonance. These qualities are fundamental to many models of cochlear operation—the resonance to peak the cochlear traveling wave and then bring it to a full halt; negative resistance to enhance the peak at low levels (deBoer, 1984; Kolston, 2000). However, there is neither a consensus for these qualities in cochlear models nor decisive experimental evidence for them. The present results inform but do not resolve the matter. Negative resistance was observed but not in all measurements on healthy cochleae. A spring-mass resonance was apparent in the phase data of several experiments at frequencies just above the best frequency (b.f.) of the response, where it is expected to be. However, at frequencies above the b.f. the driving pressure difference was close to zero, which made the analysis of this frequency region susceptible to experimental inaccuracies. (Following common usage, a region's best frequency is the frequency for which b.m. motion peaks at low sound pressure levels.)

The pressure measurements here are unique in emphasizing spatial variations in pressure close to the sensory tissue. Intracochlear pressure close to the cochlear wall has been measured to investigate the forward and reverse transfer functions of the middle ear (Nedzelnitsky, 1980; Dancer and Franke, 1980; Puria and Rosowski, 1997; Magnan *et al.*, 1999; Puria *et al.*, 1997; Decory *et al.*, 1990; Olson and Cooper, 2000), the cochlear input impedance (Lynch *et al.*, 1982; Aibara *et al.*, 1999) and distortion products (Magnan *et al.*, 1997; Avan *et al.*, 1998). Intracochlear pressure was measured in several turns and/or both scalae by Dancer and Franke (1980) and Nedzelnitsky (1980).

II. METHODS

The methods of this study were similar to those of Olson (1998) and are described in more detail there.

A. Pressure sensor construction and calibration

A pressure sensor consists of a glass capillary (inner and outer diameters 100 and 170 μm) tipped with a gold-coated polymer diaphragm. Light from an LED is delivered via an optic fiber threaded into the capillary, and reflects from the diaphragm. The amount of light retuning to the optic fiber for transmission to a photodetector varies linearly with the pressure-induced bending of the diaphragm. The acoustic

impedance of the sensors is at least an order of magnitude larger than that of the cochlea measured at the stapes.

Sensors were calibrated in water and air following assembly and in water before and after every experiment except in unusual cases in which a sensor broke. The difference between before and after calibrations was similar to what was reported previously. In the current experiments, the difference ranged from 0 to 6 dB, except experiment 12-10-98, for which the s.t. sensor calibration changed by 10 dB. For analysis, the average of the before and after calibrations was used. Calibration uncertainty was most detrimental when calculating the pressure difference across the OCC, because then the difference between two pressures measured with different sensors was taken. In a few experiments the s.t. sensor was swapped into s.v. or vice versa at the end in order to check the relative sensitivity of the sensors. A minority of sensors was found to be temperature sensitive. Therefore, following assembly, sensors were screened for temperature sensitivity and were not used if the variation with temperature between 26 °C and 38 °C was more than 3 dB.

B. Animal preparation

Animal procedures were approved by the Princeton University Institute Animal Care and Use Committee. The experimental animals were young adult mongolian gerbils (*Meriones unguiculatus*) 40–65 g in weight. Ketamine (40 mg/kg) was administered to sedate the animal, followed by the anesthetic sodium pentobarbital (initial dose 60 mg/kg). Supplemental smaller doses of sodium pentobarbital were given when deemed necessary from a toe pinch response, typically every half hour. The animal was deeply anesthetized throughout the procedure and then sacrificed with an overdose of anesthetic. The animal core temperature was maintained at 38 °C with an animal blanket. A small heater was attached to the head holder. The bulla was widely open during all data collection.

C. Sound system and calibration

Stimuli were generated and responses collected with a Tucker Davis Technologies DD1 using a 6.48 μs sampling period. The response to a click was collected and averaged with a LeCroy digital oscilloscope. Sound stimuli were produced with a Radio Shack tweeter and delivered to the ear canal via a closed sound system. In order to calibrate the stimulus, at the beginning of every experiment a pressure sensor was inserted into the ear canal via a small hole that was made in the bulla just in front of the tympanic membrane. The system was calibrated at up to 62 frequencies. In previous experiments the calibration hole was covered during and after calibration. In the current experiments that procedure was not followed without appreciable difference.

D. Intracochlear pressure measurements

In order to access the s.v. a hole just large enough for a pressure sensor was hand drilled through the bone basal to the oval window. The s.v. sensor was held in a micromanipulator and its tip was inserted 100–200 μm into the s.v. To access turn one of the s.t. a similar hole was hand drilled above the round window (r.w.) opening. Figure 1(a) shows the positions of the s.t. and s.v. holes. The s.t. sensor was positioned so that it was pointing as closely as possible toward the cochlear apex. An excised temporal bone and anatomical landmarks served as guides for positioning the s.t. sensor, which was tricky. A hole too close to the stapedial artery made it impossible to correctly angle the sensor and instead of the b.m., the spiral lamina was approached. This resulted in greatly reduced pressure gradients. A hole too far from the stapedial artery damaged the spiral ligament, ending an experiment. When the hole was all right, often several approaches were made at slightly different angles in order to change the longitudinal and/or lateral position of the sensor on the b.m. by 50–100 μm. In one of the experiments presented here, the s.t. pressure was measured in the extreme base. To access the extreme base of the s.t. the sensor was inserted through the r.w. opening following removal of the r.w. membrane.

The s.t. sensor was held in a micromanipulator capable of both manual and motorized positioning. The sensor was guided into the hole manually, and advanced within the hole using the motorized manipulator. Figure 1(b) illustrates the

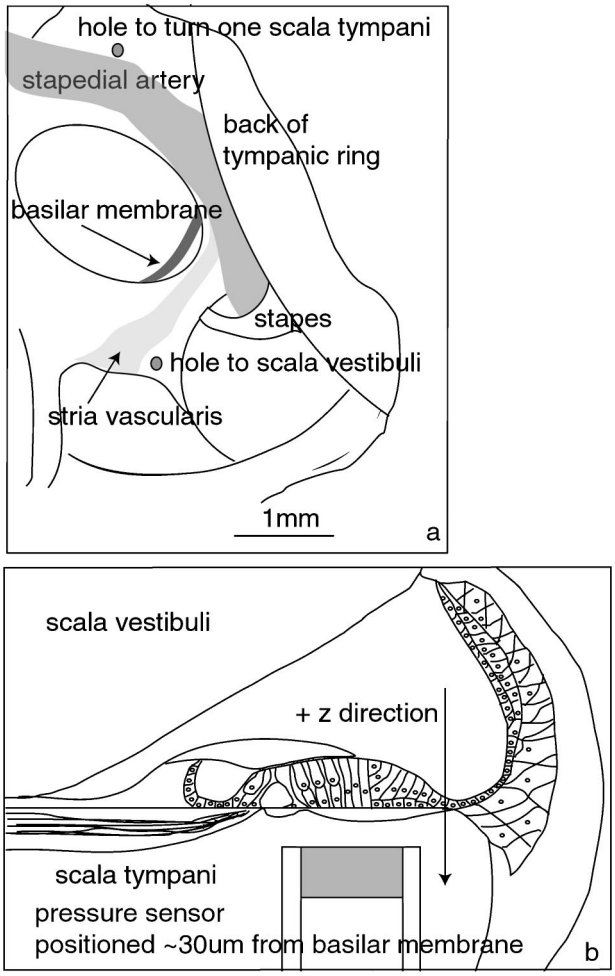


FIG. 1. (a) View of the cochlea during experiments. Not shown are the pressure sensors which were inserted into the s.v. and s.t. holes and the CAP electrode, which was positioned on the bone of the r.w. opening. (b) Idealized to-scale drawing of a pressure sensor positioned close to the b.m. in the s.t.

sensor positioned close to the b.m. When close to the b.m., measurements were usually spaced by 10 or 20 μm in the direction along the sensor axis (z axis). In later experiments the s.t. sensor was held in a piezoelectric bimorph assembly similar to that described in Olson and Mountain (1991), which in turn was held in the motorized manipulator. The bimorphs were driven with DC voltage in order to advance and retract the s.t. sensor in the direction along its axis. In these experiments at each frequency and level the s.t. pressure was measured consecutively at two positions separated by 12 μm. This procedural change was made in order to reduce the effect of slow changes (for example, in fluid level) on the calculation of fluid velocity. The distance from the b.m. was determined by touching it with the sensor, which produced a bouncy sensor response on the oscilloscope.

The pressure stimuli were pure tones, 32 ms in duration. The number of averages taken ranged from 20 to 200. The responses were stored and later analyzed via fast Fourier transform to find the magnitude and phase at the stimulus frequency. The initial 5 ms of the response was truncated before analysis in order to exclude the transient response. In one of the presented experiments the pressure stimulus was a click produced by driving the earphone with a 10 μs voltage pulse. Because of the frequency response of the speaker the acoustic pulse in the ear canal was longer than 10 μs.

E. Compound action potential

A silver wire electrode insulated to its tip was positioned on the bone near the r.w. This was used to measure the compound action potential (CAP) response to tones, as a monitor of cochlear condition (Johnstone *et al.*, 1979). CAP stimuli were tone bursts, 3 ms in duration. They ranged from 20 to 80 dB SPL in 10 dB increments and from 0.5 to 40 kHz at 10 or more frequencies. (SPL is decibels re 20 μPa.) Typically 60 responses were averaged. The polarity of alternate tones was reversed to reduce the cochlear microphonic in the averaged response. The averaged response was displayed on line and stored. ''Threshold'' CAP was defined as the visual threshold, 5 to 10 μV peak-to-peak.

III. RESULTS

In all the results, phases are shown referenced to the s.v. pressure phase measured at the stapes within seconds of each s.t. pressure measurement. The s.v. pressure at the stapes can be considered as the input pressure of the cochlea.

IV. GROUPED RESULTS

A. General description

In Fig. 2 scala tympani pressure magnitude and phase is shown from 14 turn one experiments. The stimulus level was 80 dB SPL in all cases, and these were initial data, taken with the s.t. sensor ~150 μm within the s.t., ~300 μm from the b.m. The average s.t. pressure is also shown, and the average s.v. pressure close to the stapes from these experiments. The character of the s.v. pressure was described previously (Olson, 1998), here it suffices to repeat that above 1

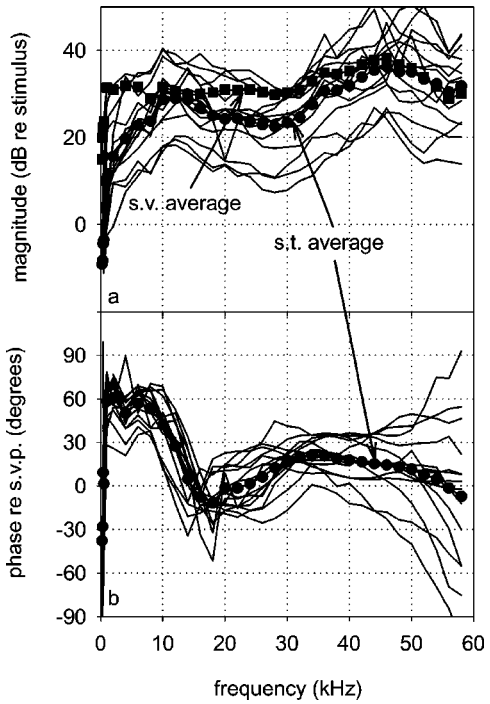


FIG. 2. Turn-one scala tympani pressures far from the b.m., 14 experiments. Also shown are the average of these measurements and the average of the s.v. pressures near the stapes from the same experiments. The stimulus level was 80 dB SPL in the ear canal. (a) Magnitude *re* stimulus level in ear canal. (b) Phase relative to the simultaneously measured pressure in s.v. near the stapes.

kHz the s.v. pressure scaled linearly (except as discussed at the end of Sec. IV B 2), and that it was nearly flat with frequency, with a gain relative to the ear canal pressure of ~30 dB. The s.t. phase is shown relative to the ~ simultaneously measured pressure in the s.v. Figure 3 shows the Fig. 2 s.t. data from the experiments that will be used in the impedance analysis, and include data from an extreme basal experiment. Referring to Fig. 2, the s.t. pressure was substantially smaller than the s.v. pressure at frequencies below 10 kHz, and had two broad peaks, centered at ~12 and 45 kHz. The lower frequency peak is just under the best frequency of this region. This peak and the phase drop between 10 and 20 kHz are likely manifestations of the traveling wave. Similar behavior occurred between 20 and 30 kHz in the extreme base (curve 2-26-97 of Fig. 3). The 45 kHz peak was also present in the s.v. pressure and might be due to a standing wave in the ear canal described in ''sound system calibration'' in Olson (1998).

The behavior of the extreme basal s.t. pressure (2-26-97 of Fig. 3) at frequencies well below the b.f. can be understood in terms of a lumped element model. The model is shown in Fig. 4 with element values in the caption. In the usual way (Beranek, 1954) mass is treated as an inductor and stiffness as a capacitor. m_v and m_t are the fluids in the s.v. and s.t. in the region between the cochlear windows. r_c is the ''transmission line'' resistance of the cochlea (Zwilslocki, 1965). The capacitor represents the stiffness of the basilar membrane in the immediate vicinity of the windows. The model is kept very simple; e.g., the mass and resistance associated with fluid flow through the helicotrema are not in-

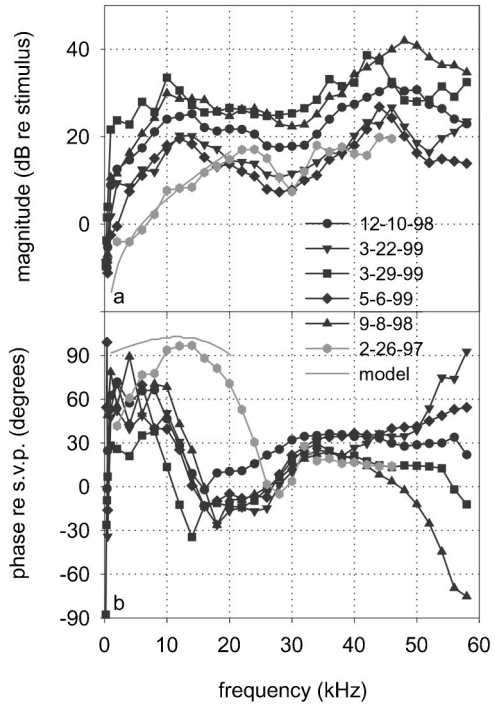


FIG. 3. Turn-one scala tympani pressures far from the b.m. Similar data as in Fig. 2, but only from those experiments used in the impedance analysis, and including an extreme basal measurement. The response of the lumped parameter model in Fig. 4 is included to shed light on the low frequency results.

cluded (Lynch *et al.*, 1982). Figure 3 shows the extreme basal s.t. pressure that the model predicts at a depth ~100 μm within the fluid of the s.t. In particular, the model explains how the interaction of fluid mass and basilar membrane stiffness can cause the phase of s.t. pressure re s.v. pressure to increase to a value greater than 90°. The model does not apply to the turn-one measurements, which were made some distance along the ''transmission line.'' (To model the turn one measurements r_c must be partly expanded into inductive and capacitive elements so the measurement position could be placed some distance along the transmission line.) Nevertheless, the well-below-b.f. turn-one results are fairly similar to the basal result, and can be loosely interpreted in a similar way.

B. Experimental uncertainties and perturbations

1. Exploring the variability

The spread of s.t. values in Fig. 2 is likely due to a combination of experimental conditions and calibration inaccuracies. The measurements rely on both the s.t. sensor and the ear canal sensor, so much of the ±12 dB variability in magnitude could be due to inaccurate calibration. An influential experimental condition is the fluid level in the r.w. opening. A higher fluid level caused an increase in the turn one s.t. pressure at frequencies above the b.f. In one experiment (7-13-98) the s.t. pressure around 40 kHz increased by ~10 dB when the r.w. opening was filled relative to when it was drained. At higher frequencies the effect was slightly less, and at frequencies below 23 kHz, the changes were ~2 dB. The s.t. pressure phases in Fig. 2(b) fan out at frequencies above 40 kHz. A 45° spread at 40 kHz grew to a 180° spread at 58 kHz. This spread *cannot* be traced to sensor variability. Comparing calibrations within experiments 3-22-99, 9-8-98, and 4-5-99 [which produced the most extreme phase in Fig. 2(b)] revealed a s.t. sensor calibration–s.v. sensor calibration difference of at most 8° at frequencies up to 40 kHz, and an overall maximum difference of 27°. The reason for the divergent high frequency phases is not known.

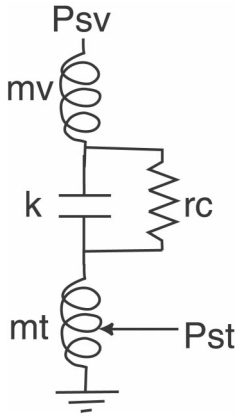


FIG. 4. Simple lumped parameter model of extreme base of cochlea at frequencies well beneath the region's best frequency. $P_{s.v.}$ is the pressure measured in the s.v. near the stapes. $P_{s.t.}$ is the extreme basal scala tympani pressure, measured about 100 μm within the s.t. fluid. The round window membrane was removed for measurements of extreme basal s.t. pressure and the sensor entered the s.t. through the r.w. opening. m_v and m_t are the fluid masses within the s.v. and s.t. in the basal region. They are taken to be equal with a value of 3.3×10^5 mks acoustic ohm s. [An acoustic ohm =pressure/volume velocity$=N/(m^5/s)$.] This value is reasonable given the depth of the fluid column between the oval and round windows (~1 mm), and the sizes of the windows. r_c is the ''transmission line'' resistance of the cochlea, and its value of 1.7×10^{11} ohm is based on measurements of gerbil cochlear input impedance (Olson and Cooper, 2000). For comparison, Lynch *et al.* (1982) found a value of 1.2×10^{11} for r_c in cat. k is the stiffness of the OCC in the region between the oval and round windows. The stiffness value was found using the 4 Pa/nm value reported in Olson (1998) for the extreme basal OCC stiffness. Similar values were reported (as b.m. compliance) from a number of sources in Table IV of Ruggero *et al.* (1990), and the results presented later in this paper are also in reasonable accord with this value. The width of the OCC in this region is ~0.2 mm, and a length of ~1 mm is in the vicinity of the stapes. From these, the OCC stiffness was estimated as $\sim2\times10^{16}$ ohm/s.

2. Perturbative effect of holes and sensors in the cochlea

The s.v. hole was expected to perturb cochlear mechanics more than the s.t. hole. This is because the s.t. hole was just above the r.w. and the effect of a small hole near such a large opening is expected to be minor relative to the effect of the s.v. hole near the stapes. In order to evaluate the effect of the s.v. hole, in several experiments the s.t. hole was made first and the pressure ~150 μm within the s.t. was measured before and after drilling the s.v. hole and inserting the s.v. sensor. The results are shown in Fig. 5, where the after–before differences in s.t. pressure magnitude and phase are shown for three experiments. The differences were rarely more than 2 dB in magnitude or 15° in phase.

The CAP response was also used to gauge the effect of introducing sensors into the cochlea. The CAP thresholds often increased at all frequencies after making the holes and

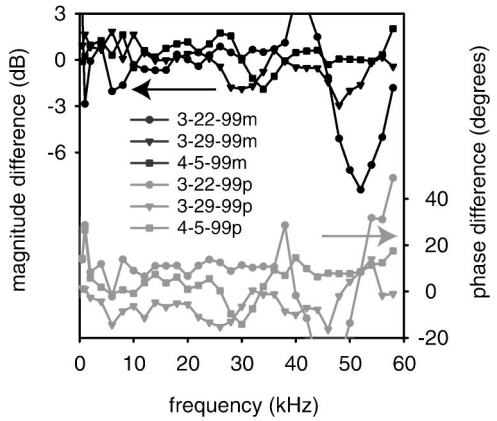


FIG. 5. Perturbation study: Change in the s.t. pressure after drilling the s.v. hole and inserting the s.v. sensor. Results from three experiments are shown. The magnitude changes (black lines) use the left axis, the phase changes (gray lines) use the right. The changes were small.

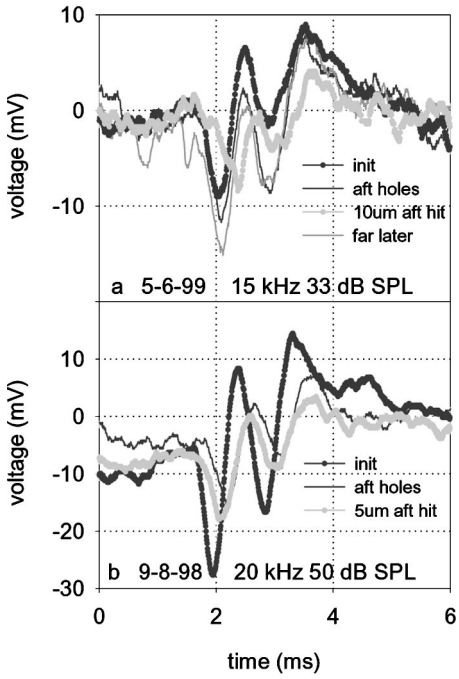


FIG. 6. Perturbation study using CAP response. CAP responses to a tone pip with frequency close to the b.f. of the longitudinal location under study were measured at various times. ''Init'' was the initial CAP, measured prior to drilling cochlear holes. ''Aft holes'' was after drilling the s.t. and s.v. holes and inserting the sensors. ''5 or 10 μm aft hit'' was the response when the s.t. sensor was 5 or 10 μm from the b.m., just after tapping it. ''Far later'' was after retracting the s.t. sensor. (a) Expt. 5-6-99. Here, the CAP response did not change upon introducing the sensors to the cochlea but was reversibly reduced when the s.t. sensor was close to the b.m. (b) Expt. 9-8-98. Here, the CAP response was reduced upon introducing the sensors to the cochlea but did not change further when the s.t. sensor was close to the b.m.

introducing the sensors. When the s.t. sensor was close to the b.m., further changes in the CAP response to tones at frequencies close to the b.f. sometimes occurred. Both these effects are illustrated in Fig. 6, which shows CAP responses to a tone close to the b.f. In the experiment of Fig. 6(a) (5-6-99), very little change occurred in the CAP following sensor introduction, but the CAP was reduced when the sensor was 10 μm from the b.m. The change was reversed when the sensor was retracted. In the experiment of Fig. 6(b) (9-8-98), the CAP response decreased after introducing the sensors, but was unchanged when the s.t. sensor was close to the b.m.

Another observation which bears on the influence of the sensor to cochlear mechanics is that the s.v. pressure measured at the stapes sometimes changed when the s.t. sensor was close to the b.m. The changes were largest at small sound levels and at frequencies close to the b.f. These changes were small, 3 dB at most, and reversible.

The conclusion drawn from these observations is that making holes and introducing sensors into the cochlea did not cause an overall reduction in the intracochlear pressure, but *usually* did traumatize the cochlea, leading to an overall decrease in sensitivity. When close to the b.m., the sensor sometimes perturbed cochlear mechanics.

V. SCALA TYMPANI PRESSURE VS POSITION

A series of pressure measurements made with spatial changes solely in the direction along the axis of the sensor is referred to as an approach. A ''run'' is a single run through of a data collection program, which comprises a series of frequencies and levels. When the piezoelectric positioner was in use, during a single run these data were collected at two positions in the s.t. which were separated by 12 μm in the z direction [see Fig. 1(b)]. In this section, approaches from the two best turn-one experiments, 9-8-98 and 5-6-99, are shown. Experiment 9-8-98 was performed before the piezoelectric positioner was in use; 5-6-99 did use the positioner. These experiments had relatively strong CAP responses, relatively strong compressive nonlinearity in s.t. pressure, stable fluid levels and repeatability of measurements over hours, especially 9-8-98. The basic observations are (i) the s.t. pressure close to the b.m. was tuned and compressively nonlinear, (ii) in some approaches the pressure variations close to the b.m. suggested that the distortion of the moving OCC was level dependent, and (iii) the pressure was composed of a sum of a traveling wave component which varied rapidly in space and a compressive component which varied very little in space. The last point is supported with an approach from a more recent experiment in which the pressure response to a click was measured.

This paragraph provides a brief guide to the figures in this section. Figure 7, from experiment 9-8-98, shows the s.t. pressure gain *re* ear canal pressure and the s.t. pressure phase re the simultaneously measured s.v. pressure at stimulus levels of 50 dB SPL (left panels) and 80 dB SPL (right panels). The measurements were made at distances ranging from 7 to 322 μm from the b.m. The complete approach included measurements at 13 positions. To improve figure clarity the data are not shown from every position. Figure 8 compares pre- and post-mortem data from 9-8-98. In Fig. 9 the results from an approach from experiment 5-6-99 are shown. For clarity only one of the intra-run positions (the closer of the two) is included in this figure. Responses were collected at 40, 50, 65, and 85 dB SPL, with the lower level responses collected over a narrower range of frequencies. The 50 dB SPL (left panels) and 85 dB SPL (right panels) results are in Fig. 9. Figure 10 shows magnitude data from the closest 5-6-99 run,

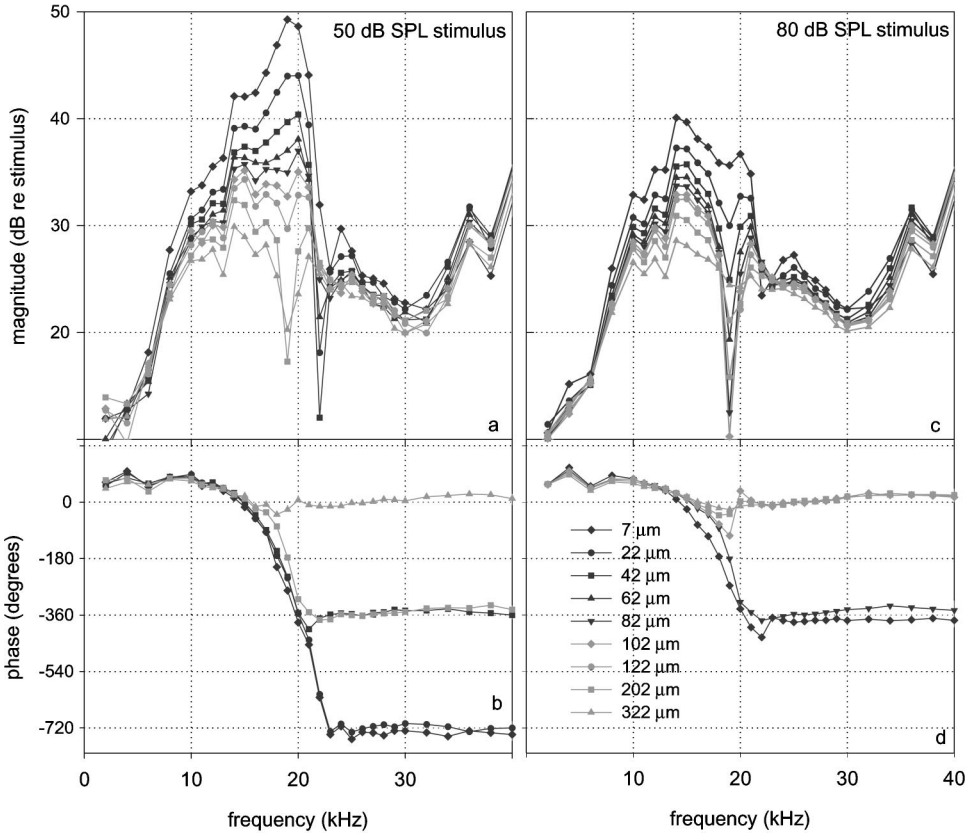


FIG. 7. The s.t. pressure as the s.t. sensor approached the b.m. Expt. 9-8-98. The key indicates the distance of the sensor from the b.m. To improve the clarity of the figures the phase data are shown at fewer positions than the magnitude data. Magnitudes are shown relative to the stimulus level in the ear canal; phases are relative to the simultaneously measured s.v. pressure. (a) Magnitude, 50 dB SPL stimulus. (b) Phase, 50 dB. (c) Magnitude, 80 dB. (d) Phase, 80 dB.

which is not included in Fig. 9 for the reason described in Sec. V C, at the two intra-run positions and all SPLs. Figure 11 shows the click response approach of experiment 4-25-00.

A. Nonlinearity

The similarity of Figs. 7 and 9 attests to the repeatability of the results in fairly healthy preparations. The results in Figs. 7, 9, and 10 show a moderate degree of compressive nonlinearity. For example, in Fig. 10, over the 40 to 85 dB range of stimulus levels the nonlinearity at the b.f. (18 kHz) was 20 dB. Nonlinearity began approximately a half octave below the b.f., at 12 kHz, and extended to just above the b.f., at 22 kHz. These nonlinear characteristics are similar to those reported for basal b.m. motion, although the degree of nonlinearity here was smaller than that of b.m. motion in the healthiest preparations (Ruggero *et al.*, 1997; Cooper, 1998; deBoer and Nuttall, 2000; Rhode and Recio, 2000). Based on the CAP thresholds even the best cochleae of this report were compromised slightly; this might account for the moderate level of compression. The data of Fig. 8 were taken just after those of Fig. 7, at a position 67 μm from the b.m. The pre-mortem data were taken, the animal was sacrificed with anesthetic, and the post-mortem data were taken minutes later. Nonlinearity disappeared post-mortem.

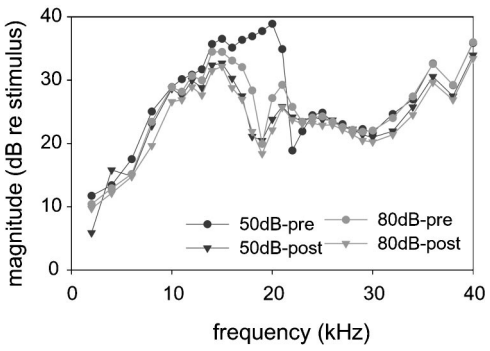


FIG. 8. The s.t. pressure magnitude re stimulus level measured at a distance 67 μm from the b.m. premortem and a few minutes postmortem. Expt. 9-8-98. Nonlinearity disappeared postmortem.

B. Suggestion of level dependent distortion of the organ of Corti

In the data of Fig. 9 the closest measurement *shown* was 20 μm from the b.m. The closest data run, 10 μm from the b.m., is not shown in Fig. 9 because, anomalously, the pressure was smaller at 10 than at 20 μm. This is illustrated in Fig. 10, which shows the pressure at the two intra-run positions of the closest run. The positions are 10 and 22 μm from the b.m. At 40 and 50 dB SPL and frequencies between the onset of nonlinearity and the b.f., the pressure at 22 μm was substantially bigger than at 10 μm. At 65 dB the effect was present but smaller. At 85 dB SPL the usual behavior, bigger pressure at the closer location, was observed. It is not difficult to imagine how the pressure at the closer location could be smaller than at the further location: For a simple beamlike radial profile of b.m. motion the pressure is greatest at the radial center (Steele and Taber, 1979). Ideally, the sensor approaches the radial center of the b.m. perpendicularly, as in Fig. 1(b). If the sensor approach is not quite perpendicular its degree of centering will change as it approaches. Then, when very close to hitting the b.m. the sensor could move from a region of relatively high pressure (more centered) to a

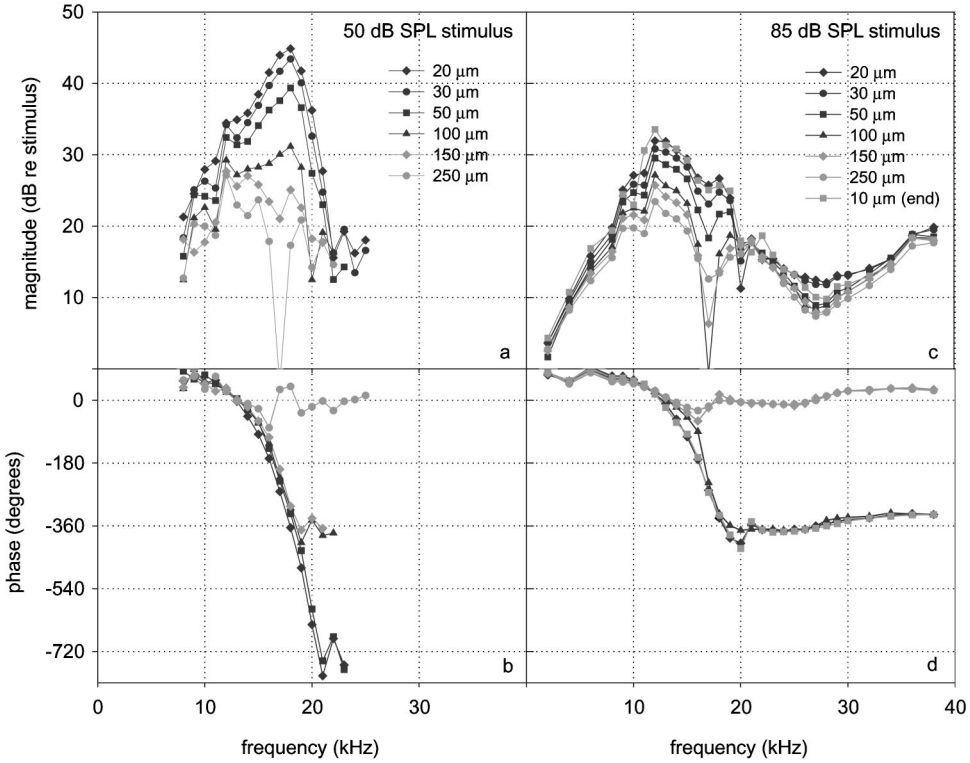


FIG. 9. The s.t. pressure as the s.t. sensor approached the b.m. Expt. 5-6-99. The key indicates the distance of the sensor from the b.m. To improve the clarity of the figures the phase data are shown at fewer positions than the magnitude data. Magnitudes are shown relative to the stimulus level in the ear canal; phases are relative to the simultaneously measured s.v. pressure. (a) Magnitude, 50 dB SPL stimulus. (b) Phase, 50 dB. (c) Magnitude, 85 dB. (d) Phase, 85 dB.

region of lower pressure (less centered). This effect was sketched in Olson (2000). What was intriguing about the reversal in the pressure gradient was that it sometimes depended on stimulus level, as in Fig. 10. This observation suggests that the radial profile of b.m. motion—in other words, the shape the b.m. took as it moved—underwent level dependent changes. The level dependent distortion could arise via a force generated from within the OC (e.g., Mountain and Hubbard, 1989; Kolston, 1999) whose strength was level dependent. In studies of the radial profile of basal b.m. motion, Cooper (2000) found a unimodal, beamlike profile of gerbil basal b.m. motion whose shape was mildly level dependent; the trimodal radial profile reported by Nilsen and Russell (1999) from the guinea pig base was also mildly level dependent. In the current study, level dependent reversals in pressure gradient occurred in several experiments, although not in experiment 9-8-98 or in a second approach of experiment 5-6-99. [Finally, the pressure sensor could be influencing the level dependence of the reversals. The perturbation of the sensor depends on the relative impedances of the sensor and the OCC. The frequency region just below the b.f. is implicated in cochlear amplification (e.g., deBoer and Nuttall, 2000 and see below), so the impedance in this region might be level dependent. Therefore, the perturbative effects of the sensor might be level dependent in this region.]

C. Multi-component nature of intracochlear pressure

Many aspects of the s.t. pressure reflect its being the sum of a traveling wave and a compressive wave. As background to this view: *The traveling wave pressure is produced by and produces the traveling wave motion of the* OCC. *It is largest near the* OCC *and spreads with decreasing amplitude into the scalae. The compressional pressure is produced by the compression of the cochlear fluid by the motion of the stapes and fills the cochlea approximately uniformly*. In this interpretation: (i) The phase accumulated at positions close to the b.m. because the traveling wave component is dominant there [Figs. 7 and 9, (b) and (d)]. (iii) The phase accumulated more at low stimulus levels because nonlinearity makes the traveling wave relatively stronger at low levels [Figs. 7 and 9, (b) and (d)]. (iii) The phase did not accumulate at frequencies above the peak because the traveling wave is small (perhaps nonexistent) relative to the compressional pressure there [Figs. 7 and 9, (b) and (d)]. (iv) Spatial variations, which register fluid motions, were large at frequencies of the peak because the fluid motions of the traveling wave are substantial (Figs. 7 and 9). (v) Spatial variations were small at frequencies above the peak because the fluid motions associated with the compressional wave are very small (Fig. 7). [In Fig. 9(c) the pressure *did* vary in space at frequencies above the peak. However, these changes were at least in part actually *time* dependent changes, as can be seen by comparing data collected with the sensor close to the b.m. at the beginning (20 μm position) and end of the approach (10 μm-end position).] (vi) Notches were caused by cancellation between traveling wave and compressional wave components [Figs. 7 and 9, (a) and (c)] (Cooper and Rhode, 1996).

Figure 11 also speaks for the two component nature of the pressure. It shows the response to a click (experiment 4-25-00, maximum level in the ear canal 77 dB) measured at several distances from the b.m. in the s.t. The initial peak of the response, which is presumably the compressional wave, did not change with position. In contrast, the ''slow'' ringing response, presumably the traveling wave, became more and more pronounced as the b.m. was approached.

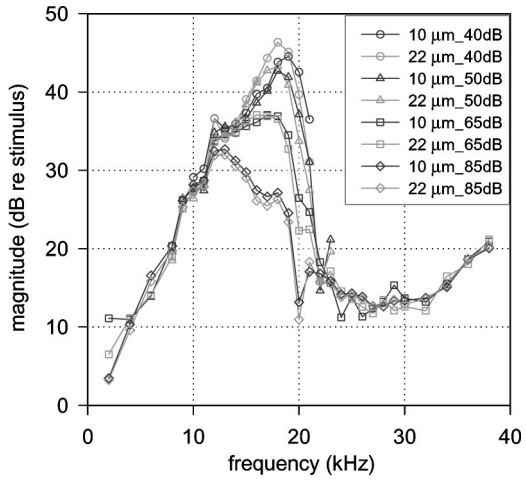


FIG. 10. The s.t. pressure magnitude *re* stimulus level 10 and 22 μm from the b.m. Expt. 5-6-99. These measurements were from the same approach as in Fig. 9; this was the closest position of that approach. The results here show level-dependent reversals in the relative magnitudes of the closer (10 μm) and farther (22 μm) pressures. Responses at 40, 50, 65, and 85 dB SPL are shown. At 85 dB SPL the pressure at the closer position was larger, as is usually the case. At 40, 50, and 65 dB SPL the pressure at the closer position was smaller over some portion of the left side of the response peak.

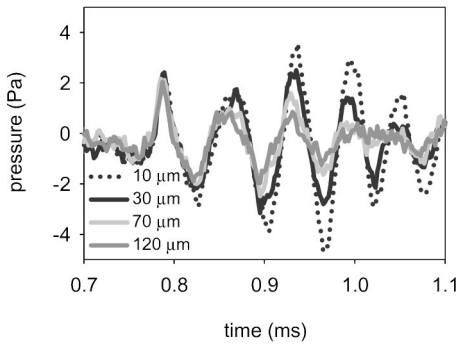


FIG. 11. The s.t. pressure response to a click measured at several distances from the b.m. Expt. 4-25-00. The click was produced by driving the earphone with a 10 μs voltage pulse, but filtering in the sound system produced a longer stimulus as measured in the ear canal. The key indicates the distance of the s.t. sensor from the b.m. The prolonged ringing response, presumably the traveling wave pressure, grew as the b.m. was approached, whereas the initial pressure peak, presumably the compressive pressure, remained almost the same.

VI. IMPEDANCE OF THE ORGAN OF CORTI COMPLEX

The specific acoustic impedance of the OCC (Z_{OC}) was derived from the pressure data. Z_{OC} is defined as the pressure across the OCC (ΔP_{OC}) divided by the *z* component of b.m. velocity ($v_{b.m.}$). [The *z* axis was defined in Fig. 1(b).]

In a passive system, Z_{OC} depends on the material properties and geometry of the OCC. For example, up to frequencies through the peak the passive part of Z_{OC} is likely, in simple terms, a combination of stiffness and damping. In an active system in which a force generator is present within the organ of Corti, that force, F_{active}, adds a term to the passive impedance which is equal to (F_{active}/area on which force acts)/$v_{b.m.}$. The Z_{OC} that these experiments find when the measured ΔP_{OC} is divided by the measured $v_{b.m.}$ is the sum of the passive part and the active part (deBoer and Nuttall, 2000). In cochlear models the active part is most successful at producing realistic b.m. tuning when it has the character of a negative damping that is large enough to cause the net damping to be negative over a limited region somewhat basal to the peak (e.g., deBoer, 1983; Neely and Kim, 1986; Kolston, 2000). *In the current experiments, performed at one place and many frequencies, this negative resistance would appear as a negative real part of the impedance at frequencies slightly below the b.f.*

In the classic traveling wave/resonant model of cochlear operation, the traveling wave, produced by the interaction of fluid inertia and OCC stiffness, ripples down the cochlear spiral. The decreasing stiffness of the OCC causes the traveling wave to slow and grow. At the point that the OCC mass begins to dominate its stiffness the traveling wave stops (e.g., Peterson and Bogart, 1950; Lighthill, 1981). The spring–mass transition is expected to occur slightly apical of the peak of the traveling wave. *Therefore, below and through the b.f. the imaginary part of the impedance is expected to be that of stiffness, and the spring–mass resonance will be looked for at frequencies slightly above the b.f. If the spring–mass resonance exists the imaginary part of the impedance will make a transition from stiffness dominated (negative) to mass dominated (positive) at the resonant frequency*.

A. Derivation

The analytic method for deriving ΔP_{OC} and $v_{b.m.}$ from the pressure measurements was described in Olson (1998). It is summarized here.

1. Basilar membrane velocity

The calculation of $v_{b.m.}$ uses one s.t. pressure measurement close to the b.m. (P_b) and a second s.t. pressure measurement a small distance from the first (P_a). The line that connects the two points of measurement is defined as the *z* direction, and it points away from the b.m., from the s.v. towards the s.t. [Fig. 1(b)].[1] At frequencies above a few kHz the *z* component of fluid velocity can be written very simply using these two pressures:

$$v_z \approx i(P_a - P_b)/(\omega \rho \Delta z). \tag{1}$$

In the expression, ω is the angular frequency, ρ is the density of the cochlear fluid, and Δz is the distance between the two pressure measurements. The fluid very close to the b.m. is expected to move with it, so when P_b is very close to the b.m. the fluid velocity approximates b.m. velocity. Then

$$v_{b.m.} \approx i(P_a - P_b)/(\omega \rho \Delta z). \tag{2}$$

2. Pressure across the organ of Corti complex

ΔP_{OC} is the pressure close to the OCC in the s.v. ($P_{s.v.-OCC}$) minus the pressure close to the b.m. in the s.t. (P_b). What was *measured* was the pressure in the s.v. near the stapes ($P_{s.v.}$) and the pressure in the s.t. close to the b.m. (P_b). With the simplifying assumptions that the cochlea is symmetric and the pressure at the r.w. is zero, $P_{s.v.} - P_{s.v.-OCC} = P_b - 0$, and

$$\Delta P_{OC} \approx P_{s.v.} - 2P_b \tag{3}$$

(see Footnote 2).

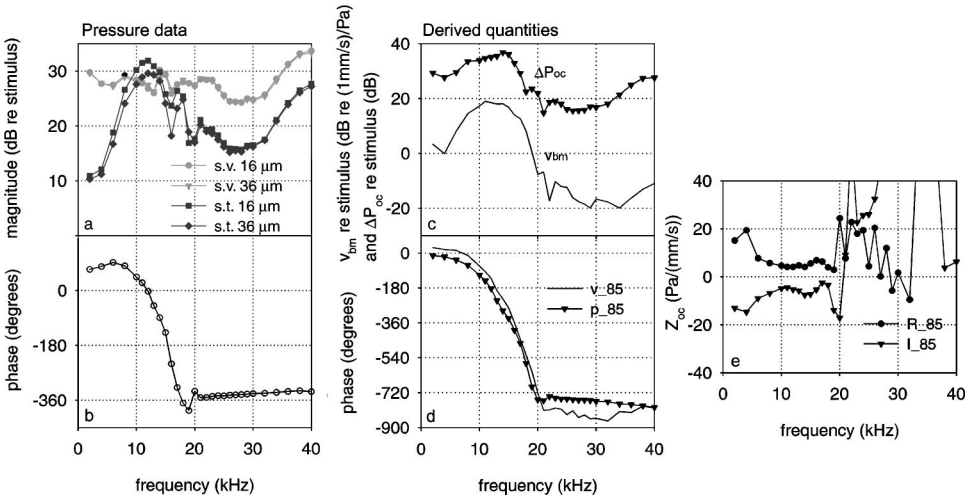


FIG. 12. Pressure data and derived quantities, 12-10-98. This was a nearly linear cochlea and only 85 dB SPL data are shown. (a) and (b) The s.t. and s.v. pressures for impedance calculation. Magnitude is shown relative to the stimulus level in the ear canal; phase (s.t. only) is relative to the simultaneously measured s.v. pressure. Distances in the key refer to the distance between the b.m. and the s.t. sensor. (a) Magnitude. (b) Phase. (c)–(e) ΔP_{OC}, $v_{\mathrm{b.m.}}$, and Z_{OC}. The usual calculation for ΔP_{OC} was used. (c) Magnitude of $v_{\mathrm{b.m.}}$ and ΔP_{OC} (*re* stimulus level). (d) Phase of $v_{\mathrm{b.m.}}$ and ΔP_{OC} (*re* simultaneously measured s.v. pressure). (e) Real and imaginary parts of Z_{OC}.

3. Specific acoustic impedance of the organ of Corti complex

$$Z_{\mathrm{OC}} = \Delta P_{\mathrm{OC}} / v_{\mathrm{b.m.}}. \tag{4}$$

B. Impedance results

Impedance results from six experiments are shown in Figs. 12–25. Impedances calculated for an additional six experiments are not shown because the results were noisier (due to time dependent variations or lower sensor sensitivity) and therefore less revealing than the six experiments that are presented. In each case the pressure measurements that went into the impedance calculation, the derived ΔP_{OC} and $v_{\mathrm{b.m.}}$ and the real and imaginary parts of Z_{OC} are plotted. Two introductory comments are in order: (1) The case-study presentation is fitting for communicating the impedance results because the meaning and authority of a particular Z_{OC} result is closely linked to the pressure measurements that generated it. Showing several case studies was necessary to demonstrate repeatability, variability, passive versus active, and turn one versus extreme base. (2) The weakest part of the analysis is the calculation of ΔP_{OC}. It was based on a symmetric cochlea, which could be an oversimplification. Further, it subtracts responses measured with two sensors, so even small calibration errors will introduce large errors when $P_{\mathrm{s.v.}} \approx 2P_b$. This weakness has been addressed by doing variations on the calculation, by finding ΔP_{OC} as $0.5P_{\mathrm{s.v.}} - 2P_b$ (contribution of s.v. pressure halved relative to usual calculation) and as $2P_{\mathrm{s.v.}} - 2P_b$ (contribution of s.v. pressure doubled) in a few cases. These variations show how a 6 dB calibration difference affects the impedance results and point out robust and fragile aspects of the results.

1. General conclusions

Taken as a whole the results lead to some general conclusions. From the ΔP_{OC} and $v_{\mathrm{b.m.}}$ plots: (i) The accumulation of the phases of both ΔP_{OC} and $v_{\mathrm{b.m.}}$ indicates that both are part of the cochlear traveling wave. (ii) Whether or not nonlinearity was present (i.e., in active and passive cochleae), both ΔP_{OC} and $v_{\mathrm{b.m.}}$ were tuned. (iii) When nonlinearity was in evidence it was usually present to nearly the same degree in both ΔP_{OC} and $v_{\mathrm{b.m.}}$. (An exception to this emerges from the analysis 9-8-98-I-double, as discussed below.) The Z_{OC} plots illustrate the relative and absolute sizes of the real and imaginary parts of the impedance, and indicate where it was stiffness dominated (imaginary part negative), mass dominated (imaginary part positive), and where the resistance was negative (real part negative). The Z_{OC} plots are most reliable in the broad region of the peak,

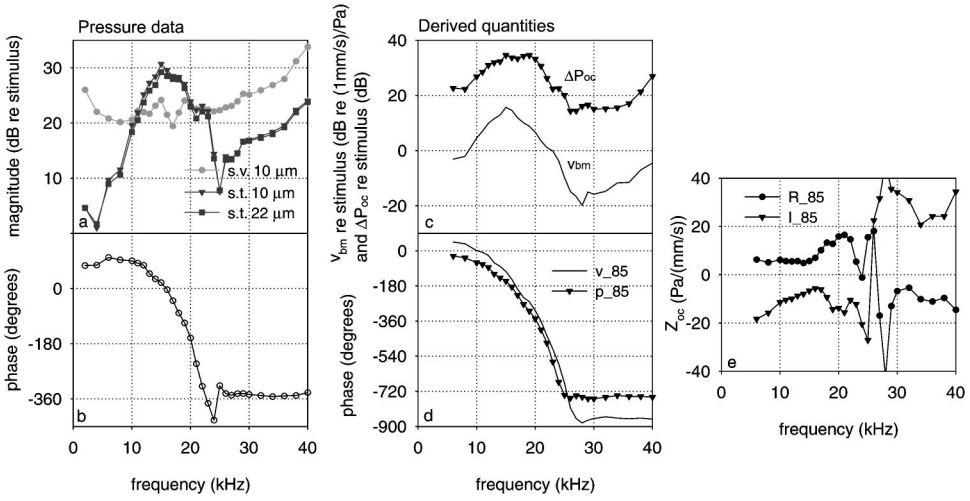


FIG. 13. Pressure data and derived quantities, 3-22-99. This was a nearly linear cochlea and only 85 dB SPL data are shown. (a) and (b) The s.t. and s.v. pressures for impedance calculation. Magnitude is shown relative to the stimulus level in the ear canal; phase (s.t. only) is relative to the simultaneously measured s.v. pressure. Distances in the key refer to the distance between the b.m. and the s.t. sensor. (a) Magnitude. (b) Phase. (c)–(e) ΔP_{OC}, $v_{\mathrm{b.m.}}$, and Z_{OC}. The usual calculation for ΔP_{OC} was used. (c) Magnitude of $v_{\mathrm{b.m.}}$ and ΔP_{OC} (*re* stimulus level). (d) Phase of $v_{\mathrm{b.m.}}$ and ΔP_{OC} (*re* simultaneously measured s.v. pressure) (e) Real and imaginary parts of Z_{OC}.

roughly 8–23 kHz (10–35 kHz for the extreme basal experiment 2-26-97), because there both the s.t. pressure and the spatial variations of the s.t. pressure were large. From these reliable regions of the plots, the additional general conclusions can be drawn: (iv) The imaginary part was negative (stiffness dominated) from low frequencies right up to and through the b.f. (v) In the ~5–10 kHz region (8–16 kHz for 2-26-97) the magnitude of the imaginary part usually decreased as frequency increased, as is expected for a stiffness. (vi) Overall, the real and imaginary parts were similar to each other in size; at 8–10 kHz they were mostly within 5–20 Pa/(mm/s) for the turn one experiments, and ~20 Pa/(mm/s) for the extreme basal experiment (2-26-97).

An additional general conclusion derives from the basic pressure data that introduces each case study, and bears on the calculation for ΔP_{OC}. Recall that ΔP_{OC} is found as $P_{\mathrm{s.v.}} - 2P_b$. From the basic pressure data it is seen that at frequencies above the b.f. in the region of the phase plateau the s.t. pressure (P_b) was nearly in phase with the s.v. pressure and about 6 dB smaller. (The calibration uncertainty of ± 6 dB makes this observation true within the uncertainty in all cases.) Therefore, the data are generally consistent with a ΔP_{OC} above the peak that is zero or close to it, which is true in most cochlear models. This is particularly pertinent to the investigation of spring–mass resonance.

2. Specific investigations

The six experiments were explored individually for (i) nonlinearity in the pressure data, (ii) evidence for negative resistance *below the b.f.*, and (iii) evidence for spring–mass resonance *closely above the b.f.* Table I, which appears at the end of the text, summarizes these results.

a. Spring–mass resonance. In the table, the ''detection'' of spring–mass resonance refers to an indication of spring–mass resonance in the phase. Resonance can be seen by inspecting the phase of $v_{\mathrm{b.m.}}$ relative to ΔP_{OC}. A 90° lead indicates stiffness, a 90° lag indicates mass, (something less than 90° indicates resistance is present as well) and the transition from leading to lagging will occur at the resonance frequency. Alternatively, the imaginary part of Z_{OC} can be inspected—it will make a transition from negative to positive at the resonance frequency. Consider experiment 12-10-98 (Fig. 12). This cochlea was just barely nonlinear due to inadvertent damage to the cochlea. Therefore, only 85 dB SPL data are shown. The ΔP_{OC} phase began at 2 kHz at $\sim -10°$ with respect to the s.v. pressure, began to accumulate rapidly at 8 kHz, and leveled off at 20 kHz and $\sim -730°$. The $v_{\mathrm{b.m.}}$ phase went through a similar accumulation, but began by leading ΔP_{OC} by about 40°, crossed the ΔP_{OC} phase at 20 kHz, and leveled off at 21 kHz, lagging ΔP_{OC} by ~70°. The relative phase behavior suggests that Z_{OC} was stiffness and resistance dominated below 20 kHz, mass and resistance dominated above 20 kHz, and had a spring–mass resonance at 20 kHz. A spring–mass resonance is used in many cochlear models to bring the cochlear traveling wave to a halt, and the signature of a resonance just where the phase plateaus is compelling. Notably however, the magnitudes did not indicate a spring–mass resonance: the velocity did not peak relative to pressure at 20 kHz. The imaginary part of

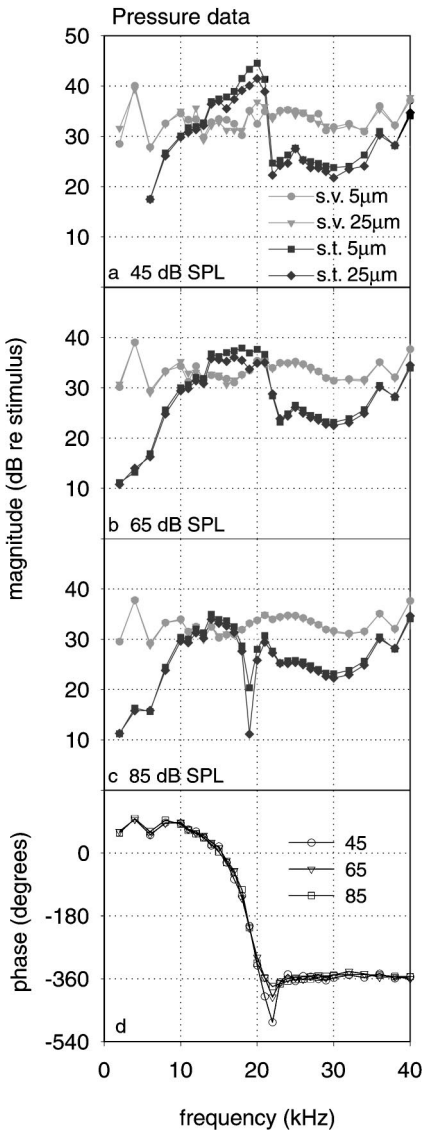


FIG. 14. Pressure data, 9-8-98-I. The s.t. and s.v. pressures for impedance calculation 9-8-98-I usual, 9-8-98-I-double and 9-8-98-I half. Magnitudes are shown relative to the stimulus level in the ear canal; phases (s.t. only) are relative to the simultaneously measured s.v. pressure. Distances in the key refer to the distance between the b.m. and the s.t. sensor. (a) Magnitude, 45 dB SPL stimulus. (b) Magnitude, 65 dB. (c) Magnitude, 85 dB. (d) Phase, all levels.

Z_{OC} from 12-10-98 of course tells the same story: the sign changed from negative to positive at ~20 kHz.

Experiment 3-22-99 (Fig. 13) was very similar to 12-10-98 and is shown primarily to demonstrate repeatability. The phases of $v_{\mathrm{b.m.}}$ and ΔP_{OC} showed a similar course with frequency in the two experiments. In experiment 3-22-99 the resonance frequency indicated by the phase crossing was 25 kHz, which like 12-10-98 was slightly above the b.f. and at the beginning of the phase plateau. Similar to 12-10-98, 3-22-99 showed no sign of the resonance in the relative magnitudes of $v_{\mathrm{b.m.}}$ and ΔP_{OC}. A difference in 3-22-99 compared to 12-10-98 is that in the final plateau $v_{\mathrm{b.m.}}$ lagged ΔP_{OC} by slightly more than 90° (which indicates a component of negative resistance—unlikely in this linear cochlea, and suspected as an experimental error). Nevertheless, the

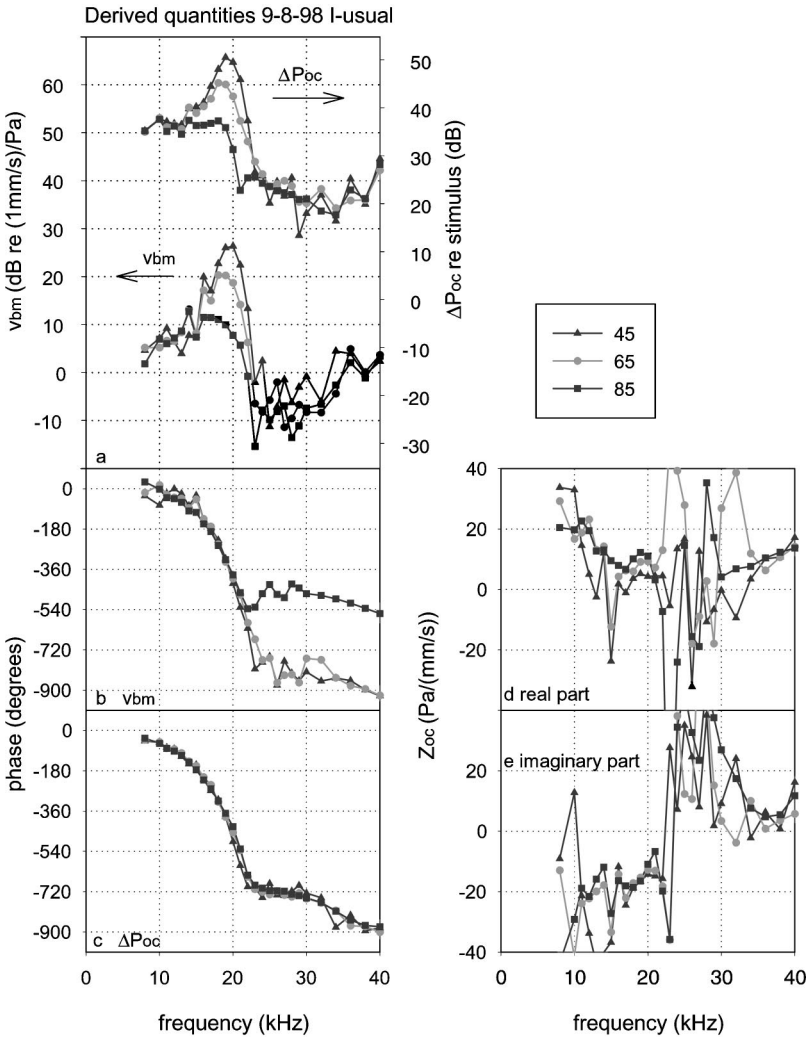


FIG. 15. Derived quantities ΔP_{OC}, $v_{b.m.}$, and Z_{OC}, 9-8-98-I-usual. The usual calculation for ΔP_{OC} was used. (a) Magnitude of $v_{b.m.}$ and ΔP_{OC} (*re* stimulus level). (b) Phase of $v_{b.m.}$ *re* simultaneously measured s.v. pressure. (c) Phase of ΔP_{OC} *re* simultaneously measured s.v. pressure. (d) Real part of Z_{OC} (e) Imaginary part of Z_{OC}.

phase data are close enough to looking like a spring–mass resonance to qualify for a yes in the table.

In the analysis of 9-8-98-I, ΔP_{OC} was calculated in the usual way and the two alternative ways, in which the contribution from s.v. pressure was halved or doubled. While spring–mass resonance was apparent in the phase of 9-8-98-I-usual (Fig. 15) and 9-8-98-I-double (Fig. 17), it was not apparent in the phase of 9-8-98-I-half (Fig. 16). The reason for the difference is that in 9-8-98-I-half at frequencies above the b.f. the s.t. pressure dominated the s.v. pressure in the calculation for ΔP_{OC} (Figs. 16 and 14, which shows the pressure data). Because of this the plateau level of the ΔP_{OC} phase changed by about 180° compared to the ΔP_{OC} phase for 9-8-98-I-usual and 9-8-98-I-double, causing the high frequency $v_{b.m.}$ to lead ΔP_{OC} by ~90° rather than lagging by ~90°. This comparison makes the point that, particularly in the plateau region, the calculated ΔP_{OC} can undergo large changes due to variations in s.v. and s.t. pressure that are within the experimental uncertainty.

On the spring–mass resonance question, the nonlinear experiment 5-6-99 (Fig. 19) is not helpful because it has an erratic $v_{b.m.}$ phase at frequencies above the b.f. Spring–mass resonance was not apparent in the analyses of experiments 3-29-99 and 2-26-97. $v_{b.m.}$ and ΔP_{OC} were nearly in phase at frequencies in the plateau region above the b.f. for 3-29-99-usual (Fig. 21), and $v_{b.m.}$ led ΔP_{OC} slightly in this region for 3-29-99-double (Fig. 22). In 2-26-97-usual (Fig. 24) and 2-26-97-half (Fig. 25) in the plateau region above the b.f. $v_{b.m.}$ led ΔP_{OC} by ~180°.

b. Negative resistance. Negative resistance was indicated when $v_{b.m.}$ led or lagged ΔP_{OC} by more than 90°. If present, it is expected at stimulus frequencies somewhat below the b.f., as these responses were passing through on the way to *their* b.f. place. For the purposes of ''detection,'' negative resistance must be in this region to be counted. (Strictly speaking negative resistance was often in evidence within the plateau region above the peak. However, the result was too erratic there to be trustworthy.) Negative resistance was cautiously detected in the two most nonlinear experiments, 5-6-99 and 9-8-98. To explore its character, consider experiment 9-8-98. 9-8-98-I (Figs. 15, 16, and 17) had a wiggle in the $v_{b.m.}$ phase and magnitude about half an octave below the b.f. The phase wiggle caused $v_{b.m.}$ to lead ΔP_{OC} by slightly more than 90° at 15 kHz. Therefore, negative resistance was indicated. Negative resistance appeared at 45 and 65, but not at 85 dB SPL for analyses 9-8-98-I-usual (Fig. 15) and 9-8-98-I-half (Fig. 16). In these cases both the frequency at which negative resistance appeared (slightly below the b.f.) and its level dependence (present at low levels,

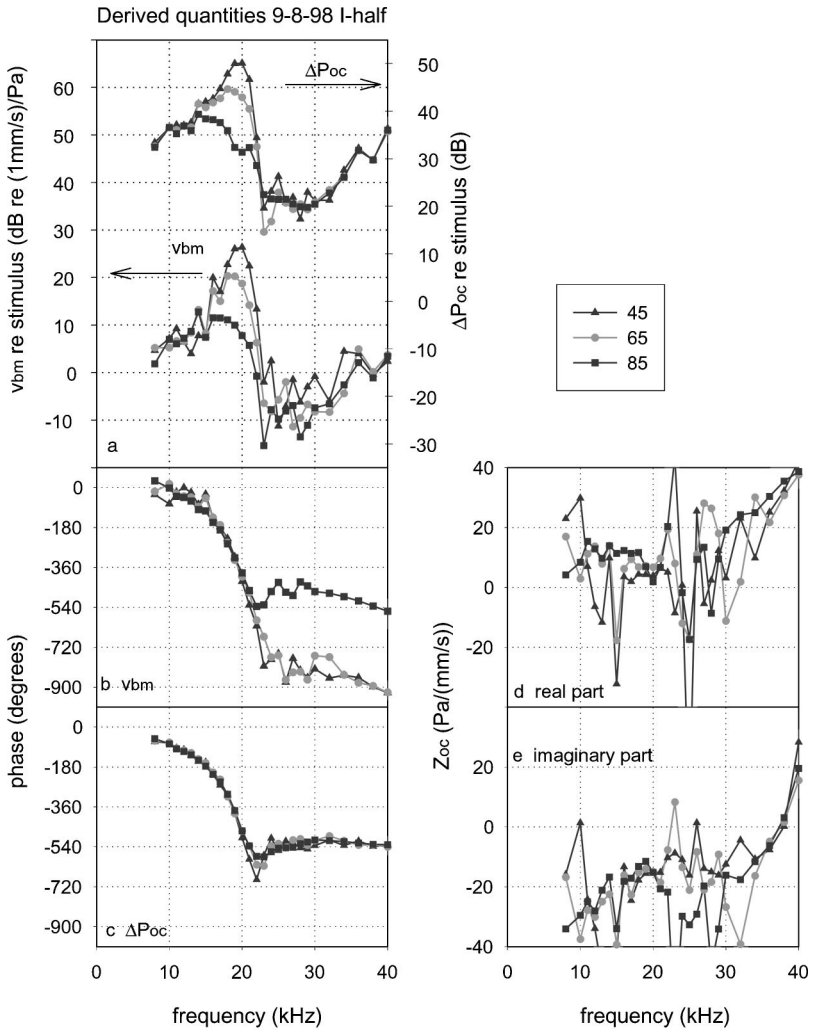


FIG. 16. Derived quantities ΔP_{OC}, $v_{b.m.}$, and Z_{OC}, 9-8-98-I-half. The contribution of $P_{s.v.}$ was half what it is in the usual calculation for ΔP_{OC}. (a) Magnitude of $v_{b.m.}$ and ΔP_{OC} (*re* stimulus level). (b) Phase of $v_{b.m.}$ *re* simultaneously measured s.v. pressure. (c) Phase of ΔP_{OC} *re* simultaneously measured s.v. pressure. (d) Real part of Z_{OC}. (e) Imaginary part of Z_{OC}.

absent at high levels) are consistent with model predictions. Negative resistance appeared at all three levels for 9-8-98-I-double (Fig. 17). Recall that in the ''double'' calculation, the contribution of s.v. pressure to the calculation of ΔP_{OC} was doubled. In the case of 9-8-98-I-double this variation has substantial effects. For one, the ΔP_{OC} phase found with the double calculation was less smooth than that found with the other calculations. This change is what caused the negative resistance to appear at all three levels. A second effect was that ΔP_{OC} was substantially less nonlinear than $v_{b.m.}$ so their ratio, Z_{OC}, *was* nonlinear. Negative resistance did not appear in 9-8-98-II (Fig. 18) a separate but equally nonlinear approach of this experiment. As can be seen in the pressure data from this approach, the s.t. pressure was relatively large in the region of the peak. Because of this, the s.t. pressure dominated s.v. in the calculation for ΔP_{OC}, and thus ΔP_{OC} was very similar in shape to $v_{b.m.}$. Therefore Z_{OC} was quite featureless, with little frequency change and no sign of negative resistance. 5-6-99 (Fig. 19) was similar to 9-8-98-I in that negative resistance was detected at low levels (40 and 50 dB, but not 65 and 85 dB) at a frequency about half an octave below the peak. Also similar to 9-8-98-I, the negative resistance stemmed from a wiggle in the velocity phase. None of the nearly linear experiments, 12-10-98, 3-22-99, and 2-26-97 showed negative resistance. In the analysis of the mildly nonlinear experiment 3-29-99, ΔP_{OC} was calculated in the usual way and with the s.v. contribution doubled. [The relative sizes of the s.v. and s.t. pressures in this experiment (Fig. 20) suggest that a calibration might have been in error and the 3-29-99-double analysis might be more accurate than the usual analysis.] ΔP_{OC} for 3-29-99-usual (Fig. 21) showed a sharp notch at 15 kHz and 85 dB SPL, which stems from a notch in the s.t. pressure. In 3-29-99-usual negative resistance was apparent only in the region of the notch. The level where negative resistance was apparent and the observation that it is related to a notch in ΔP_{OC} make it suspicious as an analytical error. In 3-29-99-double (Fig. 22), negative resistance was not apparent.

c. Extreme base compared to turn one. In gerbil the extreme basal region of the OCC is sandwiched between the cochlear windows, where it would be directly exposed to evanescent pressure modes (Steele and Taber, 1979). Because of this anatomy it is reasonable to expect that the results, particularly of ΔP_{OC}, would differ in the extreme base compared to turn one. It is notable that in healthy (chinchilla) cochleae the character of b.m. motion is quite similar in the extreme base and turn one (Narayan and Ruggero, 2000; Rhode and Recio, 2000).

2-26-97 was an extreme basal experiment, and was dis-

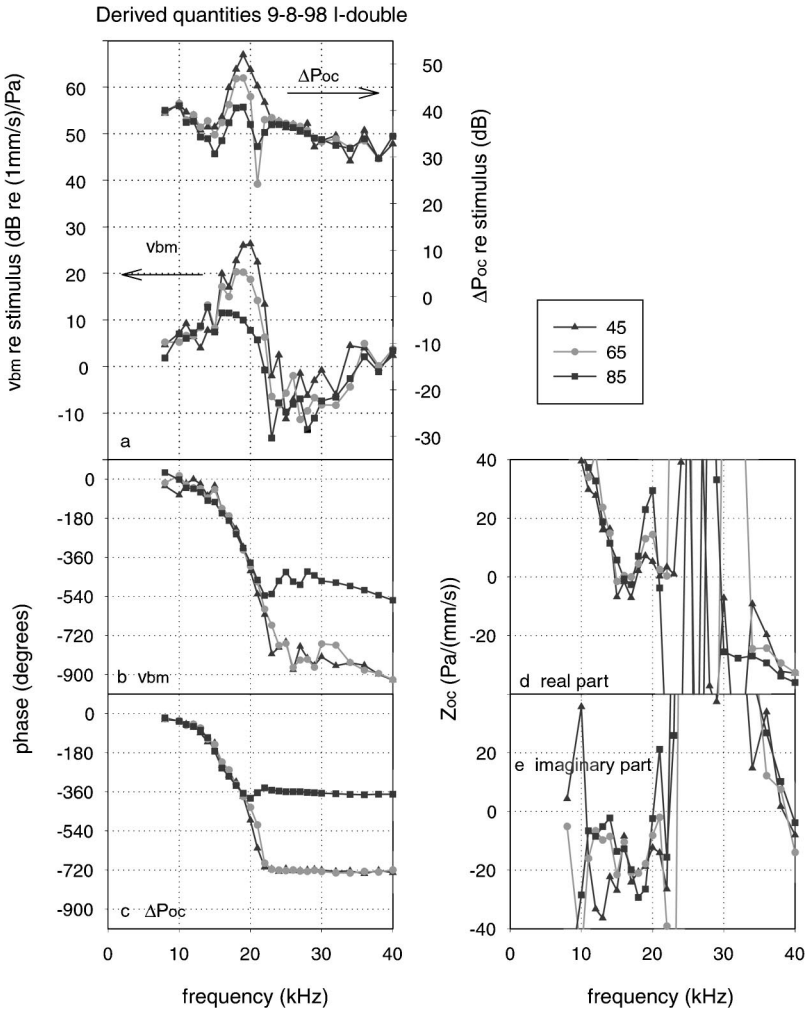


FIG. 17. Derived quantities ΔP_{OC}, $v_{b.m.}$, and Z_{OC}, 9-8-98-I-double. The contribution of $P_{s.v.}$ was twice what it is in the usual calculation for ΔP_{OC}. (a) Magnitude of $v_{b.m.}$ and ΔP_{OC} (*re* stimulus level). (b) Phase of $v_{b.m.}$ *re* simultaneously measured s.v. pressure. (c) Phase of ΔP_{OC} *re* simultaneously measured s.v. pressure. (d) Real part of Z_{OC}. (e) Imaginary part of Z_{OC}.

cussed in Olson (1998). In addition to finding ΔP_{OC} in the usual way (2-26-97-usual) it was also found with the contribution from s.v. pressure halved (2-26-97-half). This variation was calculated because the s.v. pressure was bigger than the average in this experiment (Fig. 23), suggesting that the s.v. sensor calibration might be in error. In that case the 2-26-97-half analysis might be more accurate. The 2-26-97-usual and 2-26-97-half analyses showed substantial dissimilarities. $v_{b.m.}$, ΔP_{OC}, and Z_{OC} from 2-26-97-half (Fig. 25) were quite similar, just shifted up in frequency, to those quantities in turn-one of the linear cochleae 12-10-98 and 3-22-98 (Figs. 12 and 13). In contrast, ΔP_{OC} for 2-26-97-usual (Fig. 24) was just barely tuned. Because $v_{b.m.}$ *was* tuned, Z_{OC} for 2-26-97-usual was more sharply tuned than for any of the turn one experiments. Considering the dissimilarity in the -half and -usual results, more measurements are necessary to decide whether and how cochlear mechanics differs in the region of the windows compared with other locations.

VII. DISCUSSION: CONCLUSIONS, COMPARISONS AND OTHER STRATEGIES FOR MEASURING Z_{OC}

The primary observations of this report were that the pressure close to the sensory tissue was tuned, and that it possessed a degree of nonlinearity similar to that of b.m. motion. These points were evident both in the primary s.t. pressure data, and the derived ΔP_{OC}. Therefore a primary conclusion is that compared to b.m. motion, the impedance of the OCC is relatively untuned, and is nonlinear in a different, subtler way.

The specific question of whether the cochlea provides power amplification in the form of negative resistance proved difficult to answer decisively. The two best cochleae, 9-8-98 and 5-6-99, both exhibited a brief flare of negative resistance over a 1–2 kHz range (one to two data points) within a kHz of the frequency where nonlinearity started, ~half an octave below the b.f. However, negative resistance was not detected in a separate equally nonlinear run of 9-8-98 or in the more mildly nonlinear cochlea 3-29-99. This ambiguity suggests that amplification's signature in the pressure close to the b.m. might vary radially. To address this matter requires a smaller pressure sensor. In the literature the most direct probe of negative resistance is that of deBoer and Nuttall (e.g., 1999, 2000). They derived the OCC impedance with an inverse method by coupling basilar membrane motion data to a 3-dimensional model of the cochlear fluid and geometry. Nuttall and deBoer made measurements of the frequency response at a single location, then used scaling and the cochlear map to convert the measured frequency response into an inferred spatial response. ''Below the b.f.'' in this study corresponds to ''basal to the peak'' in their analy-

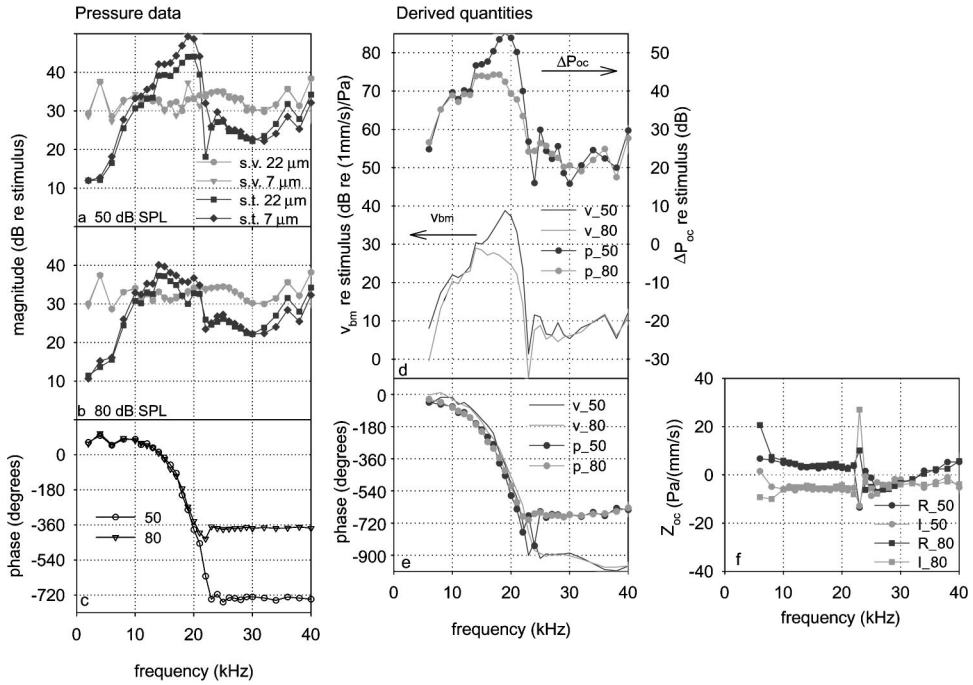


FIG. 18. Pressure data and derived quantities, 9-8-98-II. In approach 9-8-98-II the sensor was angled relative to 9-8-98-I by about 15° so that it would contact the b.m. ~100 μm from the first approach in a direction towards the lamina (a)–(c). The s.t. and s.v. pressures for impedance calculation. Magnitudes are shown relative to the stimulus level in the ear canal; phases (s.t. only) are relative to the simultaneously measured s.v. pressure. Distances in the key refer to the distance between the b.m. and the s.t. sensor. (a) Magnitude, 50 dB SPL stimulus. (b) Magnitude, 80 dB. (c) Phase, both levels. (d)–(f) ΔP_{OC}, $v_{b.m.}$, and Z_{OC}. The usual calculation for ΔP_{OC} was used. (d) Magnitude of $v_{b.m.}$ and ΔP_{OC} (*re* stimulus level). (e) Phase of $v_{b.m.}$ and ΔP_{OC} (*re* simultaneously measured s.v. pressure). (f) Real and imaginary parts of Z_{OC}.

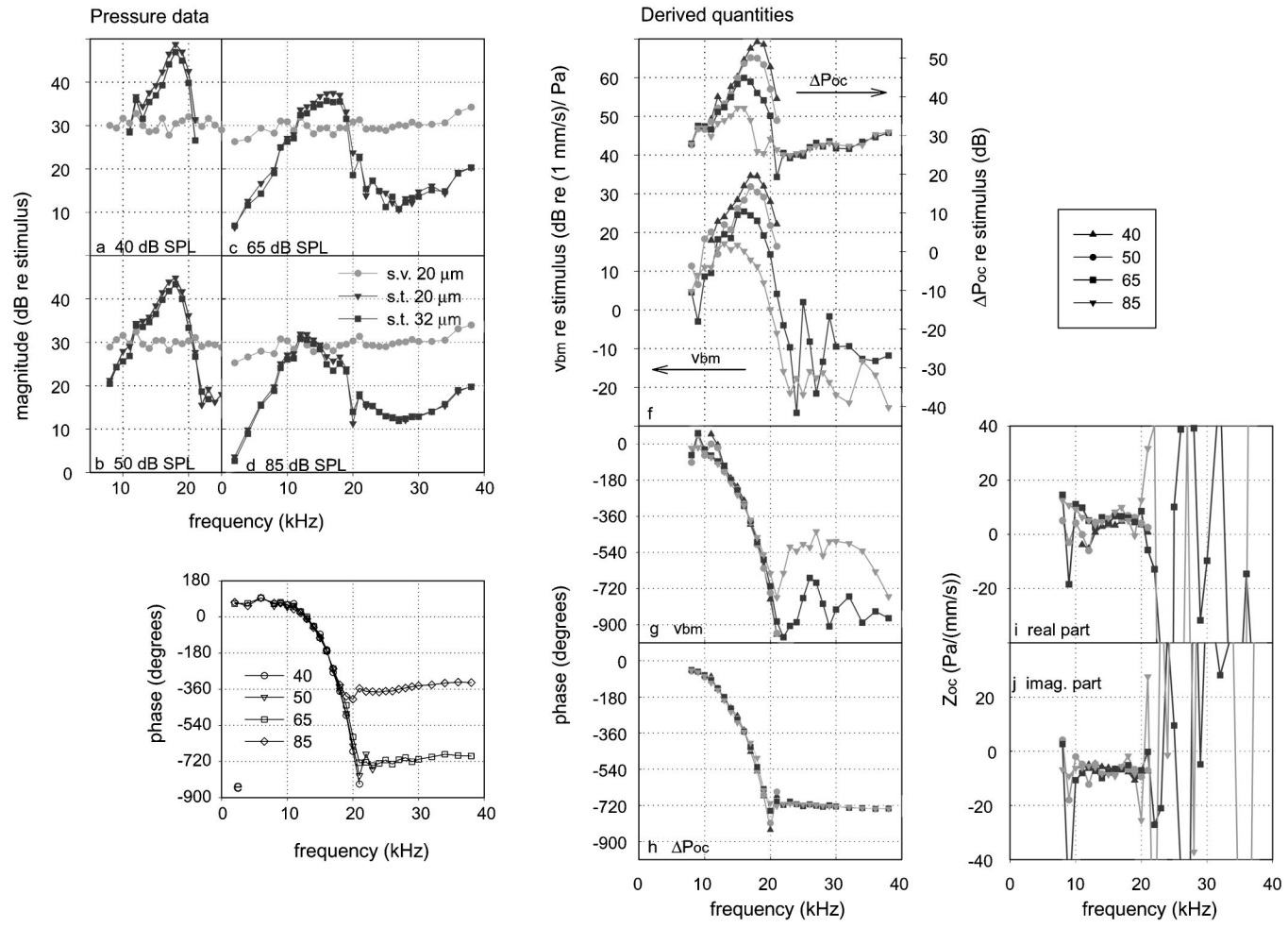


FIG. 19. Pressure data and derived quantities, 5-6-99. The analysis of this experiment used s.t. pressure data at the position not closest, but second closest to the b.m. because of the reversal in pressure magnitude that was described in Fig. 11. (a)–(e) The s.t. and s.v. pressures for impedance calculation. Magnitudes are shown relative to the stimulus level in the ear canal; phases (s.t. only) are relative to the simultaneously measured s.v. pressure. Distances in the key refer to the distance between the b.m. and the s.t. sensor. (a) Magnitude, 40 dB SPL stimulus. (b) Magnitude, 50 dB. (c) Magnitude, 65 dB. (d) Magnitude, 85 dB. (e) Phase, all levels. (f)–(j) ΔP_{OC}, $v_{b.m.}$, and Z_{OC}. The usual calculation for ΔP_{OC} was used. (f) Magnitude of $v_{b.m.}$ and ΔP_{OC} (*re* stimulus level). (g) Phase of $v_{b.m.}$ *re* simultaneously measured s.v. pressure. (h) Phase of ΔP_{OC} re simultaneously measured s.v. pressure. (i) Real part of Z_{OC}. (j) Imaginary part of Z_{OC}.

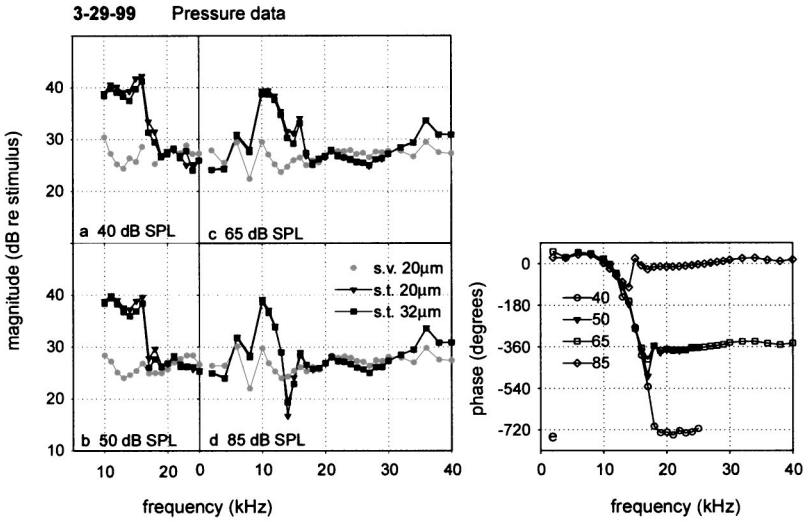


FIG. 20. Pressure data, 3-29-99. The s.t. and s.v. pressures for impedance calculation 3-29-99-usual and 3-29-99-double. The impedance analysis of this experiment used s.t. pressure data at the position not closest, but second closest to the b.m. because of a reversal in pressure magnitude at the closest position. Magnitudes are shown relative to the stimulus level in the ear canal; phases (s.t. only) are relative to the simultaneously measured s.v. pressure. Distances in the key refer to the distance between the b.m. and the s.t. sensor. (a) Magnitude, 40 dB SPL stimulus. (b) Magnitude, 50 dB. (c) Magnitude, 65 dB. (d) Magnitude, 85 dB. (e) Phase, all levels.

sis. The detections of negative resistance above agreed with deBoer and Nuttall's results in that a relatively small degree of nonlinear negative resistance at frequencies below the b.f. produced a large degree of nonlinearity in the response in the b.f. region. However, in the report of deBoer and Nuttall the region of negative resistance was broader, and more robust than in the current report. Notably, at the frequencies where they found negative resistance its magnitude was at most 20%–30% of the magnitude of the imaginary part of the impedance. These robust detections occurred at low stimulus levels in very sensitive cochleae. If this is true, negative resistance will be challenging to detect decisively with the more direct approach of the present study.

The results of de Boer and Nuttall (1999) and those of the present report agree on other points as well. In both reports well beneath the b.f. (basal to the peak) the real part and imaginary part of the impedance were usually within ~ a factor of 2 to each other in magnitude. In both reports even

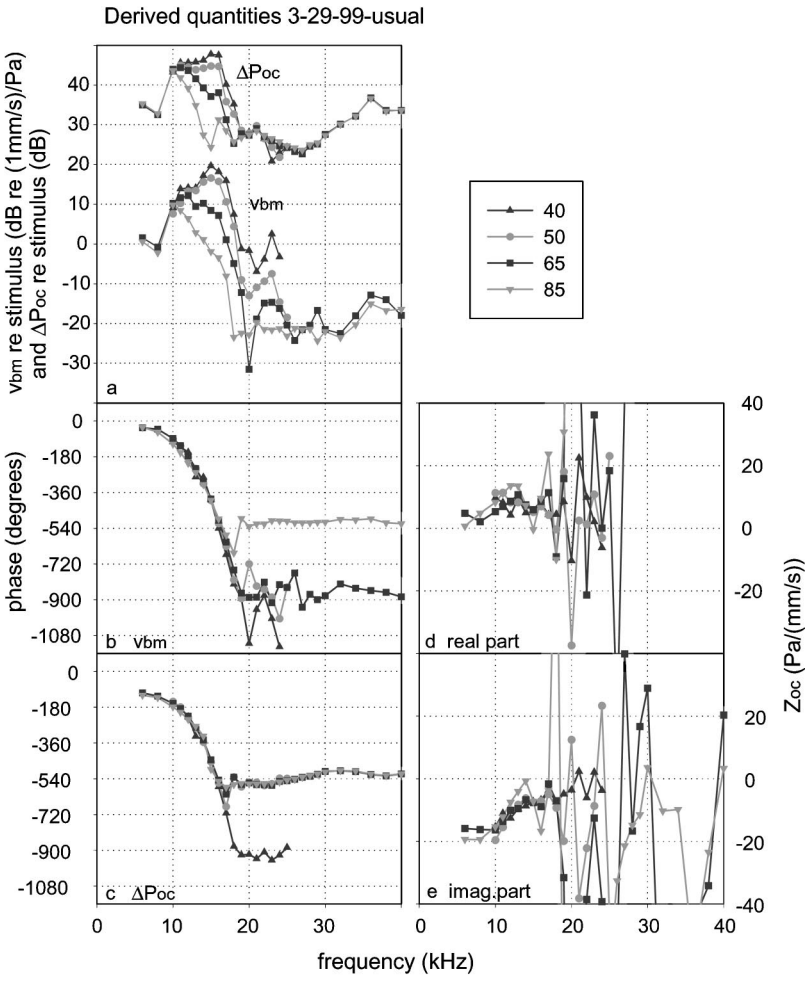


FIG. 21. Derived quantities, ΔP_{OC}, $v_{b.m.}$, and Z_{OC}, 3-29-99-usual. The usual calculation for ΔP_{OC} was used. (a) Magnitude of $v_{b.m.}$ and ΔP_{OC} (*re* stimulus level). (b) Phase of $v_{b.m.}$ *re* simultaneously measured s.v. pressure. (c) Phase of ΔP_{OC} *re* simultaneously measured s.v. pressure. (d) Real part of Z_{OC}. (e) Imaginary part of Z_{OC}.

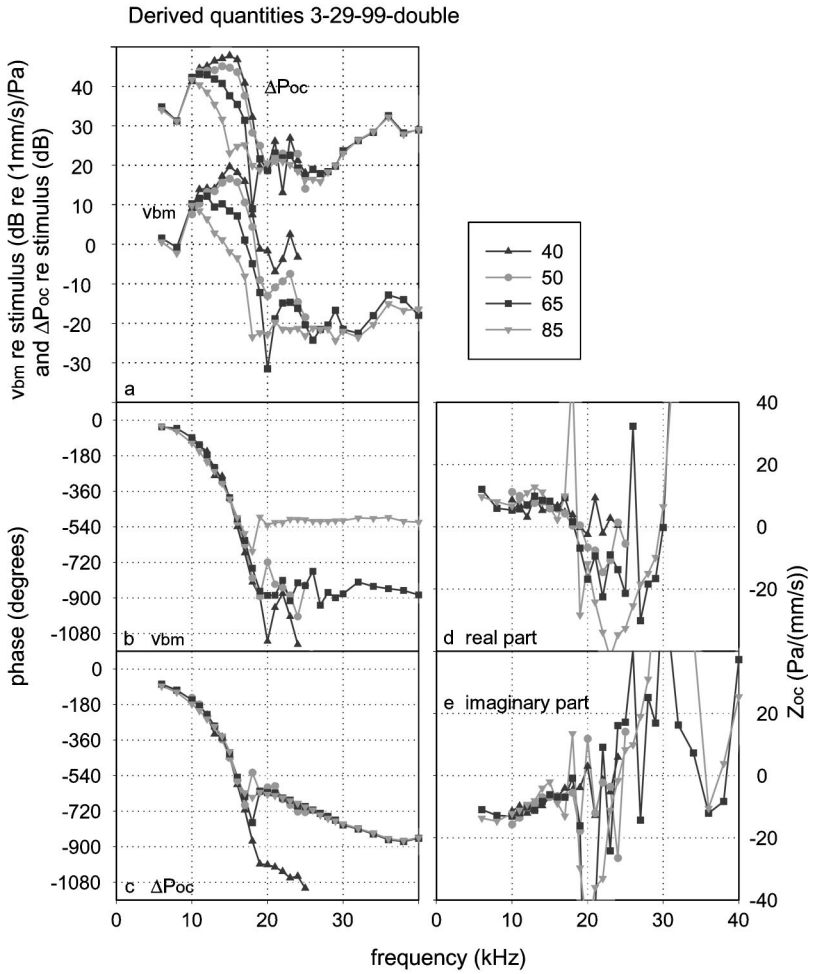


FIG. 22. Derived quantities, ΔP_{OC}, $v_{b.m.}$, and Z_{OC}, 3-29-99-double. The contribution of $P_{s.v.}$ was twice what it is in the usual calculation for ΔP_{OC}. (a) Magnitude of $v_{b.m.}$ and ΔP_{OC} (*re* stimulus level). (b) Phase of $v_{b.m.}$ *re* simultaneously measured s.v. pressure. (c) Phase of ΔP_{OC} *re* simultaneously measured s.v. pressure. (d) Real part of Z_{OC}. (e) Imaginary part of Z_{OC}.

in passive cochleae the real part of the impedance varied substantially with frequency (place).

The question of whether Z_{OC} possesses a spring–mass resonance is fundamental. If it does, the frequency map of the cochlea is established primarily by the stiffness and mass of the OCC and at frequencies above the resonant frequency (a little above the b.f.) of a particular point the traveling wave is not present at all. Many cochlear models employ a resonant Z_{OC} (e.g., Neely and Kim, 1986; Peterson and Bogart, 1950; Kolston, 2000). If the spring–mass resonance does not exist the frequency map of the cochlea is established by the stiffness of the OCC and wavelength dependent changes in fluid mass. Z_{OC} never becomes masslike, and the traveling wave is small but present at frequencies well above the b.f. Many cochlear models do *not* contain a resonant Z_{OC}. Steele and colleagues in particular maintain that the OC mass should have very little effect on cochlear mechanics (e.g., Steele, 1999; Steele and Taber, 1981). (The argument against the OC mass playing a mechanical role is that most of the cells of the OC are soft and the fluid within the cells would move almost as it would if it was not enclosed within cells.)

Experimentally, there is nothing truly compelling either for or against resonance. In the present studies and broadly in the literature a phase plateau is present at frequencies above b.f. (e.g., Rhode, 1971). At first glance this seems like evidence that the traveling wave has stopped. However, the traveling wave mode need only be small compared with the compressive mode (or other nonpropagating modes) and a phase plateau will be observed. In the presented study spring–mass resonance was detected in the impedance phase in half the measurements, appearing in both linear and nonlinear cochleae. These detections were based on the presence of a stiffness-mass transition, which occurred where the phase plateau began. The strength of this detection was compromised by the fact that in the plateau region the analysis for ΔP_{OC} was susceptible to experimental inaccuracies. Moreover, to be convincing the resonance should be apparent in both the magnitude and phase and it was not apparent in the magnitude. The inverse method of deBoer and Nuttall (e.g., 1999) did not detect a resonance. Although one could argue that their frequency range was not extended high enough above the b.f. to address resonance their results appear to weigh in against it. There is experimental evidence from linear cochleae that speaks against resonance. In linear cochleae with drained scala tympani the b.f. shifted up by about half an octave (discussed in Steele and Taber, 1981; Patuzzi *et al.*, 1982). This is predicted if fluid mass, not organ of Corti mass, determined the location of the peak. A complementary observation is that of Cooper and Rhode (1995), who in measurements of b.m. motion in the apex of guinea pig cochleae found no difference in peak location or shape when the organ of Corti was removed.

This report concludes with further strategies for probing negative resistance and resonance in Z_{OC}. The question of

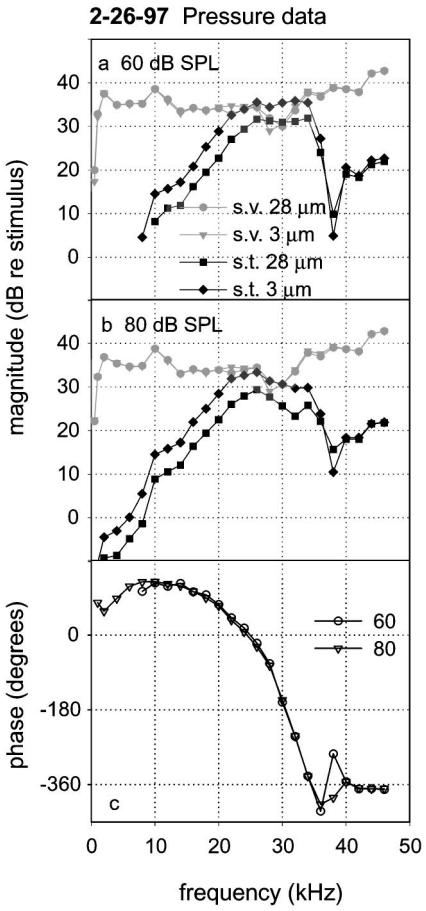


FIG. 23. Pressure data, 2-26-97. The s.t. and s.v. pressures for impedance calculation 2-26-97-usual and 2-26-97-half. This was an extreme basal experiment. Magnitudes are shown relative to the stimulus level in the ear canal; phases (s.t. only) are relative to the simultaneously measured s.v. pressure. Distances in the key refer to the distance between the b.m. and the s.t. sensor. (a) Magnitude, 60 dB SPL stimulus. (b) Magnitude, 80 dB. (c) Phase, both levels.

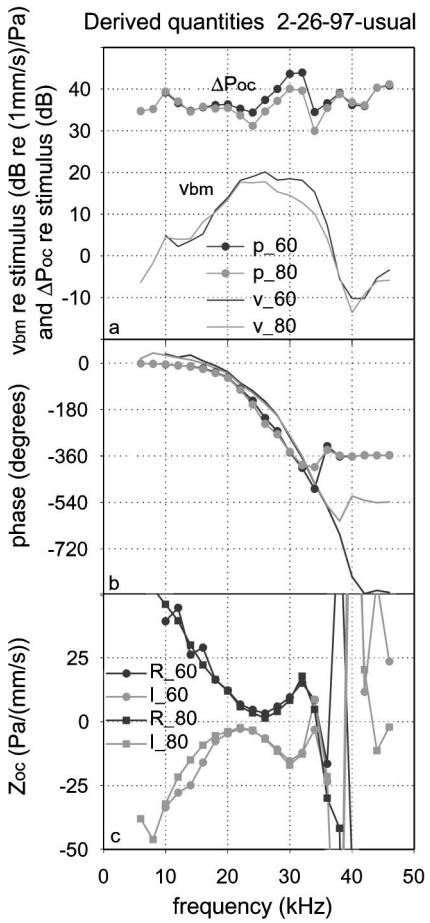


FIG. 24. Derived quantities, ΔP_{OC}, $v_{b.m.}$, and Z_{OC}, 2-26-97-usual. The usual calculation for ΔP_{OC} was used. (a) Magnitude of $v_{b.m.}$ and ΔP_{OC} (*re* stimulus level). (b) Phase of $v_{b.m.}$ and ΔP_{OC} (*re* simultaneously measured s.v. pressure). (c) Real and imaginary parts of Z_{OC}.

negative resistance might be explored via a more detailed map of the s.t. pressure. The impedance of the OCC influences the way that the fluid velocity varies with distance from the b.m. For example, in a simple two-dimensional (2D) system an impedance of stiffness is linked to an exponentially decreasing velocity-with-distance. However, a partly resistive impedance introduces oscillations into the exponential decrease (de Boer, 1984). The fluid velocity over a range of distances from the b.m. is measurable via pressure gradients (Olson, 1999). Such measurements, linked to a three-dimensional (3D) cochlear model might illuminate the question of negative resistance.

The resonance question might be explored via measurements of the longitudinal curvature of the traveling wave. Experimentally, curvature is found by taking the difference in b.m. motion phases ($\delta\phi$) between two locations spaced by a small longitudinal distance (δx). The curvature equals $\delta\phi/\delta x$ and is represented by the ''wave number,'' κ. The frequency dependence of the curvature depends on the organ of Corti mass (m_{OCC}, the OC mass/unit length) and fluid mass [$m_{eq}(\kappa)$, the equivalent mass of fluid/unit length which resists the displacement of the b.m. (Steele and Taber, 1979)]. These masses relate directly to resonance—resonance exists if and when m_{OCC} dominates $m_{eq}(\kappa)$ (Lighthill, 1981). $m_{eq}(\kappa)$ is prominent in many modeling papers [$m_{eq}(\kappa)$ is a scaled version of ''$h_{eq}(\kappa)$'' or ''$Q(\kappa)$'' found in, e.g., Steele and Taber, 1979; Lighthill, 1981; de Boer, 1984]. Its variation with κ is what distinguishes 1D, 2D, and 3D cochlear models. (In the following, ω is the angular frequency, z is the z displacement of the b.m. and s is the OCC stiffness/length at the longitudinal location of the measurement. As in most cochlear models, s is assumed independent of κ.)

By equating the potential and kinetic energies of the traveling wave at every longitudinal location the relationship between stiffness, mass and frequency is (Lighthill, 1983):

$$\tfrac{1}{2}sz^2=\tfrac{1}{2}(m_{eq}(\kappa)+m_{OCC})\omega^2z^2. \tag{5}$$

Experimentally measuring curvature at many frequencies leads to $\kappa(\omega)$, or equivalently, $\omega(\kappa)$. Equation (5) can be rewritten as

$$(m_{eq}(\kappa)+m_{OCC})/s=(\omega(\kappa))^{-2}. \tag{6}$$

The right-hand side of Eq. (6) is an experimental quantity, so the left-hand side is experimentally accessible. The idea is to find how the wave curvature varies with frequency in the region of the peak, and to use the result to ''measure'' $(m_{eq}(\kappa)+m_{OCC})/s$ vs κ. Does it begin to level off to a con-

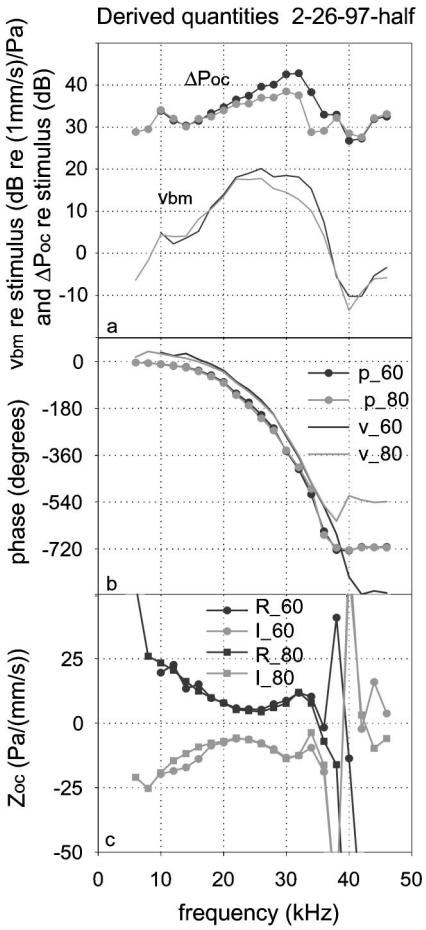


FIG. 25. Derived quantities, ΔP_{OC}, $v_{b.m.}$, and Z_{OC}, 2-26-97-half. The contribution of $P_{s.v.}$ was half what it is in the usual calculation for ΔP_{OC}. (a) Magnitude of $v_{b.m.}$ and ΔP_{OC} (*re* stimulus level). (b) Phase of $v_{b.m.}$ and ΔP_{OC} (*re* simultaneously measured s.v. pressure). (c) Real and imaginary parts of Z_{OC}.

stant value? (If yes, resonance is supported.) Does it look just like $m_{eq}(\kappa)$ from 3D cochlear models looks, so m_{OCC} is effectively zero? (If yes, resonance is contested.) Lighthill (1981) examined Rhode's (1971) measurements, longitudinally spaced by 1.5 mm, and decided that the variation of ω with κ was consistent with the presence of resonance. However, more closely spaced longitudinal measurements, such as have been appearing in the experimental literature (Russell and Nilson, 1997; Ren, 2001; Rhode and Recio, 2000) are better for measuring curvature and for addressing the question of resonance.

ACKNOWLEDGMENTS

This work was supported by NIH DC03130. Many thanks to Egbert deBoer, Hideko Nakajima and the J. Acoust. Soc. Am. reviewers, Chris Shera and ''A'' for improving the presentation.

[1]This derives from a simplification of the Navier–Stokes equation, $\nabla P = -\rho \partial v/\partial t - \rho v \nabla v + \mu \nabla^2 v$. The equation can be approximated as $\nabla P = -\rho \partial v/\partial t$ at high enough frequencies. In Olson (1998) the relative sizes of the three right-hand terms was approximated using dimensional analysis. The length scale over which fluid velocity changed by $\sim$ a factor of e was estimated as 100 μm. Then at 1.5 kHz the first right-side term was 1000 times bigger than the second term and 100 times bigger than the third term. The dominance of the first term grows with frequency. However, in recent measurements close to the b.m. the length over which the velocity dropped off by a factor of e was found to be only 15 μm (Olson, 1999). This reduced length scale reduced the dominance of the first term; at 3.5 kHz, it is 300 times bigger than the second term and 5 times bigger than the third term. Therefore, the method to find fluid velocity close to the b.m. using Eq. (2) is restricted to frequencies over several kHz.

[2]With a symmetric cochlear model, the pressure can be decomposed into symmetric and antisymmetric parts (Peterson and Bogart, 1950). For simplicity the derivation above only discussed the antisymmetric part. Including the symmetric component does not change the answer as long as the symmetric component does not vary spatially. The symmetric component is associated with the compressional wave. It is expected to vary in space

TABLE I. Summary of impedance results from six experiments. Symbols in the table refer to the following notes: *Results ambiguous, see discussion. †Roman numerals I and II in expt 9-8-98 refer to two different approaches. The suffixes half, double, and usual are appended when alternative calculations for ΔP_{OC} were done. §During the first approach of expt. 9-8-98, two pressure series were actually taken, separated by more than an hour in time, but without repositioning the sensor. The results of these two series tested repeatability. They were similar and in particular both showed negative resistance. Only one of these is presented here (9-8-98-I). In a separate approach (9-8-98-II) the sensor was angled relative to the first by about 15° so that it would contact the b.m. $\sim$100 μm from the first approach in a direction towards the lamina. Because this approach found larger s.t. pressures the sensor was probably more radially and/or perpendicularly centered on the b.m. then.

Experiment date and region	Nonlinear experiment?	Measurement and analysis name†	Figure numbers	Negative resistance?	Spring–mass resonance? (apparent in phase)
12-10-98 turn one	no (or just barely)	12-10-98	12	no	yes
3-22-99 turn one	no (or just barely)	3-22-99	13	no	yes
9-8-98 turn one§	yes	9-8-98-I-usual	14,15	yes	yes
		9-8-98-I-half	14,16	yes	no
		9-8-98-I-double	14,17	*	yes
		9-8-98-II	18	no	no
5-6-99 turn one	yes	5-6-99	19	yes	*
3-29-99 turn one	yes, somewhat	3-29-99-usual	20,21	*	no
		3-29-99-double	20,22	no	no
2-26-97 extreme base	no (or just barely)	2-26-97-usual	23,24	no	no
		2-6-97-half	23,25	no	no

much more slowly than the antisymmetric part (Lighthill, 1981) and in the section on pressure versus position this was confirmed. Therefore, it can be considered as an offset, P_c. The antisymmetric part satisfies $P'_{\text{s.v.}} - P'_{\text{s.v.–OCC}} = P'_b - P'_{\text{r.w.}}$, but these primed quantities are no longer the actual pressure at each of these positions. Adding P_c to each term returns it to a form that includes actual pressures: $(P'_{\text{s.v.}}+P_c)-(P'_{\text{s.v.–OCC}}+P_c) = (P'_b+P_c)-(P'_{\text{r.w.}}+P_c)$ [equation (i)]. $(P'_{\text{s.v.}}+P_c)$ is the pressure measured in the s.v. $(P_{\text{s.v.}})$, (P'_b+P_c) is the s.t. pressure measured close to the b.m. (P_b). Because of the r.w. boundary condition $(P'_{\text{r.w.}}+P_c)=0$ [equation (ii)]. The desired quantity is $\Delta P_{\text{OC}}=(P'_{\text{s.v.–OCC}}+P_c)-(P'_b+P_c)$. From Eqs. (i) and (ii), $\Delta P_{\text{OC}}=(P'_{\text{s.v.}}+P_c)-2(P'_b+P_c)=P_{\text{s.v.}}-2P_b$, just as in Eq. (3).

Aibara, R., Welch, J., Puria, S., and Goode, R. L. (**1999**). ''Direct measurement of cochlear input impedance in human cadaver ear,'' Abstracts of the 2nd International Symposium of Middle-ear Mechanics in Research and Otosurgery, Boston, MA.

Avan, P., Magnan, P., Smurzynski, J., Probst, R., and Dancer, A. (**1998**). ''Direct evidence of cubic difference tone propagation by intracochlear acoustic pressure measurements in the guinea-pig,'' Eur. J. Neurosci. **10**, 1764–1770.

Beranek, L. L. (**1954**). *Acoustics* (Acoustical Society of America, 1996), Chap. 3.

Cooper, N. P. (**1998**). ''Harmonic distortion on the basilar membrane in the basal turn of the guinea-pig cochlea,'' J. Physiol. (Paris) **509**, 277–288.

Cooper, N. P. (**2000**). ''Radial variation in the vibrations of the cochlear partition,'' in *Recent Developments in Auditory Mechanics*, edited by H. Wada, T. Takasaka, K. Ikeda, K. Ohyama, and T. Koike (World Scientific, Singapore), pp. 109–115.

Cooper, N. P., and Rhode, W. S. (**1995**). ''Nonlinear mechanics at the apex of the guinea-pig cochlea,'' Hear. Res. **82**, 225–243.

Cooper, N. P., and Rhode, W. S. (**1996**). ''Fast traveling waves, slow traveling waves, and their interactions in experimental studies of apical cochlear mechanics,'' Aud. Neurosci. **2**, 207–212.

Crawford, A. C., and Fettiplace, R. (**1981**). ''An electrical tuning mechanism in turtle cochlear hair cells,'' J. Physiol. (London) **312**, 377–412.

Dancer, A., and Franke, R. (**1980**). ''Intracochlear sound pressure measurements in guinea pigs,'' Hear. Res. **2**, 191–205.

de Boer, E. (**1983**). ''No sharpening? A challenge for cochlear mechanics,'' J. Acoust. Soc. Am. **73**, 567–573.

de Boer, E. (**1984**). ''Auditory physics. Physical principles in hearing theory. II,'' Phys. Rep. **105**, 141–226.

de Boer, E., and Nuttall, A. L. (**1999**). ''The inverse problem solved for a three-dimensional model of the cochlea. III. Brushing-up the solution method,'' J. Acoust. Soc. Am. **105**, 3410–3420.

de Boer, E., and Nuttall, A. L. (**2000**). ''The mechanical waveform of the basilar membrane. III. Intensity effects,'' J. Acoust. Soc. Am. **107**, 1497–1507.

Decory, L., Franke, R. B., and Dancer, A. L. (**1990**). ''Measurement of middle ear transfer function in cat, chinchilla and guinea pig,'' in *The Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer-Verlag, Berlin), pp. 270–277.

Freeman, D. M., and Weiss, T. F. (**1990**). ''Hydrodynamic analysis of a two-dimensional model for micromechanical resonance of free-standing hair bundles,'' Hear. Res. **48**, 37–68.

Johnstone, J. R., Alder, V. A., Johnstone, B. M., Robertson, D., and Yates, G. K. (**1979**). ''Cochlear action potential threshold and single unit threshold,'' J. Acoust. Soc. Am. **65**, 254–257.

Kolston, P. J. (**2000**). ''The importance of phase data and model dimensionality to cochlear mechanics,'' Hear. Res. **145**, 25–36.

Kolston, P. J. (**1999**). ''Comparing *in vitro*, *in situ*, and *in vivo* experimental data in a three-dimensional model of mammalian cochlear mechanics,'' Proc. Natl. Acad. Sci. U.S.A. **96**, 3676–3681.

Lighthill, J. (**1981**). ''Energy flow in the cochlea,'' J. Fluid Mech. **106**, 149–213.

Lighthill, J. (**1983**). ''Advantages from describing cochlear mechanics in terms of energy flow,'' in *Mechanics of Hearing*, edited by E. de Boer and M. A. Viergever (Delft University Press), pp. 63–71.

Lynch, T. J., Nedzelnitsky, V., and Peake, W. T. (**1982**). ''Input impedance of the cochlea in cat,'' J. Acoust. Soc. Am. **72**, 108–123.

Magnan, P., Avan, P., Dancer, A., Probst, R., and Smurzynski, J. (**1997**). ''A new approach to cochlear mechanics and cubic distortion tones by intracochlear acoustic pressure measurements in the guinea pig,'' in *Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. Hecht-Poinar (World Scientific, Singapore), pp. 333–338.

Magnan, P., Dancer, A., Probst, R., Smurzynski, J., and Avan, P. (**1999**). ''Intracochlear acoustic pressure measurements: Transfer functions of the middle ear and cochlear mechanics,'' Aud. Neurootol. **4**, 123–128.

Mountain, D. C., and Hubbard, A. E. (**1989**). ''Rapid force production in the cochlea,'' Hear. Res. **42**, 195–202.

Narayan, S. S., and Ruggero, M. A. (**2000**). ''Basilar membrane mechanics at the hook region of the chinchilla cochlea,'' in *Recent Developments in Auditory Mechanics*, edited by H. Wada, T. Takasaka, K. Ikeda, K. Ohyama, and T. Koike (World Scientific, Singapore), pp. 95–101.

Nedzelnitsky, V. (**1980**). ''Sound pressures in the basal turn of the cat cochlea,'' J. Acoust. Soc. Am. **68**, 1676–1689.

Neely, S. T., and Kim, D. O. (**1986**). ''A model for active elements in cochlear biomechanics,'' J. Acoust. Soc. Am. **79**, 1472–1480.

Nilsen, K. E., and Russell, I. J. (**1999**). ''Timing of cochlear feedback: spatial and temporal representation of a tone across the basilar membrane,'' Nat. Neurosci. **2**, 642–648.

Olson, E. S. (**1998**). ''Observing middle and inner ear mechanics with novel intracochlear pressure sensors,'' J. Acoust. Soc. Am. **103**, 3445–3463.

Olson, E. S. (**1999**). ''Direct measurement of intracochlear pressure waves,'' Nature (London) **402**, 526–529.

Olson, E. S. (**2000**). ''The use of intracochlear pressure to find the impedance of the organ of Corti,'' in *Recent Developments in Auditory Mechanics*, edited by H. Wada, T. Takasaka, K. Ikeda, K. Ohyama, and T. Koike (World Scientific, Singapore), pp. 144–150.

Olson, E. S., and Mountain, D. C. (**1991**). ''In vivo measurement of basilar membrane stiffness,'' J. Acoust. Soc. Am. **89**, 1262–1275.

Olson, E. S., and Cooper, N. P. (**2000**). ''Stapes motion and scala vestibuli pressure in gerbil,'' Abstract #4415 from the Midwinter Meeting of the Association for Research in Ololaryngology.

Patuzzi, R., Sellick, P. M., and Johnstone, B. M. (**1982**). ''Cochlear drainage and basilar membrane tuning,'' J. Acoust. Soc. Am. **72**, 1064–1065.

Peterson, L. C., and Bogart, B. P. (**1950**). ''A dynamical theory of the cochlea,'' J. Acoust. Soc. Am. **22**, 369–381.

Puria, S., and Rosowski, J. J. (**1997**). ''Measurement of reverse transmission in the human middle ear: preliminary results,'' in *Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. Hecht-Poinar (World Scientific, Singapore), pp. 151–157.

Puria, S., Peake, W. T., and Rosowski, J. J. (**1997**). ''Sound-pressure measurements in the cochlear vestibule of human-cadaver ears,'' J. Acoust. Soc. Am. **101**, 2754–2769.

Ren, T. (**2001**). ''Direct measurement of the traveling wave using a scanning laser interferometer in sensitive gerbil cochlea,'' Abstract #555 from the Midwinter Meeting of the Association for Research in Ololaryngology.

Rhode, W. S. (**1971**). ''Observations on the vibration of the basilar membrane in squirrel monkey using the Mössbauer technique,'' J. Acoust. Soc. Am. **49**, 1218–1231.

Rhode, W. S., and Recio, A. (**2000**). ''Study of mechanical motions in the basal region of the chinchilla cochlea,'' J. Acoust. Soc. Am. **107**, 3317–3332.

Ruggero, M. A., Rich, N. C., Robles, L., and Bhagyalakshmi, G. S. (**1990**). ''Middle ear response in the chinchilla and its relationship to mechanics at the base of the cochlea,'' J. Acoust. Soc. Am. **87**, 1612–1629.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (**1997**). ''Basilar membrane responses to tones at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **101**, 2151–2163.

Russell, I. J., and Nilsen, K. E. (**1997**). ''The location of the cochlear amplifier: Spatial representation of a single tone on the guinea pig basilar membrane,'' Proc. Natl. Acad. Sci. U.S.A. **94**, 2660–2664.

Steele, C. R., and Taber, L. A. (**1979**). ''Comparison of WKB calculations and experimental results for three-dimensional cochlear models,'' J. Acoust. Soc. Am. **65**, 1007–1018.

Steele, C. R., and Taber, L. A. (**1981**). ''Three-dimensional model calculations for guinea pig cochlea,'' J. Acoust. Soc. Am. **69**, 1107–1111.

Steele, C. R. (**1999**). ''Toward three-dimensional analysis of cochlear structure,'' ORL **61**, 238–251.

Zwislocki, J. (**1965**). ''Analysis of some auditory characteristics,'' in *Handbook of Mathematical Psychology, Vol. III*, edited by R. D. Luce, R. R. Bush, and E. Galanter (Wiley, New York), pp. 3–97.

Auditory nerve fiber responses to electric stimulation: Modulated and unmodulated pulse trains[a)]

Leonid Litvak[b)]
Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, Massachusetts 02114-3096; Speech and Hearing Sciences Program, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139; and Cochlear Implant Research Laboratory, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, Massachusetts 02114-3096

Bertrand Delgutte
Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, Massachusetts 02114-3096; Speech and Hearing Sciences Program, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139; and Research Laboratory of Electronics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Donald Eddington
Cochlear Implant Research Laboratory, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, Massachusetts, 02114-3096; Speech and Hearing Sciences Program, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139; and Research Laboratory of Electronics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

(Received 12 October 2000; revised 5 March 2001; accepted for publication 3 April 2001)

Many modern cochlear implants use sound processing strategies that stimulate the cochlea with modulated pulse trains. Rubinstein *et al.* [Hear. Res. **127**, 108 (1999)] suggested that representation of the modulator in auditory nerve responses might be improved by the addition of a sustained, high-rate, desynchronizing pulse train (DPT). In addition, activity in response to the DPT may mimic the spontaneous activity (SA) in a healthy ear. The goals of this study were to compare responses of auditory nerve fibers in acutely deafened, anesthetized cats elicited by high-rate electric pulse trains delivered through an intracochlear electrode with SA, and to measure responses of these fibers to amplitude-modulated pulse trains superimposed upon a DPT. Responses to pulse trains showed variability from presentation to presentation, but differed from SA in the shape of the envelope of the interval histogram (IH) for pulse rates above 4.8 kpps (kilo pulses per second). These IHs had a prominent mode near 5 ms that was followed by a long tail. Responses to modulated biphasic pulse trains resembled responses to tones in intact ears for small ($<10\%$) modulation depths, suggesting that acousticlike responses to sinusoidal stimuli might be obtained with a DPT. However, realistic responses were only observed over a narrow range of levels and modulation depths. Improved coding of complex stimulus waveforms may be achieved by signal processing strategies for cochlear implants that properly incorporate a DPT. [DOI: 10.1121/1.1375140]

PACS numbers: 43.64.Me, 43.64.Pg [LHC]

I. INTRODUCTION

In the continuous interleaved sampling (CIS) strategies used in many modern cochlear implant processors, temporal information about incoming sounds is encoded in the amplitude modulations of pulse trains (Wilson *et al.*, 1991). Proper representation of modulation in temporal discharge patterns of the auditory nerve is an important goal in these strategies.

Despite the popularity of CIS schemes, the responses of auditory nerve fibers to a sinusoidal modulation of an electric pulse train can be very different from responses to a tone in a healthy ear. For modulation frequencies below 500 Hz, virtually every stimulated neuron is likely to entrain to the modulator (i.e., to produce a spike discharge for every modulator cycle). In contrast, neurons of a healthy ear responding to an acoustic tone fire at random multiples of the stimulus period. For example, there may be 1, 2, 3 or more cycles between successive spikes (Rose *et al.*, 1967). The situation is even worse at higher frequencies, because, with electric stimulation, neurons may fire on every other cycle or even higher multiples of the modulation period. If most stimulated neurons fire together, then the population of auditory neurons would code a submultiple of the modulator frequency rather than the actual frequency (Wilson *et al.*, 1997). Rubinstein *et al.* (Rubinstein *et al.*, 1999) proposed that the coding of modulation waveforms might be improved by introducing a sustained, high-frequency, ''desynchronizing'' pulse train (DPT) in addition to the modulated pulse train. The rationale for the DPT is that across-fiber differences in refractory, sen-

[a)]Portions of this work were presented as a poster in the ASILOMAR Conference in Monterey, California, 1999.
[b)]Electronic mail: lmlitvak@cirl.meei.harvard.edu

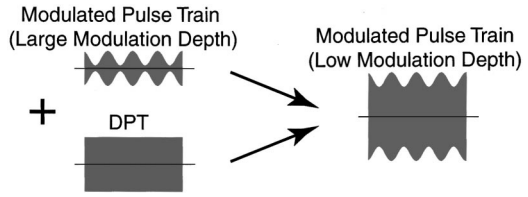


FIG. 1. The left panel shows two stimuli generated by a cochlear implant to implement a DPT protocol, as suggested by Rubinstein *et al.* (1999). This stimulus is composed of two pulse trains: one, strongly modulated, is the CIS signal, and the other, unmodulated, is the DPT. The carrier for the two signals need not be identical. We assume that the DPT-enhanced stimulus can be modeled by a carrier of the same frequency as the DPT that is more weakly modulated than the original CIS stimulus (right). This approximation is exact if the CIS and the DPT have identical carriers.

sitivity and other properties, as well as noise present in the neural membrane will result in the responses across fibers being desynchronized after the first few hundred milliseconds of DPT stimulation. Such desynchronization would lead to an improved representation of the modulator in temporal discharge patterns by allowing an ensemble of neurons to encode the true modulator frequency rather than a submultiple.

In this paper, we report the responses of auditory nerve fibers to both modulated and unmodulated electric pulse trains that were recorded to test the ideas underlying the DPT. We focused on two specific questions:

(1) Do the responses to a sustained high-frequency pulse train resemble spontaneous activity? Specifically, we characterized interval histograms (IH) for pulse trains and compared them to the nearly exponential histograms observed for spontaneous activity in an intact ear (Kiang *et al.*, 1965). We also quantified the variability in the spike count from presentation to presentation, and compared it to the variability expected for normal spontaneous activity.
(2) Does use of a high-frequency DPT result in a better representation of modulation frequency? We used modulated high-frequency pulse trains with low modulation depths ($\leqslant 10\%$) to imitate the effect of a DPT. We assumed that neural responses to a high-frequency pulse train with a low modulation depth are similar to those elicited by a stimulus that is a sum of a sustained DPT and a highly modulated pulse train (Fig. 1). This assumption may hold if the membrane time constant is large compared to the intervals between pulses. We compared period and interval histograms of responses to electric, modulated pulse trains with acoustic responses to tones.

II. METHODS

A. Animal preparation

Cats were first anesthetized using dial in urethane (75 mg/kg). Co-administration of kanamycin (subcutaneous, 300 mg/kg) and ethacrinic acid (intravenous, 25 mg/kg) was then used to deafen the animals (Xu *et al.*, 1993). The bulla was opened to expose the round window. An intracochlear stimulating electrode was inserted about 10 mm into the cochlea through the round window. The electrode was a 400 μm Pl/Ir ball. A similar electrode was inserted into the base of the cochlea for compound auditory potential (CAP) recordings. The round window was then sealed using connective tissue. The ear bar was used as a reference electrode for both stimulation and recording.

In order to verify that the animal was deafened, we measured a CAP in response to acoustic clicks (condensation, 100 μs). The CAP was measured in the implanted ear. In all cases, no CAP was noted for the highest click levels ($\sim$90 dB SPL) investigated.

B. Stimuli

Stimuli were delivered through an isolated current source and were either (1) unmodulated pulse trains (150 ms or 250 ms duration) with pulse rates of 1.2, 2.4, 4.8 or 24 kpps (kilo pulses per second) or (2) ''modulated'' 4.8 kpps pulse trains (first 50 ms or 150 ms unmodulated; last 100 ms modulated; modulation frequency: 400 Hz) of varying modulation depth. In all cases, pulse trains consisted of cathodic-anodic (CA) biphasic pulses (20.8 μs per phase). Modulated stimuli were modulated ''down'' such that the peak amplitude was equal in the modulated and the unmodulated portion of the stimulus. Stimuli were presented at a repetition rate of 1 per second. Stimulus level was adjusted to obtain discharge rates of 50 to 400 spikes/s. All stimulus levels reported in this study are peak currents.

C. Recording techniques

Standard techniques were used to expose the auditory nerve via a dorsal approach (Kiang *et al.*, 1965). We measured from single units in the auditory nerve using glass micropipettes filled with 3M KCl. A digital signal processor (DSP) was used to separate neural responses from the stimulus artifact (voltage excursions recorded at the micropipette that are not due to neural discharges). First, we recorded the ''artifact'' at a subthreshold stimulus level. Then, a scaled version of the recorded ''artifact'' was subtracted from the incoming waveform in real time. The gain applied to the recorded waveform was adjusted to optimally match the recorded and the incoming waveforms. The operation of recording the artifact was repeated for each neuron and for each stimulus studied. An important assumption of this technique is that the artifact grows linearly with stimulus level. Consequently, nonlinearities in the conducting medium, stimulation system or the recording equipment decrease the effectiveness of the cancellation for levels that are significantly above the neural threshold. In a saline solution, the stimulus artifact could be cancelled effectively at up to 6 dB above the recorded level. In an actual experiment, however, time constraints in finding the highest level at which there are no spikes, instability of the artifact waveform, the nonlinearity of biological tissue and contributions of nonlinear gross evoked responses limited the levels that could be investigated to no more than 2.5 dB above fiber threshold.

Times of the spike peaks were measured with 1 μs precision, and recorded in computer files for both on-line and off-line analysis.

D. Unit selection criteria

Possible hair-cell mediated activity (''electrophonic hearing'') complicates the interpretation of responses of ANFs to electric stimulation. Hair cells might not be completely eliminated by the acute deafening protocol used in this study. To minimize the effect of any remaining hair cells, only units that (1) had no spontaneous activity, (2) did not respond to an acoustic click at 90 dB SPL and (3) had unimodal, short-latency (<1 ms) PSTs in response to an electric biphasic pulse (CA, 20 μs per phase) were included for further analysis. The last criterion is based on the observation that when the hair cells are intact, responses to single pulses may contain late components that are hair-cell mediated (Moxon, 1967, 1971; van den Honert and Stypulkowski, 1984; Javel *et al.*, 1987).

Electrically stimulated ANFs can exhibit long-term adaptation with the time scale of seconds (Moxon, 1967). This implies that for relatively short repetition times, the responses can undergo a steady change in discharge rate from one run to the next. Although we used reasonably long (750 ms) pauses between successive presentations, some units still showed statistically significant adaptation in discharge rate throughout the measurement. Records were included in the analysis only if the correlation between the number of spikes per presentation and presentation time was not significantly different from 0 ($P>0.01$).

Electrically stimulated ANFs also adapt on a shorter time scale. Responses of ANFs to sustained electric stimuli typically show a strong initial response followed by a gradual decrease in discharge rate over the course of 30–100 ms (Moxon, 1967; Killian, 1994). For analysis of interspike intervals, we selected a window in which the discharge rate is nearly constant. We developed a recursive algorithm for selecting such a window automatically. The algorithm begins by selecting a 10 ms window centered 135 ms after the onset of the 150 or 250 ms stimulus. For each step, the mean and variance of the spike count in the current window is compared to the mean and the variance of the spike count in the adjacent 10 ms windows using the permutation test (Efron and Tibshirani, 1993). If the two measures are not significantly different in the two windows [two-sided achieved significance level (ASL) >0.03], the adjacent window is appended to the steady state window, and the procedure is repeated. The algorithm stops when both adjacent windows are rejected, or when the end of the stimulus is reached. Responses to electric pulse trains in the computed window were used to compute histograms that were compared with those for spontaneous activity in a healthy ear.

III. RESULTS

We recorded from 106 single units in 5 acutely deafened, anaesthetized cats. Spontaneous activity was present in 13 units. Spontaneous activity can be detected in deafened preparations with no remaining hair cells (Shepherd and Javel, 1997). Nevertheless, units with spontaneous activity were not included in the consequent analysis. No unit responded to acoustic stimulation or exhibited a multi-modal response to a single biphasic pulse. These results suggest that the deafening protocol was largely effective in eliminating hair-cell mediated activity.

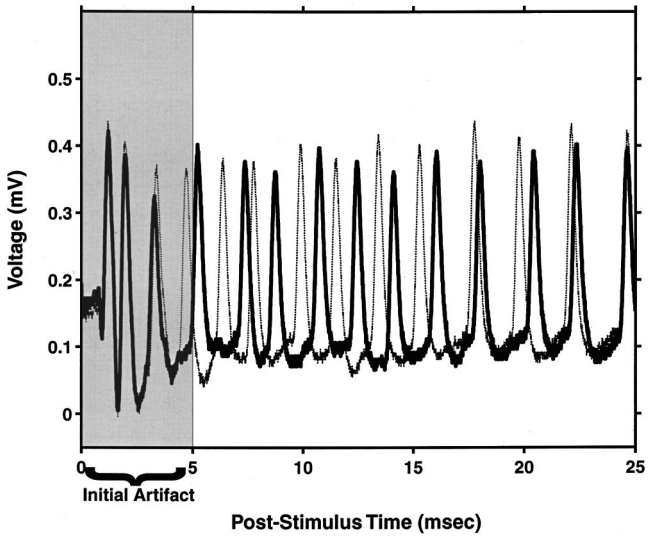


FIG. 2. Example of the raw spike waveforms derived after applying the cancellation strategy to remove the artifact. Responses to two presentations of the stimulus are shown (dark and light lines). The artifact template was learned at approximately 1 dB below the stimulus level used to generate the plotted responses. The gray area shows the portion of the initial response where the artifact is typically large. Both the initial peak, as well as the change in the envelope of the response during the first 5 ms are related to the residual artifact.

Artifact cancellation was applied in real time to the neural recordings. An example of the output of the cancellation is shown in Fig. 2. For the data presented in this paper, the artifact was at most 5% of the spike height. Because the cancellation technique was not effective in canceling the artifact for the earliest responses to high-rate pulse trains, responses that occurred within 5 ms of the stimulus onset were not analyzed.

The presentation-to-presentation stability test (Methods, Sec. D) was applied to 430 spike records. The test rejected nearly 40% of the records. Thus, despite our relatively long interstimulus times (greater than 750 ms for 250 ms stimulus) adaptation of single units from presentation to presentation can still be significant. Responses to longer stimuli (250 ms) were less stable than responses to shorter stimuli (150 ms), as were responses at higher discharge rates.

For the records rejected, we frequently observed that the response rate in the first 10–20 presentations was significantly different from the later responses. For these records, the stability analysis was repeated for the late responses only. Using this less stringent criterion, we were able to include the late responses for an additional 7% of the records. It should be pointed out that although the results in this paper are based only on analysis of the records that passed the latter stability test (67% of the data), the conclusions are not changed if the rejected records are included in the analysis. The primary purpose for excluding the unstable records is to demonstrate that instability in the recordings cannot account for the results reported in this study.

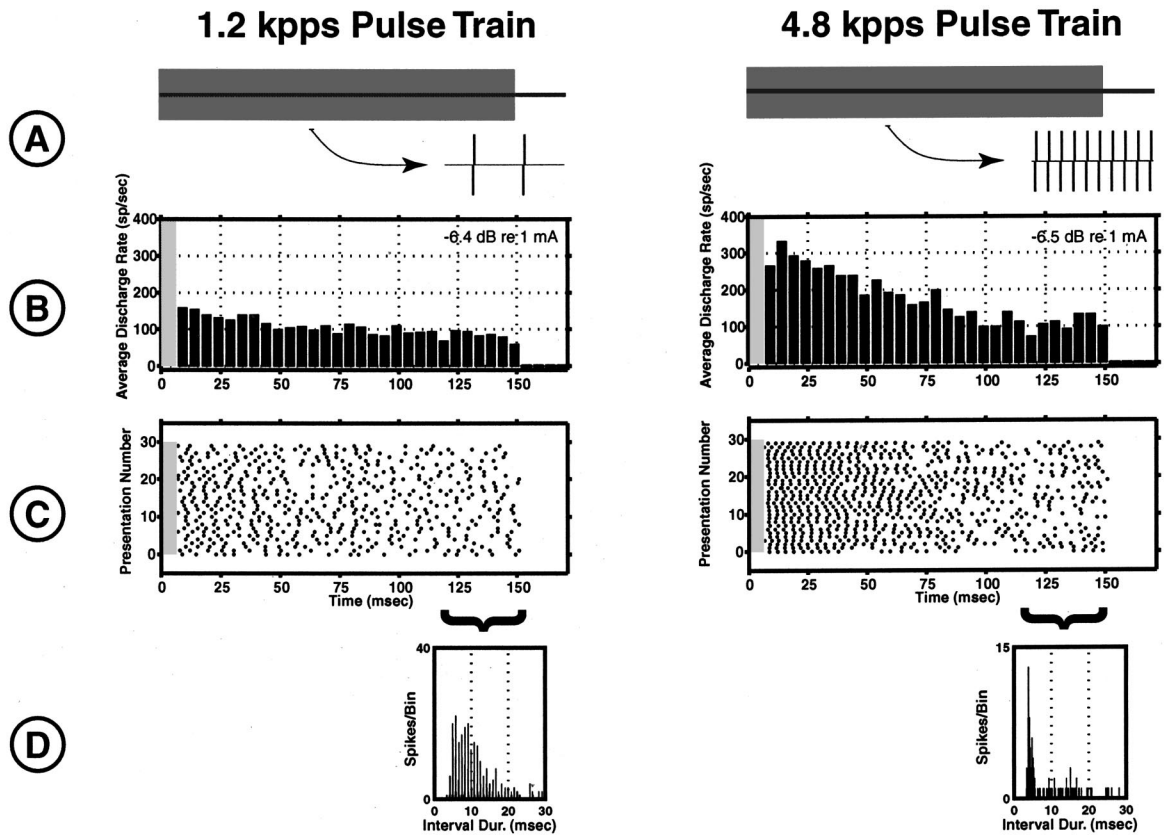


FIG. 3. Responses of a fiber to two unmodulated pulse trains of similar levels with pulse rates of 1.2 kpps (left) and 4.8 kpps (right). Panel A shows each stimulus with an expanded time segment of 2 ms. Panel B shows the PST histogram of the unit's responses. Gray bars indicate areas where spike data were discarded due to a large stimulus artifact. The response to the 4.8 kpps pulse train shows greater adaptation during the 150 ms window than the response to the 1.2 kpps stimulus. Panel C shows the dot raster plots of the responses. The desynchronization of the responses towards the later part of the stimulus is apparent. Finally, the last panel shows the interval histogram computed from responses that occurred during the last 30 ms of each stimulus. While the envelope of the IH computed from responses to the 1.2 kpps pulse train is similar to Poisson, the envelope of the IH computed from the responses to the 4.8 kpps pulse train is very different, showing a strong mode at 5 ms, followed by a longer tail.

A. Unmodulated pulse trains

Figure 3 shows responses of a fiber to two unmodulated pulse trains with pulse rates of 1.2 kpps (A, left) and 4.8 kpps (A, right). At this level, both stimuli evoke sustained responses from the unit. However, there is more adaptation in response to the higher-rate stimulus (B). This is a common finding in our data. For both pulse rates, responses are initially highly synchronized across trials, and become desynchronized over the course of the stimulus. This can be seen in the scatter of the response times from trial to trial in the dot raster plots (C). The interval histogram (IH, panel D) for the 1.2 kpps pulse train exhibits phase locking to the pulses, and a roughly exponential envelope. In contrast, the IH for the 4.8 kpps pulse train has a nonexponential envelope, with a pronounced mode at 5 ms. This mode is not related to the stimulus period, but is inversely related to the average discharge rate. At this coarse bin width (0.208 ms), phase locking to the pulse train is not apparent.

1. Adaptation

As used here, *adaptation* refers to a slow (on the order of 30 to 100 ms) change in the response discharge rate over the course of the stimulus. We found that adaptation is a function of pulse rate. Figure 4 shows the final rate (the discharge rate in the 10 ms window centered at 145 ms after the onset of the stimulus) versus the initial rate (the rate in a 10 ms window centered at 15 ms from stimulus onset). The responses during the first 10 ms were not included in the analysis. The solid black line indicates where the initial and the final rates are equal. Records falling on this line would show no adaptation.

The scatter in the points indicates that adaptation varies greatly across units, as reported in previous studies (Dynes and Delgutte, 1992; Killian, 1994). This variability is large even when the stimulus evoked comparable initial discharge rates. The dashed lines represent linear regressions for 1.2, 4.8, and 24 kpps stimuli. The slope of the 1.2 kpps regression line is significantly steeper than those for the 4.8 and 24 kpps ($p<0.001$, permutation test for 4.8 and 24 kpps). Thus, for stimuli that evoke similar initial discharge rates, the response tends to adapt less for pulse trains of lower pulse rate.

2. Dynamic range

In this study, level of the stimulus was adjusted for each fiber. While future cochlear implants might stimulate the auditory nerve more selectively, in the current designs a single stimulating electrode excites many fibers. If such an electrode is used to present DPT stimulation, differences in the responses of the stimulated fibers must be considered. Figure 5 plots the response rate in a 10 ms window centered 145 ms

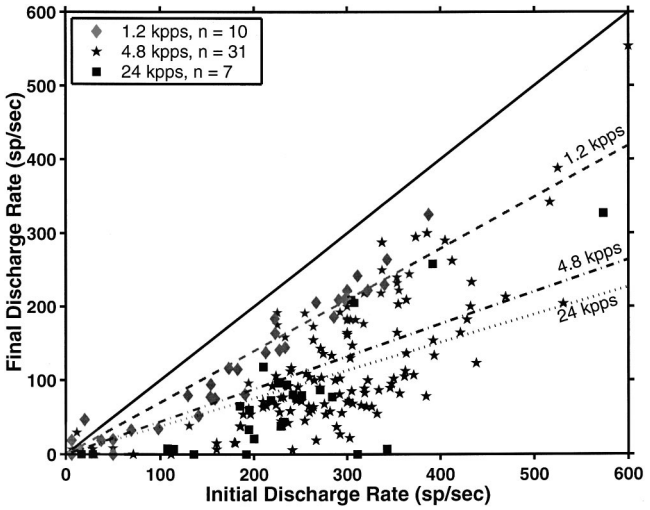


FIG. 4. Adaptation of an auditory nerve fiber response to an electric pulse train is a function of pulse rate. Adaptation is represented by the fact that the final discharge rate (rate at 140–150 ms) is lower than the initial discharge rate (rate at 10–20 ms). Each point is based on an average response over 20–40 stimulus presentations for one unit and stimulus level. In the legend, ''n'' is the number of units for which at least one record was included. Different symbols represent responses to different pulse rates of the electric stimulus. Linear regressions (represented by broken lines) were computed for each pulse rate and were constrained to include the origin. The 1.2 kpps data show significantly less adaptation than the responses to 2.4 or 4.8 kpps. The difference between the 4.8 and 2.4 kpps data is not statistically significant.

after stimulus onset as a function of level for eight fibers from the same animal. Each fiber responds at a rate appropriate for the range of spontaneous activity (gray area in the plot) over a limited range of levels (about 2 dB). Because the range of threshold across fibers for electric stimulation is 10–15 dB, a single electrode stimulus will, at best, result in only a small fraction (~20%) of the available ANFs responding at a rate appropriate for spontaneous activity.

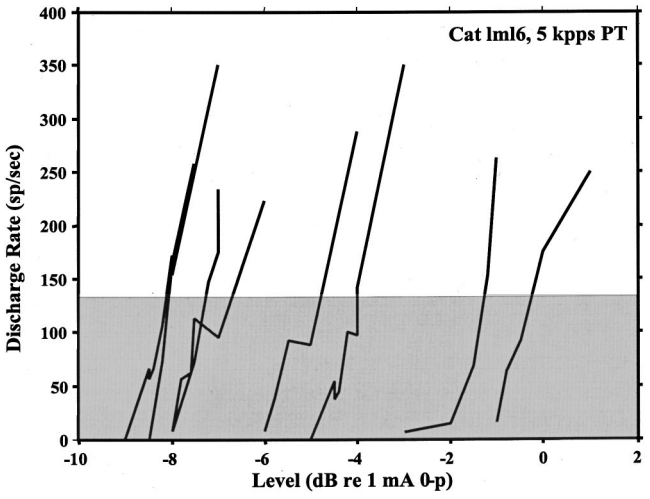


FIG. 5. Rate-level functions for eight fibers recorded from one animal as a function of stimulus current. Discharge rate is computed in a 10 ms window centered 145 ms after stimulus onset. The gray area indicates the range of rates that have been reported for spontaneously responding ANFs in an intact ear (Liberman, 1978). For each fiber, only a narrow (1–2 dB) range of stimulus levels result in a discharge rate that is appropriate for spontaneous activity.

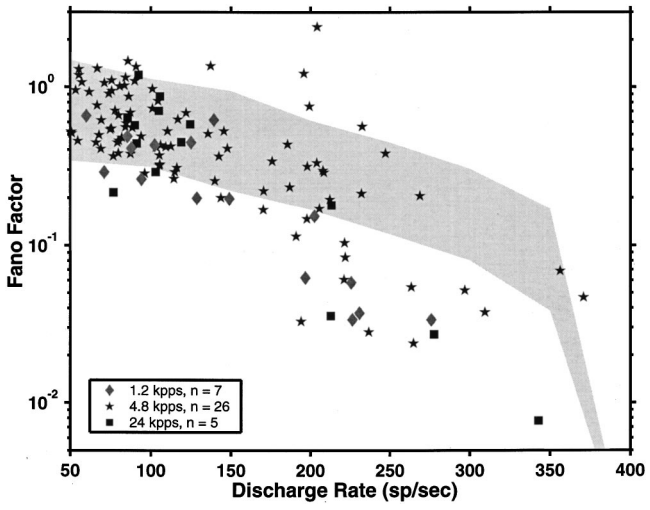


FIG. 6. Fano Factor (which characterizes variability in the stimulus from presentation to presentation) as a function of discharge rate. Measurements were made in the analysis window determined as described in the Methods section. The gray area represents the 99% confidence interval for the distribution expected from a Poisson process.

3. Variability

When stimulated acoustically, auditory nerve responses show pronounced variability in the number of spikes elicited from trial to trial. Similarly, spontaneous activity shows variability in the spike count from one time interval to another (Kelly *et al.*, 1993; Teich and Khanna, 1985). The variability can be quantified by the Fano Factor:

$$FF \equiv \frac{Variance[N_i]}{Mean[N_i]},$$

where N_i represents the number of spikes on trial *i*. For short (<100 ms) time intervals, the FF for normal spontaneous activity is consistent with a Poisson model with dead time of 2 ms (Kelly *et al.*, 1996). Figure 6 plots the FF for responses to unmodulated pulse trains of 1.2, 4.8 and 24 kpps as a function of average discharge rate in the ''steady-state'' window. For rates below 180 spikes/s, most points fall within the 99% confidence interval for a Poisson model with a dead time of 2 ms (shaded area). For higher discharge rates, the data tend to fall below the predicted range, indicating that there is less variability than expected for a Poisson model. In any case, for low and moderate discharge rates, variability in spike count from trial to trial is comparable with that for spontaneous activity.

4. Interval histograms

As indicated in Fig. 3, the shape of the interval histogram (IH) can depend on pulse rate. For the lowest pulse rate (1.2 kpps), IHs have an exponential envelope (Fig. 7, upper inset). An exponential shape is expected for Poisson discharges and is approximately consistent with the IHs for spontaneous activity in an intact ear. For high pulse rates, some but not all IH envelopes are clearly nonexponential, showing a sharp mode followed by a long tail (Fig. 7, left inset). To quantify the shape of the interval histogram, we fit the interval histogram with both a single exponential (dashed

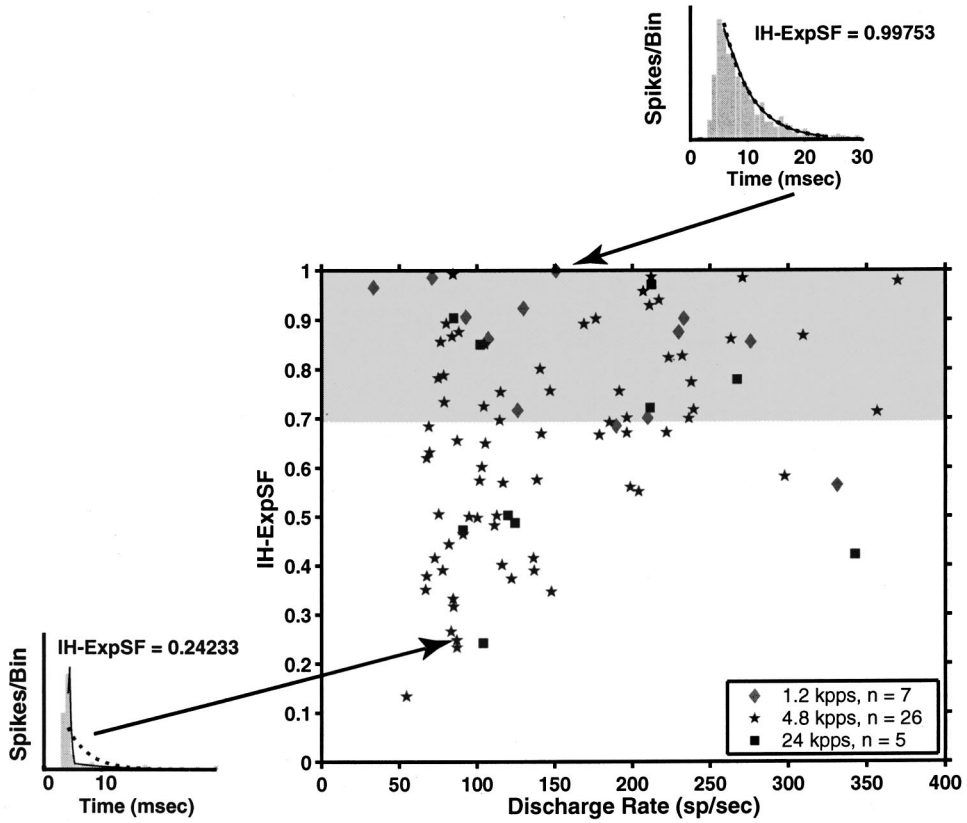


FIG. 7. Distribution of the IH-ExpSF as a function of the average discharge rate in the steady-state portion of the response. The filled area shows the region that would contain 99% of the units if their responses could be described by the Poisson process (see the Appendix).

line in the insets) and piecewise, with three exponentials (solid line in the insets). The numerical procedures used in fitting the data are described in the Appendix. We measured the root mean squared error of each fit to the data, and defined an IH exponential shape factor (IH-ExpSF) as the ratio of the error for the piecewise fit to that for the single exponential fit. The IH-ExpSF for samples from a Poisson process is approximately 1.

The scatter plot of Fig. 7 shows IH-ExpSF versus the average discharge rate for measurements made with pulse rates of 1.2, 4.8 and 24 kpps. For a pulse rate of 1.2 kpps, most points fall in the region expected for a Poisson model (shaded area, Appendix). For higher pulse rates, 50% of the data points are outside of the range expected for a Poisson process. Thus, only lower pulse rates consistently produce IHs that resemble spontaneous activity in intact ears.

One possible explanation for the difference between the 1.2 kpps and higher rate responses is the difference in adaptation during the analysis window (the window over which the rate is assumed to be approximately steady state). To test this possibility, Fig. 8 plots the IH-ExpSF versus adaptation in the analysis window. Adaptation is defined here as the average decrease of discharge rate. There is no obvious trend, suggesting that adaptation in the analysis window does not significantly alter the IH.

B. Modulated pulse trains

Figure 9 shows responses from a single unit to a sinusoidally amplitude modulated (400 Hz) pulse train (4.8 kpps) for modulation depths of 1% (left) and 10% (middle). These modulation depths might be representative of the modulation depths that would be used in a DPT-enhanced strategy. The smaller modulation depth is comparable to the psychophysical threshold to modulation in cochlear implant patients (Shannon, 1992). Pulse trains were modulated during the last 100 ms of the 150 ms train duration (row A). The levels of the two stimuli were adjusted to produce similar response rates during their modulated segment. For both stimuli, the response adapts over the first 50 ms while the pulse train is unmodulated. At the onset of modulation, average discharge

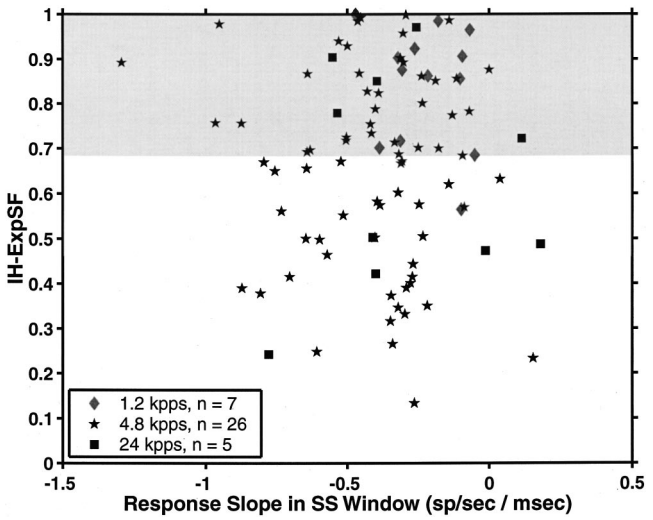


FIG. 8. Relationship between the measured IH-ExpSF and the adaptation that occurs within the analysis window. Adaptation is measured as the slope of the regression line to the PST histogram computed in the analysis window (bin width of PST computation 10 ms). The analysis window (referred in the legend as the SS window) is a window over which the discharge rate is considered approximately constant.

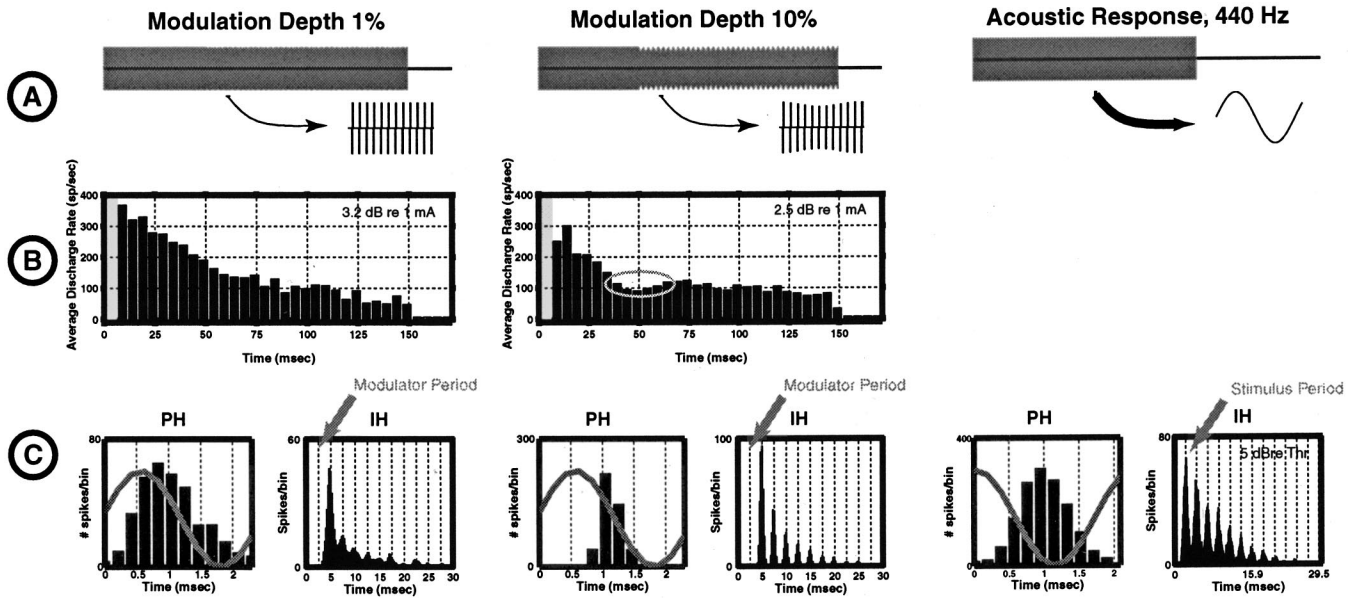


FIG. 9. The first two columns show the response of a neuron to a modulated pulse train for two modulation depths (1% in the left and 10% in the middle). For comparison, the right panel illustrates a response of a neuron (CF near 400 Hz) to an acoustic stimulus. Row A shows the stimulus waveforms. In the case of the left and the middle columns, the stimulus is unmodulated for the first 50 ms and is down modulated during the last 100 ms. Row B shows period histograms of the responses to the two electric stimuli. For the 10% modulation depth, the response actually increases after the onset of the modulation (gray oval). Finally, row C shows the period and the interval histograms computed during the modulated portion of the stimulus for the two electric stimuli. These histograms are also plotted for the responses of a fiber to the acoustic tone. Both electric responses are broadly similar to the acoustic response in their temporal properties.

rate increases for the modulation depth of 10% (row B, middle column; also, Fig. 10). This increase in rate is interesting because the rms current actually decreases (by 0.9 dB for 10% modulation) when modulation begins since the peak amplitude remains constant. Row C shows period and interval histograms computed from the responses measured during the modulated portion of the stimulus. For comparison, the right column shows both the interval and the period histogram computed from responses to a 440 Hz tone at a moderate level in a normal ear (from McKinney and Delgutte, 1998). For both modulation depths, the period histograms show pronounced modulation, although spikes are more precisely phase locked for the higher modulation depth. Even when the modulation depth is only 1%, the response is nearly

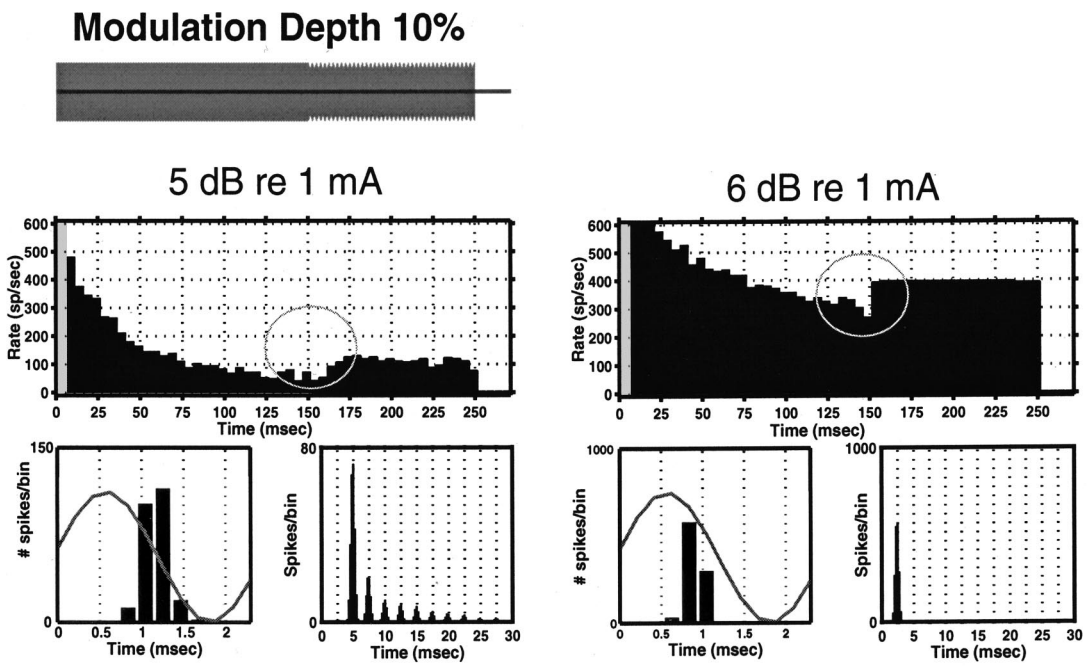


FIG. 10. Response of a fiber to a stimulus with a large (10%) modulation depth at supra-threshold (left panels) and saturation (right panels) levels. The topmost plot shows the stimulus. The middle left and right plots show responses at the lower (left) and at the higher (right) levels. Note that in both cases the rate increases after the modulation onset. The bottom row of plots shows the period (left) and the interval (right) histograms at each level for the responses measured during the modulated portion of the stimulus. Note that at the higher level, the neural response is entrained to the stimulus.

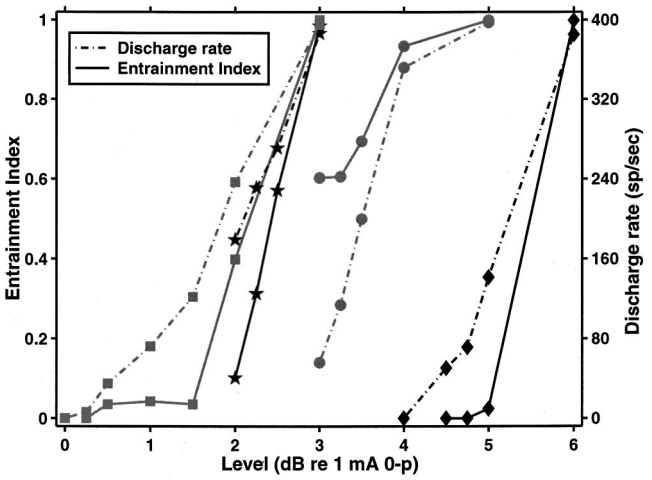


FIG. 11. Entrainment index (solid line) and discharge rate (dashed line) as a function of stimulus level for four units from the same animal. Entrainment index is defined as the ratio of intervals that fall between 1/2 and 3/2 of the stimulus period to the total number of intervals. An entrainment index of 1 represents perfect entrainment. At the highest levels, all four fibers entrain to the stimulus. This entrainment does not occur in a healthy ear in response to pure tones. At stimulus levels below those that evoke discharge rates of 200 spikes/s, the entrainment index is 0 for two units, indicating that the mode at the stimulus period is missing in the units' responses. In contrast, responses to pure tones in a healthy ear have entrainment indices between 0.6 and 0.9 (Joris *et al.*, 1994).

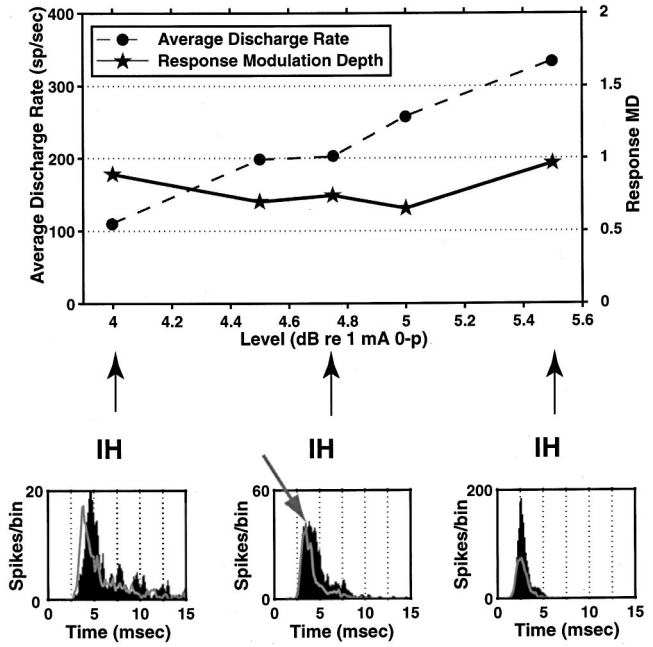


FIG. 12. Responses of a unit to a stimulus with modulation depth of 1%, modulation frequency of 400 Hz, at several levels. The top panel shows both the response rate and the modulation depth of the response as a function of level. Both the response rate and the response modulation depth are computed during the modulated portion of the stimulus. Note that the modulation depth of the response is nearly independent of level. The lower panels show the IHs computed from the responses at three different levels. The solid areas represent the IH computed from responses during the modulated portion of the stimulus while the gray lines are the IHs computed from responses preceding the modulation. At the lowest level, the mode of the IH for the modulated portion is related to the modulation frequency. At the higher levels, however, the mode of the IH for the modulated portion is similar to the mode described earlier for the unmodulated responses and not to the modulation frequency.

fully modulated. Furthermore, for this modulation depth, the period histogram is nearly sinusoidal in shape. This suggests that the responses may be representing the details of the sinusoidal modulator waveform.

Phase locking can be seen in the interval histogram as the clustering of modes around multiples of the stimulus period, which are shown here as dashed lines. For both modulation depths, the mode distribution is broadly similar to that for tones. Close examination reveals several differences. First, a pronounced mode at the modulation period is absent in the electrical case (three arrows). Second, the mode at twice the modulation period is strongly exaggerated for the smaller modulation depth. This exaggeration may be related to the preferred interval near 5 ms found for unmodulated pulse trains.

1. Large modulation depth: Entrainment

Figure 10 shows the response of another unit to two different levels of a pulse train modulated at a depth of 10%. The pulse train was 250 ms long, and was modulated only in the last 100 ms. The increase in rate at the onset of modulation is more pronounced for this unit than for the unit in Fig. 7. For this modulation depth, we observed increases in rate at the onset of modulation in 80% of the units studied.

At the lower level, the distribution of the modes in the interval histogram is similar to that for the acoustic responses to tone shown in Fig. 9. However, the mode at the stimulus period is again missing in the response to the electric stimulus. At the higher level, responses entrain to the modulator frequency, as indicated by a single mode at the modulation period in the IH.

To quantify the difference between acoustic and electric interval histograms, we computed an *entrainment index* (Joris *et al.*, 1994). The entrainment index is defined as the ratio of the number of the intervals that are between 1/2 and 3/2 of the stimulus period to the total number of intervals in the interval histogram. Figure 11 plots the entrainment index (solid line) along with the average discharge rate (dashed line) as a function of stimulus level for four units from one animal. The stimulus was the same as in Fig. 10. At levels that evoked a moderate discharge rate (less than 200 spikes/s), two out of four units had an entrainment index near zero, indicating that the mode at the stimulus period was either very small or entirely missing from the interval histogram of these units. In contrast, the entrainment index computed from responses to tones recorded from ANFs in a healthy ear is between 0.6 and 0.9 for tone frequency near 400 Hz (Joris *et al.*, 1994). When the stimulus level was at least 2 dB over the level that evoked a discharge rate of 100 spikes/s, units entrained to the modulator (had an entrainment index of 1). Such entrainment is never seen in the responses of ANFs in an intact ear to tones of that frequency (Joris *et al.*, 1994; Rose *et al.*, 1967). Thus, except for an extremely small range of levels, interval distributions for responses to electric modulated pulse trains can differ from the distributions for responses to pure tones recorded in an intact ear.

2. Small modulation depth

Figure 12 shows average discharge rate and response

modulation depth (which is a measure of modulation of the period histogram, and is twice the synchronization index) for a single unit as a function of level. The stimulus was a 250 ms pulse train (4.8 kpps) that was modulated by a 400 Hz sinusoid in the last 100 ms (modulation depth 1%). Average discharge rate increased three-fold over the 1.5 dB range of levels. In contrast, response modulation depth was nearly constant at 0.7–0.8 for all levels, indicating robust phase locking. The first mode in the interval histogram shifts to the left as the level increased, reflecting the increase in average rate (lower panels). While the first mode is a multiple of the modulator period at the lowest and at the highest level shown, at the intermediate level of 4.75 dB *re*: 1 mA, it falls nearly halfway between the first and the second multiple of modulator period (arrow). This result suggests that at this level, the first mode is not related to the modulator frequency. We hypothesize that for the higher levels, the first mode is related to the preferred firing periods demonstrated earlier for high-frequency unmodulated stimuli. To test this hypothesis, we computed IHs from responses that immediately precede the modulation onset (gray line). As predicted by the hypothesis, the location of the mode of the IH for these responses roughly matches the mode observed during the modulated segment at levels above 4.75 dB *re* 1 mA. At the lower level, however, the first mode of the modulated segment differs from the first mode of the pre-modulation responses and is related to the modulation frequency.

Thus, for very low modulation depths, interval histograms can differ from IHs evoked by acoustic tones because their IHs can include modes that are not multiples of the modulator period.

IV. DISCUSSION

A. Adaptation

The adaptation that we report here is generally consistent with previous studies of electrically stimulated ANFs (Moxon, 1967; van den Honert and Stypulkowski, 1987; Javel *et al.*, 1987; Dynes and Delgutte, 1992; Killian, 1994). The result that at least short-term (<150 ms) adaptation depends on pulse rate of the electric stimulus is new. This result may have important consequences for models of electrically stimulated ANFs.

Our working hypothesis is that the slow decrease in discharge rate over time is evidence either of depletion of an ''excitatory'' agent or accumulation of an ''inhibitory'' agent. For example, intracellular sodium and extracellular potassium might be the relevant agents for the electrically stimulated ANFs. The changes in the concentrations of these substances may occur because of change in the membrane conductance resulting from the increased spike activity induced by the electric stimulus. Alternatively, concentrations may change even without increased spike activity because the electric stimulus can induce voltage changes across the neural membrane that do not lead to spikes. If the increased activity itself were the primary influence on accumulation or depletion of the agent, we would expect that adaptation would depend primarily on the initial response rate and not on the pulse rate of the electric stimulus. We showed that adaptation of ANFs to electric stimulation depends not only on the initial rate, but also on stimulation frequency. This implies that significantly more agent is depleted due to voltage changes across the membrane that do not lead to spikes during the higher-frequency stimulus. Whether neurons will be able to readjust to the new homeostatic balance if the high-frequency DPT is presented continuously is a topic for future investigations.

B. Temporal response patterns to unmodulated pulse trains

We found that the detailed response pattern of auditory nerve fibers to electric pulse trains also depends strongly on stimulation rate. In particular, while the envelope of the IH of responses to a 1.2 kpps pulse train is nearly exponential, the envelope can strongly differ from an exponential for pulse rates above 4.8 kpps. This conclusion is consistent with published data. Van den Honert and Stypulkowski (1987) described ANF responses to moderate-level electric sinusoids with frequencies between 100 and 1000 Hz. The interval histograms in these data appear exponential. IHs reported for responses to low-rate (<2000 pps) pulse trains appear exponential as well (Javel *et al.*, 1987). Dynes and Delgutte found Gaussian-like interval histograms of responses of ANFs to electric sinusoids for frequencies above 4 kHz (Dynes and Delgutte, 1992). The mode of their IHs is at about 5–6 ms, as in our data. Unfortunately, they do not report IHs for responses with discharge rates below 200 spikes/s, where the tail may be more apparent.

C. Does the response to a desynchronizing pulse train (DPT) resemble spontaneous activity?

Rubinstein *et al.* (1999) suggested introducing a continuous pulse train into CI strategies to produce neural responses resembling normal spontaneous activity. Our results indicate that if the level of the DPT can be adjusted for each fiber, the responses to sustained high-frequency pulse trains resemble spontaneous activity in some, but not all respects. The variability across stimulus presentations is comparable with that expected for spontaneous activity. For a relatively low pulse rate (1.2 kpps) the envelope of interval histograms resembles those for spontaneous activity for most units. However, for higher-rates (4.8 kpps and above) interval histograms can clearly deviate from those for spontaneous activity, showing a sharp mode that is followed by a long tail. On the other hand, a DPT with a low pulse rate may introduce psychophysically significant periodicity in neural responses that is related to the DPT period. Thus, intermediate pulse rates (above 1.2 and below 4.8 kpps) may be optimal for imitating spontaneous activity.

A further difficulty with the DPT idea is that electrically stimulated fibers respond with discharge rates that are appropriate for spontaneously responding neurons in a healthy ear only over a narrow range of stimulus levels. The small dynamic range reported in this study is consistent with earlier reports of responses of ANFs to electric pulse trains (Javel *et al.*, 1987; Parkins, 1989). Fibers can differ in threshold by 10 dB or more at a single cochlear location near the stimu-

lating electrode (van den Honert and Stypulkowski, 1984). If a DPT is presented at a level that stimulates a large percentage of the fibers, most fibers will respond at rates that are higher than the rates appropriate for producing spontaneous-like responses.

While the differences in DPT sensitivity across fibers are large compared to the level range producing spontaneous-like responses, the DPT idea should not necessarily be ruled out based on the results of this study of short-term (less than 250 ms) stimulation. Electric responses are known to adapt over a course of minutes (Moxon, 1967; Killian, 1994). This longer adaptation may selectively decrease the response discharge rates of fibers with low thresholds to the DPT.

D. Does a high-frequency DPT help encode modulation frequency?

We showed that a modulated pulse train with low ($<10\%$) modulation depth can produce interspike and period histograms resembling responses to tones in intact ears. If we interpret those stimuli as consisting of a DPT plus a highly modulated signal, this result suggests that realistic responses to sinusoidal stimuli might be obtained with a DPT. However, the realistic, tonelike responses are only observed over a narrow range of stimulus levels. Furthermore, for very low (1%) modulation depths, there can be preferred intervals unrelated to the modulation frequency which may be confusing to the central processor. Interestingly, these intervals may co-exist with a high level of phase locking to the modulator frequency.

E. Implications for mechanisms

Our results with unmodulated and modulated stimuli also pose important questions for models of ANF responses to electric stimulation. Published reports based on biophysical models have not described the complex shape of interval histograms (a pronounced mode with a long tail) we observed in responses to high-frequency pulse trains (e.g., Rubinstein *et al.*, 1999). In evaluating mechanisms that may account for this discrepancy it is important to determine whether the relevant processes occur on an interval-by-interval basis or on a slower time scale (e.g., bursting). In a limited number of units for which sufficient data are available, we failed to detect a correlation between consecutive intervals. Thus, the relevant processes appear to occur on an interval-by-interval basis.

Responses that are qualitatively similar to the ones that we report for cat ANFs, in which rapid firing randomly alternates with longer pauses, have also been observed in the squid giant axon that is stimulated by a dc current injection (Guttman and Barnhill, 1970; Guttman *et al.*, 1980). The mechanism for these responses may be similar to the mechanism responsible for the nonexponential responses reported here. A biophysical model that has Hodgkin–Huxley channel dynamics, and that explicitly models random channel noise can account for the observed responses in that preparation (Schneidman *et al.*, 1998). Therefore, it is possible that a biophysical model of ANFs that correctly captures ANF channel dynamics and that explicitly models channel noise may predict the responses that we observed as well.

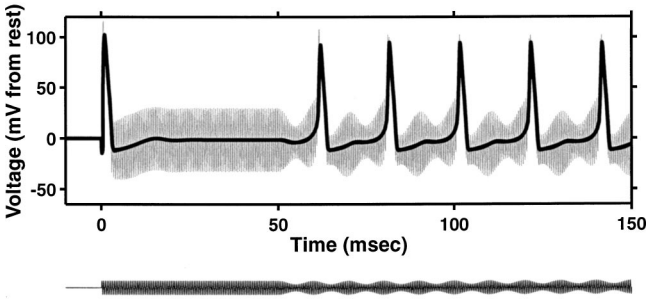


FIG. 13. Response of a current-clamped Hodgkin–Huxley model to a high-frequency sinusoid that is modulated in the last 50 ms of the 100 ms stimulus (bottom). Parameters for this model are summarized in Weiss (1996), p. 191. Simulation temperature was 6.3 °C. The model was implemented in MATLAB and simulated numerically using a 5 μs step. The stimulus was presented as intracellular current injections. The carrier was a 2 kHz sinusoid (400 μA/cm^2 0-peak) that was modulated ''down'' (sinusoidal 50% modulation, 100 Hz modulation frequency). This stimulus was chosen to represent the modulated pulse trains that were used in the physiological study. However, both the shape of the ''pulse'' and the carrier frequency were adjusted to account for the slower dynamics of the Hodgkin–Huxley model. The time course of the membrane voltage is shown by the gray line. The membrane voltage is shown relative to its resting state. We also show the time course of the membrane voltage averaged over a period of the sinusoidal carrier (dark line). The unmodulated response can be characterized by a single response at the stimulus onset. However, when down modulation is added, the model responds in a sustained manner to every second modulation cycle.

Although high-frequency pulses used in this paper differ from dc injections used in the study of the squid giant axon, some parallels may exist between the responses evoked by both. In particular, rectification and low pass filtering, both of which are known to occur in neural membrane, may transform the extracellular pulse train into an intracellular stimulus with a significant dc component. Intracellular recordings near the site of excitation are necessary to determine mechanisms underlying the observed responses.

Another intriguing observation is the increase in average discharge rate after the ''down'' modulation is turned on. The increase in average discharge rate is surprising because the rms stimulus current actually decreases at the onset of the modulator. To our knowledge, such an increased response to modulation has not been reported for biophysical models. Our simulations using the Hodgkin–Huxley (HH) model indicate that an increased response rate can be observed at the onset of the ''down'' modulation of a high-frequency carrier (Fig. 13), although the model does not mimic the data quantitatively.

One difference between the HH model and the data is that the HH model exhibits an onset-only response to the sustained high-frequency stimulus. It is unknown whether introducing channel noise into the model would improve the fit between the model and the data. While the Frankenhaeuser–Huxley model (Frankenhaeuser and Huxley, 1964) of the frog sciatic nerve predicts the sustained response during the high-frequency stimulus, it does not show an increase of firing rate at the onset of the down-modulation (personal observations). Future modeling studies would also be needed to determine whether mammalian-

based models produce responses to modulated stimuli that more closely resemble our data for ANFs.

F. Implications for the cochlear implant processor

The purpose for introducing a DPT is to improve the coding of modulation in auditory nerve responses. Some aspects of our data appear promising in that respect. A DPT can produce responses that are desynchronized across trials, suggesting that the responses may also be desynchronized across different fibers, as occurs for normal spontaneous activity (Johnson and Kiang, 1976). Desynchronization of auditory nerve responses may lead to an improved temporal coding of stimuli with rapid onsets and high frequencies. For low pulse rates of 1.2 kpps, responses to a DPT imitate some characteristics of spontaneous activity. Over a range of levels and modulation depths, a DPT can help provide temporal discharge patterns for amplitude-modulated stimuli that resemble responses to tones. Other aspects of the data are less promising. For example, we found that DPTs that accurately encode sinusoidal waveforms in the period histogram can show modes in the interval histogram that are not related to the modulator frequency. In addition, for some modulation frequencies and levels, the mode of the interspike distribution that corresponds to the stimulus period can be entirely absent in the interval histogram of the response. It appears, therefore, that successful use of a DPT depends on what exact aspects of the temporal discharge pattern (e.g., intervals versus cross-fiber synchrony) are most important to the central processor for extracting information about the stimulus. Because the exact mechanisms used by the central processor are not known, it is difficult to predict whether adding the DPT will result in an overall improvement in the reception of speech by cochlear implant users.

ACKNOWLEDGMENTS

We thank Dr. T. Weiss for providing his MATLAB implementation of the Hodgkin–Huxley model. Dr. Cariani assisted with the early experiments. Martin McKinney provided the recordings of responses of the auditory nerve to acoustic stimulation that were used in this paper. These experiments would not be possible without the surgical skills of Leslie Liberman.

APPENDIX

The interval histogram exponential shape factor (IH-ExpSF) is defined as the ratio of the error in the fit of the IH with the piecewise, three exponential function to the error of the fit with a single exponential. First, a number drawn from a uniform distribution from 0 to 0.832 ms was added to each interval to eliminate the effect of locking to the 1.28 kHz pulse frequency. Next, the mode of the interval distribution was estimated. Intervals that were longer than the mode were considered to constitute the tail and were included in the analysis; the remaining intervals were discarded.

Next, the IH was recomputed based on the remaining intervals. Instead of computing the interval histograms with a fixed bin size, however, we changed the bin size from bin to bin, such that exactly five intervals fell into each bin. This normalization is convenient because the coefficient of variation of the interval histogram for any given bin is proportional to the number of values that fall into that bin (Johnson, 1996). By normalizing the interval histogram in this way, one achieves equal coefficient of variation in each of the bins. Finally, a logarithmic transformation was applied to the IH so that straight-line fits generated exponential fitting function.

The resulting logarithmic histogram was fit by a linear and by a piecewise linear function by minimizing the least squares error. The piecewise linear function was constrained to be continuous. The times at which the function changed its slope were chosen such that each segment covered an equal number of bins.

To estimate IH-ExpSF distribution expected from a Poisson-with-dead time model, the IH-ExpSF was computed for 1000 interval histograms, each histogram composed of 200 independent ''intervals'' drawn from an exponential with dead time distribution (dead time of 3 ms). 99% of the generated data fell between 0.68 and 1. This boundary was nearly independent of the event rates of the simulated distribution for rates of 50 to 250 events/s.

Dynes, S. B., and Delgutte, B. (**1992**). ''Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea,'' Hear. Res. **58**, 79–90.

Efron, B., and Tibshirani, R. (**1993**). *An Introduction to the Bootstrap* (Chapman & Hall, New York).

Frankenhaeuser, B., and Huxley, A. F. (**1964**). ''The action potential in the myelinated nerve fibre of xenopus laevis as computed on the basis of voltage clamp data,'' J. Physiol. (London) **171**, 302–315.

Guttman, R., and Barnhill, R. (**1970**). ''Oscillation and repetitive firing in squid axons. Comparison of experiments with computations,'' J. Gen. Physiol. **55**, 104–118.

Guttman, R., Lewis, S., and Rinzel, J. (**1980**). ''Control of repetitive firing in squid axon membrane as a model for a neuroneoscillator,'' J. Physiol. (London) **305**, 377–395.

Javel, E., Tong, Y. C., Shepherd, R. K., and Clark, G. M. (**1987**). ''Responses of cat auditory nerve fibers to biphasic electrical current pulses,'' Ann. Otol. Rhinol. Laryngol. **96**, 26–30.

Johnson, D. H. (**1996**). ''Point process models of single-neuron discharges,'' J. Comput. Neurosci. **3**, 275–299.

Johnson, D. H., and Kiang, N. Y. S. (**1976**). ''Analysis of discharges recorded simultaneously from pairs of auditory nerve fibers,'' Biophys. J. **16**, 719–734.

Joris, P. X., Carney, L. H., Smith, P. H., and Yin, T. C. (**1994**). ''Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency,'' J. Neurophysiol. **71**, 1022–1036.

Kelly, O. E., Johnson, D. H., Delgutte, B., and Cariani, P. (**1996**). ''Fractal noise strength in auditory-nerve fiber recordings,'' J. Acoust. Soc. Am. **99**, 2210–2220.

Kiang, N. Y. S., Watanabe, T., Thomas, E. C., and Clark, L. F. (**1965**). *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (The MIT Press, Cambridge, MA).

Killian, M. J. P. (**1994**). ''Excitability of the electrically stimulated auditory nerve,'' Ph.D. thesis, University of Utrecht, Utrecht.

Liberman, M. C. (**1978**). ''Auditory-nerve response from cats raised in a low-noise chamber,'' J. Acoust. Soc. Am. **63**, 442–455.

McKinney, M. F., and Delgutte, B. (**1998**). ''Correlates of the subjective octave in auditory-nerve fiber responses: Effect of phase-locking and refractoriness,'' Abstract, Midwinter Meeting of the Association for Research in Otolaryngology, Florida.

Moxon, E. C. (**1967**). ''Electric stimulation of the cat's cochlea: A study of discharge rates in single auditory nerve fibers,'' Ph.D. Thesis, MIT, Cambridge, MA.

Moxon, E. C. (**1971**). ''Neural and mechanical responses to electrical stimulation of the cat's inner ear,'' MIT, Cambridge.

Parkins, C. W., (**1989**). ''Temporal response patterns of auditory nerve fibers to electrical stimulation in deafened squirrel monkeys,'' Hear. Res. **41**, 137–168.
Rose, J. E., Brugge, J. R., Anderson, D. J., and Hind, J. E. (**1967**). ''Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey,'' J. Neurophysiol. **30**, 769–793.
Rubinstein, J. T., Wilson, B. S., Finley, C. C., and Abbas, P. J. (**1999**). ''Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation,'' Hear. Res. **127**, 108–118.
Schneidman, E., Freedman, B., and Segev, I. (**1998**). ''Ion channel stochasticity may be critical in determining the reliability and precision of spike timing,'' Neural Comput. **10**, 1679–1703.
Shannon, R. V. (**1992**). ''Temporal modulation transfer functions in patients with cochlear implants,'' J. Acoust. Soc. Am. **91**, 2156–2164.
Shepherd, R. K., and Javel, E. (**1997**). ''Electrical stimulation of the auditory nerve. I. Correlation of physiological responses with cochlear status,'' Hear. Res. **108**, 112–144.
Teich, M. C., and Khanna, S. M. (**1985**). ''Pulse-number distribution for the neural spike train in the cat's auditory nerve,'' J. Acoust. Soc. Am. **77**, 1110–1128.
van den Honert, C., and Stypulkowski, P. H. (**1984**). ''Physiological properties of the electrically stimulated auditory nerve. II. Single fiber recordings,'' Hear. Res. **14**, 225–243.
van den Honert, C., and Stypulkowski, P. H. (**1987**). ''Temporal response patterns of single auditory nerve fibers elicited by periodic electrical stimuli,'' Hear. Res. **29**, 207–222.
Weiss, T. F. (**1996**). *Cellular Biophysics* (MIT Press, Boston, MA).
Wilson, B., Finley, C. C., Lawson, D. T., and Zebri, M. (**1997**). ''Temporal representations with cochlear implants,'' American Journal of Otology **18**, S30–S34.
Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (**1991**). ''Better speech recognition with cochlear implants,'' Nature (London) . **352**, 236–238.
Xu, S. A., Shepherd, R. K., Chen, Y., and Clark, G. M. (**1993**). ''Profound hearing loss in the cat following the single co-administration of kanamycin and ethacrynic acid,'' Hear. Res. **70**, 205–215.

Evidence for a behavioral significance of saccular acoustic sensitivity in humans

Neil Todd[a)]
Department of Psychology, University of Manchester, Manchester M13 9PL, United Kingdom

(Received 26 January 2001; revised 30 March 2001; accepted 3 April 2001)

In this article the results are reported of an experiment to provide direct evidence for a perceptual and behavioral significance of human saccular acoustic sensitivity. Ten human subjects were stimulated monaurally with 100-ms trains of 10-ms tone pulses with pulse repetition rate of 40 Hz, and were required to rate the pleasantness of the stimuli on a nine-point scale. The design included three within-subject factors: carrier frequency (two levels, 200 and 4000 Hz), intensity [13 levels from 55 to 115 dB(A) in 5-dB steps] and ear (left and right). For intensities above 90 dB myogenic vestibular evoked potentials (MVEP) were also obtained from the ipsilateral sternocleidomastoid muscle from which it was possible to obtain thresholds by linear regression of MVEP amplitudes against intensity. A further between-subjects factor was added which assessed subjects' attitude to vestibular sensations. The results indicate that across subjects there is a general trend of decreasing pleasantness with increasing intensity, but for the 200-Hz condition there is a significant positive departure from monotonicity in pleasantness ($p<0.05$) above the mean saccular threshold. However, when split by the between-subjects factor, the positive departure was only evident for those subjects who have a positive attitude to vestibular sensations ($p<0.01$). Implications of these results for human responses to loud sound and the possible evolutionary significance of saccular acoustic sensitivity are discussed.
[DOI: 10.1121/1.1373662]

PACS numbers: 43.64.Ri, 43.64.Tk, 43.64.Wn, 43.64.Cd [BLM]

I. INTRODUCTION

The conventional view in auditory science is that whatever the role of the otolith organs in lower vertebrates, in mammals hearing is mediated entirely by the cochlea. Evidence accumulated over the last few decades, though, supports the case that the sacculus has conserved an acoustic sensitivity throughout vertebrate phylogeny. For example, auditory sensitivity of the sacculus has been demonstrated in fish (Popper *et al.*, 1982), amphibians (Moffat and Capranica, 1976), reptiles (Lewis *et al.*, 1985), birds (Wit *et al.*, 1984), and, among mammals, in guinea pigs (Cazals *et al.*, 1983b), cats (McCue and Guinan, 1995), and squirrel monkeys (Young *et al.*, 1977). The acoustically responsive saccular fibers have irregular spontaneous discharge rates [and hence are termed acoustically responsive irregular discharge (ARID) units], a range of best stimulus frequencies lying between about 200 and 1000 Hz in the cat, and a rate threshold at about 90 dB SPL, while the synchronization threshold is about 20 dB lower (McCue and Guinan, 1994, 1995; Young *et al.*, 1977).

In humans saccular sensitivity may be demonstrated by means of myogenic vestibular eveoked potentials (MVEP) in a variety of muscles from the neck, arms, and legs. Following early work (Bickford *et al.*, 1964; Townsend and Cody, 1971), Colebatch *et al.* (1994) provided evidence for an otolithic component of an acoustically evoked, myogenic response by measuring EMG responses of the sternocleidomastoid (SCM) muscle evoked by loud clicks. They found two main complexes in the averaged response: p13-n23 and n34-p44. It was established in normal subjects that p13-n23 represented an ipsilateral response to unilateral stimulation, whereas n34-p44 reflected a bilateral response to unilateral stimulation. Further recordings from patients with unilateral and bilateral inner-ear deficits showed that the p13-n23 peaks were present in patients with bilateral sensorineural deafness, while n34-p44 peaks were absent. On the other hand, p13-n23 peaks were not present in recordings from patients with loss of vestibular function ipsilateral to the stimulation. These authors went on to conclude that whereas the first component, p13-n23, is generated by activation of vestibular afferents, the second component n34-p44 did not depend on the integrity of the vestibular nerve.

Since this work, there has been considerable other recent interest by a number of workers in myogenic vestibular evoked potentials (MVEPs). According to Ferber-Viart *et al.*'s review (1999) there has been general agreement (1) on the timing of the wave pattern in cervical muscles, (2) that the first component is more consistent than the second, and (3) that there is a linear dependence of response amplitude with sound pressure level. While there is still some considerable uncertainty about the origin of the second component, there is more agreement of the likely vestibular origin of the p13-n23 component. The most likely mechanism is that acoustic stimulation of the sacculus gives rise to inhibitory postsynaptic potentials in cervical flexor motor neurones (Uchino *et al.*, 1997) via an inhibitory interneurone in the vestibular spinal tract. This manifests as a positive-going wave in monopolar recordings.

The question remains, however, whether saccular acous-

a)Electronic mail: todd@fs4.psy.man.ac.uk

tic sensitivity in humans has any perceptual or behavioral significance. It is quite possible that acoustic sensitivity of the sacculus does not have a role or function in higher vertebrates, but is simply an epiphenomenon of evolution, a consequence of there being no selection pressure for it to disappear. And indeed it has been argued that MVEPs do not seem to play a role in balance control (Di Lazzaro *et al.*, 1995), i.e., they do not have a normal vestibular function. However, several hypotheses have been advanced to account for saccular acoustic sensitivity in higher vertebrates. McCue and Guinan's (1995) hypothesis is that, in addition to cervical motoneurones, the saccular afferents may also give rise to a response in middle-ear muscles such as stapedius which might serve an ''antimasking'' function of low-frequency on high-frequency tones. While it is possible that there is a vestibular pathway to stapedius, since the saccular tuning curve partially, but not completely, overlaps stapedius tuning curves in cats, there is no direct evidence to support this hypothesis. Further, the postauricular response, which is mediated by the facial (VII) nerve in common with the middle-ear response, does not appear to have an otolithic component (Townsend and Cody, 1971). In contrast the SCM is innervated from the cervical spinal cord and the accessory (XI) nerve.

An alternative (though not necessarily mutually exclusive) hypothesis proposed by Todd *et al.* (2000) is that saccular acoustic sensitivity contributes to the affective quality of loud sounds. Circumstantial evidence for this in humans comes from the coincidence of saccular acoustic sensitivity with the distribution of frequencies and intensities of natural acoustic signals (Todd and Cody, 2000; Todd *et al.*, 2000). In particular, given that humans seek out pleasurable sensations of self-motion by normal inertial stimulation, such as at ''fun parks,'' it is possible that human compulsion to exposure to loud, low-frequency sounds is a kind of acoustic equivalent of vestibular self-stimulation.

In the work presented in this article we seek to address the question of whether human saccular sensitivity has any perceptual or behavioral significance. One way in which this may be tackled is by looking for any evidence of a change in the pattern of frequency or intensity discrimination in the region of saccular sensitivity, since there is anatomical evidence of a projection from the saccular nerve into the cochlear nucleus (Burian *et al.*, 1989). For example, Cazals and co-workers (Cazals *et al.*, 1982, 1983a) have determined both intensity difference thresholds and frequency selectivity in guinea pigs after total or selective destruction of the organ of Corti. However, in practice it would prove very difficult in humans to dissociate any specifically saccular contribution from cochlear input to discrimination, even in profoundly deaf subjects.

An alternative approach is to look for a change in the quality of sensation above the saccular threshold. The study of affective or hedonic responses to sensory stimuli can be traced to the work of Wilhelm Wundt. In his book *Introduction to Physiological Psychology* (1910), Wundt proposed a generalized theoretical function or affect-curve to describe the course of affective responses on a pleasantness–unpleasantness dimension as a function of stimulus intensity

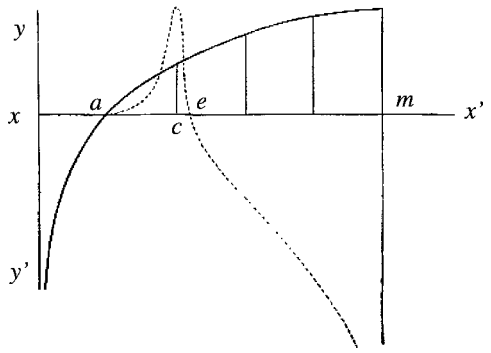


FIG. 1. (Figure 242, p. 323 in Wundt, 1910). The continuous curve shows the logarithmic growth of sensation with intensity (Fechner's law). The ordinate y-y' represents the intensity of sensation, the abscissa x-x' the stimulus level. Point *a* is the absolute threshold. Point *m* is the maximum stimulus height. The broken curve is the affective response curve (*Gefüleskurve*) as a function of stimulus intensity. The ordinate y-y' in this case represents the strength of the affective or hedonic response on a pleasantness–unpleasantness (*Lust–Unlust*) dimension. Point *c* is the value of stimulus level which gives rise to a maximal pleasure response. Point *e* is the value of stimulus level at which the affect response changes from positive to negative, the ''nullpoint'' (*Nullpunkt*).

above the absolute threshold (Fig. 1) (see also Vitz, 1973). The affect-curve has a kind of asymmetrical bell shape, initially growing slowly with stimulus intensity, reaching a maximal positive value and then inverting, subsequently passing into the negative (unpleasantness) above a certain point of indifference. Instances of pleasantness/unpleasantness ratings as a function of stimulus intensity conforming to the Wundt affect-curve can in fact be found, most notably for gustatory or olfactory stimuli (Pfaffmann, 1960).

In the case of specifically auditory stimuli several attempts have been made to characterize affective responses as a function of intensity and frequency, although the results depend to a certain extent on the method used. Paired comparison of tones at the same intensity has yielded the result that frequencies between about 200 and 1000 Hz are either least annoying (Laird and Cove, 1929) or most pleasant (Vitz, 1973). Paired comparison of tones of different intensities at the same frequency show that subjects prefer tones of 50 dB in the range 40–90 dB with the preference steadily decreasing with intensity. Below about 60 dB low frequencies are rated higher than high frequencies, and above 60 dB there is a crossover in the relative preference of high and low frequencies.

The method of adjustment (Pollack, 1952) yields the result that the most comfortable listening levels are about 40 dB above threshold (60 dB at 1 kHz), and generally follow equal loudness contours, though with differences consistent with the result from the paired comparison data that high frequencies are less preferred than low frequencies. The upper limit of most comfortable levels is some 20 dB higher, but less than the threshold of discomfort (Silverman, 1947). A third method using pleasure rating scales (Singer and Young, 1941a, b) indicates a similar decay function of rating with intensity, but does not show a peak in rating as a function of frequency between 200 and 1000 Hz. This data was later used to generate pleasantness scales (Guilford, 1954) as

a function of frequency and intensity, but has been criticized for producing distorted results (Vitz, 1973). On the other hand, the paired comparison method does not allow a zero point, and hence the distinction between pleasant and unpleasant.

Returning to the present study, however, we should note that in all the previous studies all intensities were below the saccular threshold of 90 dB. The experiment we report here, then, investigates the relationship between the pleasure ratings of pulse trains, on the Singer and Young scale, and the amplitude of MVEP as a function of intensity up to 115 dB. Two ''carrier'' frequency conditions were used: a 200-Hz experimental condition, which is near the peak of saccular acoustic sensitivity, and a 4-kHz control condition, which is beyond the region of saccular acoustic sensitivity. We predict that above the threshold of saccular acoustic sensitivity the pleasure rating of the experimental condition should depart significantly in a positive direction from the control condition. This should manifest itself as a significant ANOVA interaction between carrier and intensity within-subject factors.

There were several reasons for choosing pulse-trains in preference to isolated pulses. First, it is desirable to use extended stimuli which can give rise to a sustained response in the saccular nerve, since normal vestibular responses are sustained over a certain duration. Through extensive piloting work it has been found that a sustained pulse-following MVEP response can be obtained at rates up to at least 40 Hz without significant adaptation over five repetitions (see Fig. 2). Both the spontaneous rates and the normal operating discharge rates of otolithic units range between about 20 and 100 spikes/s (Fernandez and Goldberg, 1976). Further, the resting discharge rate of ARID fibers in the cat is about 22 spikes/s and stimulation by tones may produce a mean response rate of about 40 spikes/s at 100 dB, rising to a mean of about 130 spikes/s at 110 dB (McCue and Guinan, 1994). Thus, the demonstration of a sustained pulse-following MVEP response in this range indicates a pulse-locked response across units in the saccular nerve in their normal range. Although it is almost certain that continuous tones produce phase-locking in the saccular nerve in humans, it is difficult to observe this by means of the p13-n23 wave. In contrast the p13-n23 wave can be easily resolved and quantified for pulse trains at 40 Hz.

A second reason for believing pulse trains of this kind to be valid stimuli is that the vocalizations of lower vertebrates, including both fish and amphibians, are often in the form of pulse trains (Pitcher, 1986), suggesting that these are effective stimuli for saccularly mediated communication (see later discussion and Fig. 7). Additionally, we have shown that pulse trains of this kind are subjectively preferred to isolated pulses.

An issue of further interest is raised by our previous work (Todd and Cody, 2000) in which we found evidence of a relationship between attitudes to normal vestibular sensations, such as may be obtained from swings, rocking-chairs, etc., and loud music. We therefore also administered a simple questionnaire to obtain between-subjects information concerning subject attitudes to such sensations. We predict that the more positive a subject attitude to sensations of self-motion, the greater the positive departure of the experimental condition from the control.

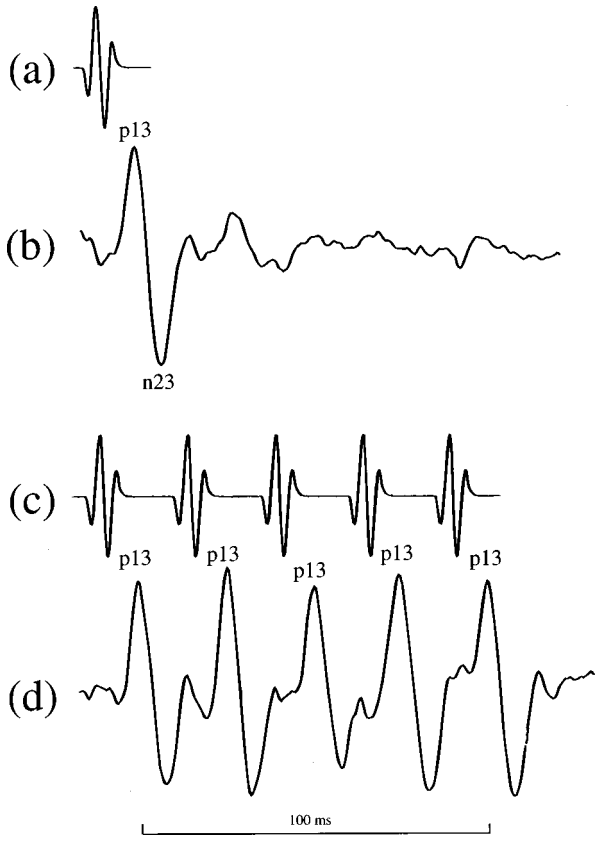


FIG. 2. Myogenic vestibular evoked potentials (MVEP). (a) A 10-ms 200-Hz tone pulse with rounded cosine envelope. All stimuli presented monaurally. (b) Averaged unrectified EMG response recorded from the right sternocleidomastoid muscle, ipsilateral to the stimulation, of one subject in response to a 200-Hz, 110-dB tone pulse. Note the presence of a diphasic wave at approximately 13 and 23 ms. The sampled, unrectified EMG was averaged over a run of 256 multiple repetitions of the stimuli. (c) A 100-ms train of 10-ms 200-Hz tone pulses. The tone repetition frequency was 40 Hz. (d) MVEP recorded from the right sternocleidomastoid muscle of the same subject in response to a 110-dB, 100-ms train of 10-ms 200-Hz tone pulses. Note that even at a 40-Hz repetition rate the p13-n23 wave is well-defined, indicative of a sustained following response in the saccular nerve.

II. METHOD

A. Subjects

Ten Manchester University undergraduate and postgraduate students, five males and five females were used as subjects. Subjects were first screened for any neurological impairments, vestibular impairments, and chronic medication. All subjects had pure tone thresholds in the normal range. Nine of the ten subjects were right handed. (The effect of handedness was not further investigated in this study.)

B. Apparatus

Stimuli were generated by CED Signal software and delivered via a pair of Sennheiser HD420 headphones. Stimulus intensity was calibrated by means of a B&K 2260 in conjunction with a B&K 4163 artificial ear and 4134 pressure microphone. Differential recordings of EMG were made

using two small, surface electrodes. A CED 1902 was used for EMG amplification and analog filtering. EMG was sampled by means of a CED 1401plus. Visual feedback for maintenance of steady background activity was provided by means of an oscilloscope. Hedonic responses were recorded by means of a nine-point scale (Appendix A). Between-subject factors were assessed by means of a simple questionnaire.

C. Stimuli

The stimuli were 100-ms trains of 10-ms tone pulses with rounded cosine envelope. The effective rise-fall times were 5 ms. The pulse repetition rate of the stimuli was 40 Hz while the train repetition rate was 5 Hz. ''Carrier'' frequencies were 200 and 4000 Hz with a sampling rate of 10 kHz. The duration and envelope of individual pulses were chosen so as to be sufficiently localized in time in order to resolve the p13-n23 wave, but at the same time to be also sufficiently localized in the frequency domain to obtain adequate frequency resolution.

Stimuli were presented at 13 intensity levels from 55 to 115 dB in 5-dB steps. These levels were chosen so as to provide a useful range, but to remain below the limits set by the *Noise at Work Regulations (1989)* issued by the *Health and Safety Executive* in order to avoid the possibility of permanent damage. These limits are a daily personal noise exposure of 90 dB(A) and a peak sound level (PSL) of 140 dB *re* 20 μPa. For this purpose two other measures were monitored, in addition to SPL, the equivalent continuous level L_{Aeq} and the PSL.

D. Procedure

Before starting the experiment pure tone audiograms were obtained and a handedness questionnaire was also administered for each subject. Subjects were informed (*i*) the purpose of the experiment, (*ii*) that it would involve EMG recording from muscles of the neck, (*iii*) that they would be required to maintain the neck muscles in a state of tension during a recording which may involve some discomfort, and (*iv*) that they would be stimulated by repetitive loud sounds. Subjects were then asked to read and sign a consent form in order to conform with local ethical committee guidelines. The local ethical guidelines conform with the principles of the British Psychological Society.

After subjects had understood and signed the consent form, two small, surface electrodes were placed over the left SCM muscle (and over the right SCM for a second recording session). The positive electrode was placed on the belly of the muscle, the negative electrode just caudal to the positive, while the common electrode was placed on the sternum. With this differential or bipolar recording arrangement both electrodes are active relative to the common electrode (Todd *et al.*, 2001).

After amplification and low-pass filtering at 1 kHz by means of the CED 1902, the EMG was sampled at 2.5 kHz by means of the CED 1401plus. Rectified EMG was used to provide feedback to enable the subject to maintain a suitable level of background muscular contraction, fixed at 200 μV, by voluntarily turning the head to one side, contralateral to the stimulation. The sampled, unrectified EMG was averaged over a run of 256 multiple repetitions of the stimuli. Before the averaging run, subjects were given an opportunity to find a sitting position which would minimize discomfort.

Once the subject had found a suitable position, the experimental procedure was started. For each of the carrier conditions pulse trains were presented at 13 intensity levels from 55 to 115 dB(A) in 5-dB steps. At each intensity level, the stimuli (pulse trains) were repeated 256 times at a 5-Hz stimulus repetition rate. At the end of the 256 repetitions of the stimulus, subjects were asked to rate quality of the sound according to whether it was pleasant or unpleasant on a scale from -4 (extremely unpleasant) to $+4$ (extremely pleasant), where 0 corresponds to an indifferent attitude (see Appendix A). Although it is quite possible that there was an order effect by this procedure, this is kept constant across all subjects. The gradual ordered increment also ensures that the correct ordinal relation is maintained between adjacent responses, and further avoids any possible startle response.

From 90 dB and upwards, in addition to the subjective rating, MVEP was obtained ipsilateral to stimulation. Between each recording run subjects were allowed to relax. Due to the length of the procedure and in order to ensure the sound exposure was kept below the set limits (see above), the procedure was repeated for the right SCM on a different day. In order to be sure that no permanent damage was being done by this procedure, one subject was monitored extensively before and after measurements. There was no evidence of any significant change in pure tone thresholds for this subject.

III. RESULTS

In order to illustrate the nature of the stimuli and physiological response we show in Fig. 2 MVEP obtained from a single 200-Hz tone pulse and a train of 200-Hz tone pulses with a 40-Hz pulse repetition rate. These responses are ipsilaterally dominant, consistent with Colebatch *et al.* (1994).

Figure 3(a) shows the results of the MVEP recording for the main experiment. The calculated values are the mean p13-n23 amplitude over each of the five repetitions within the pulse train. Since there was no significant difference between left and right ears over the ten subjects only the mean value over both ears is shown (numerical values including standard deviations are shown in Table BI, Appendix B). Note that values are only shown for the 200-Hz experimental condition since no MVEP was apparent for the 4.0-kHz control condition. Extrapolation indicates a mean threshold over all ten subjects of about 95 dB(A), but a threshold of 85–90 dB(A) would be more consistent with previous work (McCue and Guinan, 1994, 1995). The most likely reason for this discrepancy is that the measured potentials involve the attenuation of at least two synapses.

Figure 3(b) shows the pleasure rating as a function of intensity and carrier frequency (numerical values are shown in Table BI, Appendix B). For intensities less than about 75 dB(A) the mean response across all subjects shows no difference between carrier conditions. Between 75 and 85 dB(A) a transitional region indicates that on average the re-

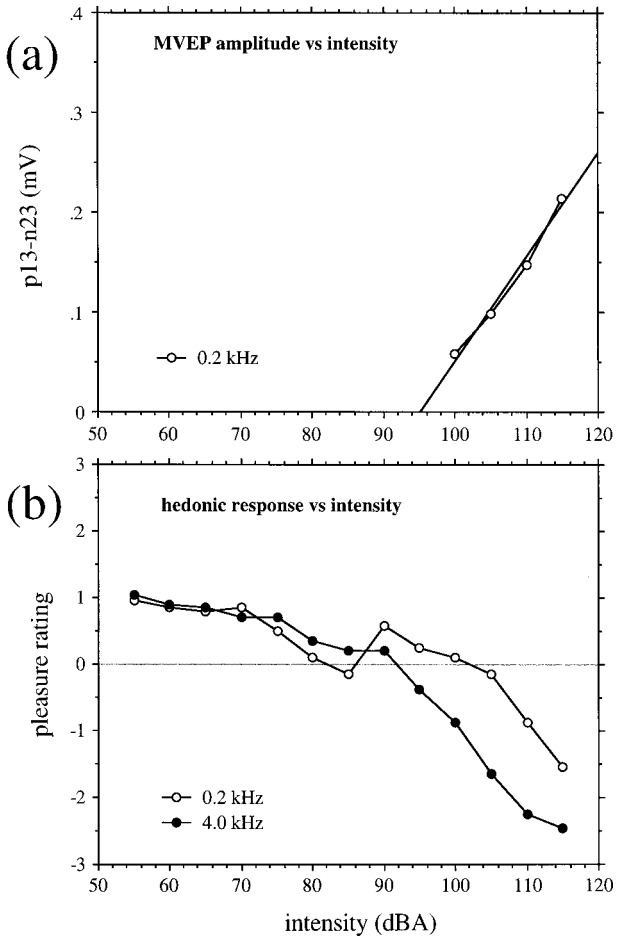


FIG. 3. MVEP versus subjective ratings of pulse trains. (a) MVEP amplitude in mV as a function of intensity. MVEP values were averaged over the five repetitions. Consistent responses in all subjects were only obtainable above 95 dB, hence the average MVEP values are only shown for 100 dB and above. Linear regression of as a function of sound pressure level gives a good approximation ($R^2=0.987$, $p<0.01$, $n=4$). The regression equation is $V(I)=aI+b$ where $V(I)$ is the p13-n23 amplitude in μV, I is the intensity in dB(A), and the parameters are $a=10.4$ and $b=-986$. (b) Hedonic rating as a function of intensity and carrier frequency ($p<0.05$).

sponse in the 200-Hz condition becomes very slightly negative. Above 85 dB(A), however, there is a distinct positive departure in the pleasure rating for the 200-Hz condition relative to the 4-kHz condition, which continues to become more negative as intensity increases. ANOVA gives an interaction F-ratio between carrier and intensity of $F(12,96)=3.97$ (Greenhouse-Geisser $\varepsilon=0.162$, $p<0.05$, corrected for violations of sphericity). Thus the elevation in subjective rating across all subjects for the 200-Hz condition appears to coincide closely with the mean threshold across all subjects for acoustic sensitivity of the sacculus.

While the overall interaction between carrier and intensity is statistically significant, the size of the positive departure from monotonicity is not large when considering the average over all ten subjects. We therefore consider the effect of between-subjects factors. Subjects were asked two questions (1) ''Do you like the sensation of swings?,'' yes, neutral, or no, and (2) ''Do you like the sensation of loud music?,'' yes, neutral, or no. Visual inspection of the responses (see Table I) suggests that there is a strong association between the two factors. If the no responses are combined with the neutral response, then a two-by-two contingency table shows that eight of the ten subjects agree across the two questions. However, the Fisher exact p-value$=0.19$. We cannot therefore conclude that the association is statistically significant from this sample. Nevertheless, a single between-subjects factor was created whereby subjects were labeled as S+ if they answered yes to both questions, else they were labeled as S−, which produced a balanced five subjects in each category (see Table I).

TABLE I. Attitude to swings and loud music. Columns 2 and 3: 1=yes, 0=neutral, −1=no. Column 4: S+ is 1, if columns 2 and 3 are both yes, else it is 0.

Subject	Like swings?	Like loud music?	S+
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	−1	0	0
6	1	1	1
7	0	1	0
8	−1	0	0
9	0	0	0
10	1	0	0

Figure 4(a) shows MVEP amplitude as a function of intensity, split by subject type (numerical values are shown in Table BII, Appendix B). Mean values of the p13-n13 amplitude indicate very little to distinguish subjects from a physiological viewpoint. However, when split into S+ and S− groups the subjective responses [Fig. 4(b) and Table BII, Appendix B] show a rather dramatic difference for the 200-Hz experimental condition. Above 85 dB(A), close to the saccular threshold, the S+ group show a pronounced positive increase in subjective rating, reaching a peak at 100 dB(A) and only becoming negative above 110 dB(A); whereas S− group showed an almost continual decrease in subjective rating with increasing intensity. In contrast there was no significant difference between the two subject types in the control condition. ANOVA gives an F-ratio for the interaction between subject type and carrier of $F(1,8)=12.4$, $p<0.01$. In order to demonstrate that this effect is not just a consequence of the way the between-subject factor was defined, an ANOVA run with the ''like-swings'' and ''like-loud'' factors independently also produces significant interactions, respectively, $F(1,8)=8.1$, $p<0.05$ and $F(1,8)=6.8$, $p<0.05$.

The interaction between subject type and carrier is illustrated in Fig. 5(b) which shows the difference between the experimental and control conditions for each of the subject types. For comparison the mean response value for the 4-kHz control condition is shown in the same figure [Fig. 5(a)]. When viewed from this perspective the S+ response looks remarkably like Wundt's idealized affect-curve.

IV. DISCUSSION

The data presented above, then, would seem to support the hypothesis that there is a change in the quality of sensation at or above the physiological threshold for saccular

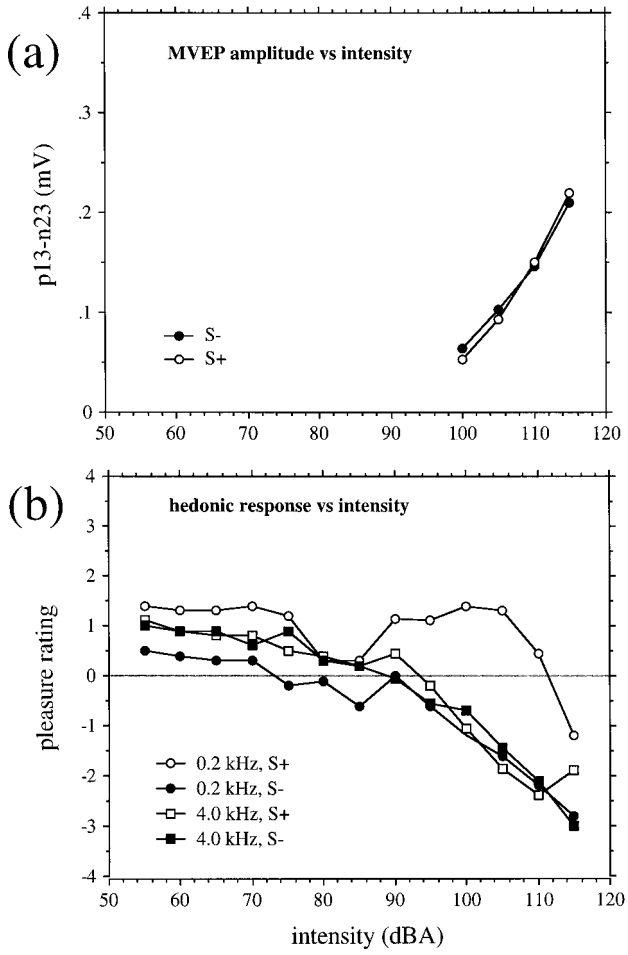


FIG. 4. MVEP versus subjective ratings of pulse trains, effect of subject type. (a) MVEP amplitude as a function of intensity in mV, split by subject type. Mean values of the p13-n13 amplitude indicates very little to distinguish subjects from a physiological viewpoint. (b) Hedonic rating as a function of intensity and carrier frequency, split by subject type ($p<0.01$).

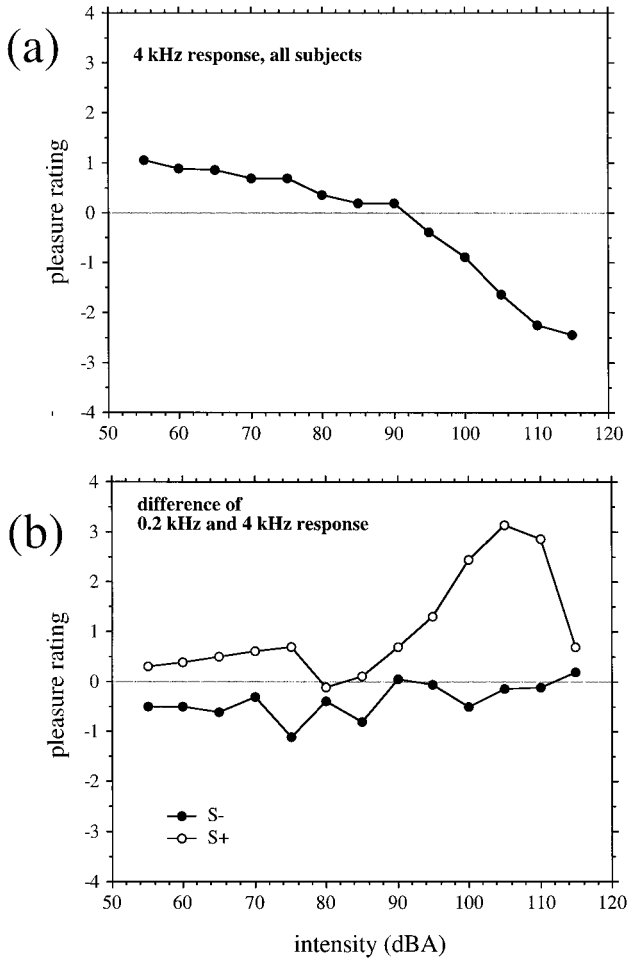


FIG. 5. Subjective ratings of pulse trains. (a) Mean hedonic rating across all subjects for 4.0-kHz condition. (b) Interaction between carrier and subject type, expressed as the difference between the 0.2- and 4.0-kHz response for each subject type ($p<0.01$).

acoustic sensitivity, and further the data indicate that the positivity of the change depends on the subject attitude to vestibular sensations. However, we need to consider alternative explanations for the above effect. One explanation might be that the onset of unpleasantness occurred later in the 200-Hz condition due to the outer-middle ear transfer function. However, all intensity measures used an ''A-weighting'' which approximates the transfer function. Another explanation of the result could be that since the physiological procedure was only started at 90 dB it might be the case that subjects attention was distracted from the unpleasantness of the acoustic sensation by the discomfort of the head position during EMG recording. However, there is little or no sign of any such effect in the 4-kHz control condition, thus even if there was such an effect it is only quite small and constant. A third argument against the case for a saccular contribution to the quality of sensation is that, on average, the pleasure ratings at 90–105 dB(A) are not significantly greater than at 55–75 dB(A), thus it might be that the apparent dip at 80–85 dB(A) in Fig. 4(b) is spurious. However, this overlooks the degree of departure from the control condition above the saccular threshold for the S+ group as shown in Fig. 5(b).

A. A two-component model of acoustic affect

We are left with the conclusion, then, that the coincidence of the departure from monotonicity in the pleasure response with the physiological threshold for saccular acoustic sensitivity means that a causal link between the two cannot be ruled out. Is it possible to quantify this link? One way in which this might be done is to assume that any affective or hedonic response $H(I)$ is the result of the linear superposition of two components which both conform to Wundt's affect-curve, i.e.,

$$H(I)=H_C(I)+H_S(I), \tag{1}$$

where $H_C(I)$ is a ''cochlear'' component, represented by the mean value of the control condition across all subjects, and $H_S(I)$ is a ''saccular'' component represented by the departure from the control [Fig. 6(a)]. The application of parsimony would suggest ideally a single generalized function for both components. However, in practice hedonic psychometric functions do not conform to a single mathematical description (Pfaffmann, 1960). Given the consistency in judgments of all subjects on the control condition and the relative smoothness as a function of intensity, $H_C(I)$ may be ap-

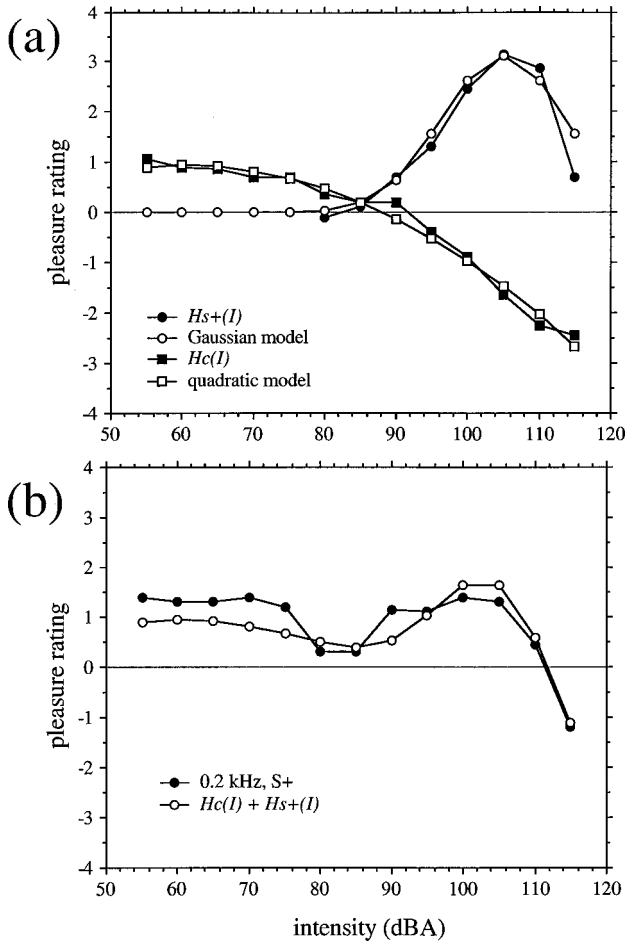


FIG. 6. Model function fits of subjective ratings. (a) 4.0-kHz response, $H_c(I)$, fitted with a quadratic model (peak response at 60 dBA) ($R^2=0.98$, $p<0.0001$, $n=13$). The 0.2-kHz difference response for S+ subjects, $H_{s+}(I)$, fitted with Gaussian model (peak response at 105 dBA) ($R^2=0.94$, $p<0.0001$, $n=8$). (b) Simulation of the 0.2-kHz response for S+ subjects by summation of quadratic and Gaussian models, $H_c(I)+H_{s+}(I)$.

proximated accurately by a quadratic function [see Fig. 6(a)], i.e.,

$$H_C(I)=a+bI+cI^2, \tag{2}$$

where I is again intensity in dB(A). A least squares fit gives parameter estimates of $a=-3.36$, $b=0.143$, $c=-0.00119$. According to this interpretation $H_C(I)$ has a maximum of about one (on the Singer–Young scale) at about 60 dB, close to the previously identified values for maximal preference (Vitz, 1973) and passes into the negative at about 90 dB(A), near to the classical threshold of discomfort. The saccular component for the S+ subjects, $H_{S+}(I)$ (ignoring for the moment values below the saccular acoustic threshold), may be approximated by a Gaussian function [see Fig. 6(a)], i.e.,

$$H_{S+}(I)=a\exp\left[-\left(\frac{I-b}{c}\right)^2\right]. \tag{3}$$

Fitting by eye gives approximate parameter estimates of $a=3$, $b=105$, $c=12$, which may be interpreted as $H_{S+}(I)$ having a maximum of about three on the Singer–Young scale at 105 dB(A). Linear extrapolation of the high-intensity part of the curve suggests that it would pass into the negative at about 120 dB(A).

Simple addition of (2) and (3), shown separately in Fig. 6(a), produces a simulation as shown in Fig. 6(b). As a first approximation, then, the simple two-component model gives at least a fair qualitative description of the data for the S+ subjects but with two notable departures. Below 80 dB(A) the measured response to the 200-Hz condition is slightly more positive than the control condition for the S+ subjects, and conversely slightly less positive than the control condition for the S− subjects. The departure at low intensities for the S+ subjects is, however, consistent with the previous work that indicates that low frequencies produce higher ratings than high frequencies at low intensities (Vitz, 1973).

B. Implications for human responses to loud sounds

If we accept the interpretation of the above data, then, the difference in the hedonic gain of the two systems, a value of one for the cochlear system and three for the saccular system, would explain the compulsion to exposure to loud music (Florentine *et al.*, 1999; Todd and Cody, 2000). Even if we correct for a possible difference in the hedonic response as a function of frequency in the cochlear system (Guilford, 1954), the saccular hedonic gain remains larger, and hence intrinsically more highly motivating. The estimated intensity of 105 dB (A) (fast time-weighting) for the maximal affective response corresponds to about 95 dB L_{Aeq} and is therefore also close to the value of 96 dB L_{Aeq} suggested by Dibble (1995) for the minimum level necessary for rock music to work.

A common experience of people who regularly attend loud music venues is temporary threshold shift (TTS) (Dibble, 1995; Moore, 1989). One might speculate that, perversely, TTS is almost a desirable state to achieve in these circumstances since, by reducing the input gain of the cochlea, the balance between any pleasant saccular sensation and the unpleasant cochlear sensation will be swung in favor of the former. This will be enhanced by the fact that vestibular acoustic adaptation is much less than cochlear adaptation.

Several questions remain to be answered. What possible central mechanism might link saccular acoustic sensitivity to a positive change in quality of sensation? Why should some individuals apparently find such sensations rewarding and others aversive? With respect to the first question, it is generally established that all senses have projections via the mesencephalic reticular formation to the limbic system (Kraus and McGee, 1992), which is thought to mediate affective responses (Ervin and Martin, 1986). However, the vestibular system in particular has well-documented specific projections to the autonomic nervous system (Yates, 1996), including higher centers up to the hypothalamus (Katafuchi *et al.*, 1987). It is therefore very likely that these central vestibular projections include the mesolimbic dopamine system (Rolls, 1975), the neural substrate of reward and addictive behavior, which would account for why the saccular system apparently has a greater hedonic gain than the cochlear system.

As for the second question, there has been considerable recent research which points to a neurobiological basis of personality (Eysenk, 1990). In particular modern trait theories of personality include a dimension of extraversion or positive emotionality (PE) which is based on sensitivity to signals of incentive–reward. Several studies now indicate that PE trait levels are dependent on brain dopamine activity (Depue and Collins, 1999). Thus the difference between S+ and S− subjects may simply be one of inherited individual differences in brain dopamine functioning, which would manifest as individual difference in levels of the extraversion trait.

Alternatively the characterisation of subject differences may require more than a single dimension. For example, according to Pickering and Gray (1999) it is the personality trait of impulsive sensation seeking (ISS) which is the more fundamental from a neurobiological perspective, i.e. it is the ISS trait which corresponds to the functioning of the mesolimbic dompamine system, which Gray refers to as the ''behavioural activation system'' (BAS). Gray has suggested that the ISS trait lies in Extraversion/Neuroticism plane of Eysenk's system such that subjects who score high on ISS are neurotic extraverts. Such subjects show strong responses to stimuli which activate the BAS or reward system. In contrast subjects who show strong responses to aversive stimuli are neurotic intraverts in the Eysenk system. Such responses are mediated by the ''behavioral inhibition system'' (BIS), another basic brain system which Gray (1987) suggests corresponds to the septo-hippocampal system. Thus, given that above the saccular threshold subjects may be receiving both pleasant or rewarding saccular sensations and unpleasant or aversive cochlear sensations, the differences between the *S*+ and *S*– subjects may lie in a more complex interaction between the BAS and the BIS.

We may speculate further that, given the powerful role that the mesolimbic structures play in learning, it could be that all subjects learn which are the kinds of sounds (e.g., pulsed low frequencies) which are likely to produce saccular sensations. Thus even at low intensities, below the saccular acoustic threshold, such sounds may give rise to affective responses by association. If it were the case that the S− subjects had learned a negative association with such saccular sensations, then by negative conditioning they would develop an aversive response to the corresponding cochlear sensations, even at low intensity, and hence the above result for S− subjects. Clearly, however, further data are required to resolve such speculations.

C. On the evolutionary significance of saccular acoustic sensitivity

One final question remains, the answer to which is necessary for a complete account of these phenomena. What is the evolutionary significance of saccular acoustic sensitivity? As suggested earlier, it is quite possible that there is no evolutionary significance, and that this phenonemon is just a byproduct of normal vestibular function. However, such an account is unsatisfactory for several reasons, not least that it does not explain the high hedonic gain of either intertial or acoustic vestibular stimulation. An alternative theory is that

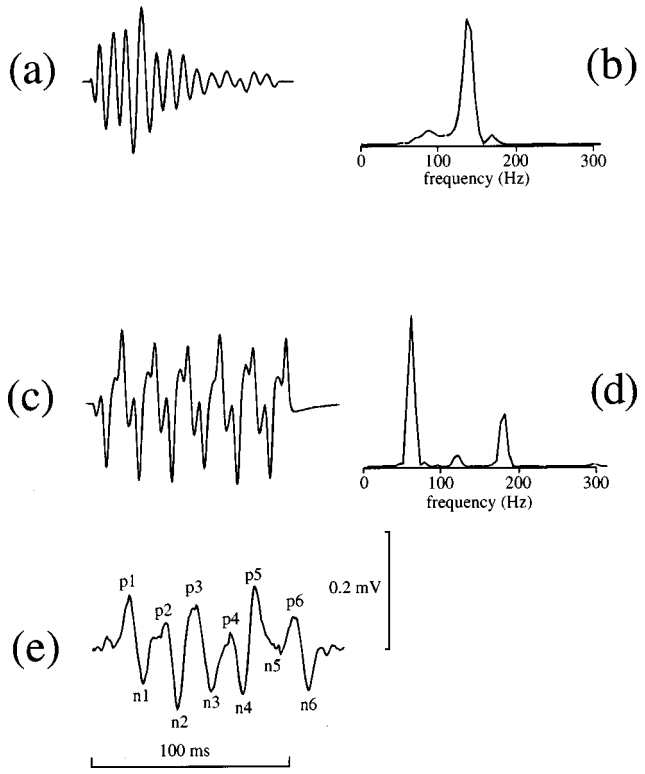


FIG. 7. Haddock (*Melanogrammus aeglefinus*) courtship vocalization. (Recorded by A. D. Hawkins, made available by the National Sound Archive at the British Library.) (a) The waveform of a single pulse sampled from a repetitive sequence of pulses made during a vocalization of a male haddock. The pulses are produced by ''drumming'' or ''knocking'' on the swimbladder, hence the above waveform represents its impulse response. (b) A spectral analysis of the waveform shows the power is concentrated around 140 Hz. (c) The waveform of a 100-ms sample of repetitive ''knocking'' produced by a male haddock. During courtship the repetition frequency of the knocking increases until it becomes more like a continuous ''hum''. Sound production ceases when eggs and sperm are released (Pitcher, 1986). (d) A spectral analysis of this waveform shows there to be predominant power at 60 and 180 Hz where the 60-Hz component corresponds to the pulse repetition frequency. This signal therefore has some similarity with the pulse-train stimuli in Fig. 2, except that the repetition frequency and ''carrier'' frequency at 40 and 200 Hz are more separated, by a factor of 5 rather than 3. (e) MVEP recorded from the right sternocleidomastoid muscle of one subject in response to the 100-ms haddock vocalization at 115 dB. In spite of the reduced temporal resolution, there are six well-defined positive–negative waves corresponding with the six cycles of the stimulus.

the sacculo-mesolimbic pathway discussed above is the product of the evolutionary conservation of a primitive central acoustic pathway, in addition to the conservation of a peripheral sensitivity. In fact, there is considerable evidence to support this theory.

In teleost fish, vocalizations are produced by ''drumming'' on the swimbladder. A classic example is the mating call of the haddock (*Melanogrammus aeglefinus*) [Figs. 7(a) and (c)] (Pitcher, 1986). Such sounds have much in common with human drumming, i.e., they consist of repetitive, low-frequency impulses, and indeed can be shown to be as effective in producing a response in the human sacculus as the stimuli used in the experiment above [Fig. 7(e)]. In the case of amphibians, many anuran (frogs and toads) mating calls also consist of pulse trains. Such calls usually require both a low-frequency and high-frequency component to be effective (Lewis and Lombard, 1988). Thus the sacculus is likely to

play a role in processing the low-frequency component, particularly if seismically mediated (Lewis and Narins, 1985). For example, the bullfrog (*Rana catesbeiana*) call has components at 200 Hz and 1.5 KHz and the acoustic/seismic sensitivity of *R. catesbeiana* sacculus is well documented (Lewis and Lombard, 1988).

The amphibian system is of interest here since it represents the transition from ''old'' to ''new'' systems. The anuran auditory pathway in particular is relevant since, of the three amphibian orders (caecilians, salamanders and anurans), they are the most vocal. While evolving new acoustic sensors which respond to higher frequencies (the basillar papilla for high frequencies, the amphibian papilla for middle and low frequencies), the anuran system also clearly conserved an acoustic function for the sacculus (Lewis and Lombard, 1988). Both sacculus and papillae send projections to the midbrain and forebrain via brainstem nuclei, the saccular projection representing the primitive central system inherited from fish. Of particular interest are several forebrain projections, including to the striatum, preoptic area, and hypothalamus, which are thought to be the neural substrate for reproductive vocal behavior (Neary, 1988). A medullary structure, the nucleus isthmi, is also implicated in both the anuran and otophysine acoustic hypothalamic pathways (McCormick, 1988, 1992). These structures correspond to structures within the mesolimbic dopamine system in higher vetebrates and thus provide an evolutionary precursor for a sacculomesolimbic pathway in higher vertebrates.

To summarize then, the evolutionary significance of saccular acoustic sensitivity in higher vertebrates, according this theory, is that a primitive central mechanism has also been conserved, through which vocally mediated sexual selection has continued to operate (Darwin, 1871; Miller, 2000) by providing a direct pathway to reward centers in the brain (which is not provided by cochlear pathways). The existence of loud vocalizations in primates is almost universal. Not least of these vocalizations are ''duetting'' in pair-bonded Gibbons (*Hylobates spp.*) and the Chimpanzee (*Pan troglodytes*) ''pant-hoot'' (Geissmann, 2000), both of which are associated with sexual behavior and are accompanied by movements, and are of sufficient intensity and in the appropriate frequency range to evoke an acoustic saccular response. Such loud synchronized vocalizations in primates are considered to be the precursor of loud music in human culture (Merker, 2000). While amplified music is clearly a modern invention, both human vocalizations and percussive instruments are of sufficient acoustic power to activate the sacculus (Todd *et al.*, 2000). Thus saccular acoustic sensitivity may still play an important role in perceptual and behavioral responses to loud music.

V. SUMMARY AND CONCLUSION

In this article we have provided data which indicates a link between saccular acoustic sensitivity and behavior. The data can be explained if it is assumed that cochlear and saccular affective response are additive and both conform approximately with Wundt's affect-curve, the cochlear affective response peaking at about 60 dB and the saccular affective response peaking at about 105 dB. The hedonic gain of the saccular system appears to be higher than the cochlear system, which would explain the compulsion to exposure to loud sounds, even at the expense of TTS in the cochlear system. Several questions remain to be answered, though, which at this stage must remain a matter of some speculation. However, the proposed sacculo-mesolimbic theory explains (a) the high hedonic gain of the saccular system, (b) individual differences in hedonic sensitivity, via inherited differences in dopamine activity, and (c) the evolutionary significance of the conservation of saccular acoustic sensitivity in higher vertebrates. Clearly, further investigations in these matters are warranted.

ACKNOWLEDGMENTS

Research was supported by a grant from the Royal Society. I am grateful to Fred Cody and Jon Banks for help in the electrophysiological aspect of this work, to Chris Lee for comments on earlier drafts of this article, and to anonymous reviewers for many useful suggestions for improvement of the submitted manuscript. I am grateful also to the National Sound Archive for permission to publish analysis of the haddock recordings.

APPENDIX A: SUBJECT RESPONSE SHEET

Rate each sound according to whether the sensation of the sound is pleasant or unpleasant on the following scale. In making your judgment focus on the basic feeling of the sensation rather than any perceived patterns or associations of the sound. Also try to dissociate the pleasantness/unpleasantness from any discomfort due to the tension in your neck.

4 extremely pleasant
3 very pleasant
2 quite pleasant
1 slightly pleasant
0 neutral
−1 slightly unpleasant
−2 quite unpleasant
−3 very unpleasant
−4 extremely unpleasant

	unpleasant						pleasant		
1.	−4	−3	−2	−1	0	1	2	3	4
2.	−4	−3	−2	−1	0	1	2	3	4
3.	−4	−3	−2	−1	0	1	2	3	4
4.	−4	−3	−2	−1	0	1	2	3	4
5.	−4	−3	−2	−1	0	1	2	3	4
6.	−4	−3	−2	−1	0	1	2	3	4
7.	−4	−3	−2	−1	0	1	2	3	4
8.	−4	−3	−2	−1	0	1	2	3	4
9.	−4	−3	−2	−1	0	1	2	3	4
10.	−4	−3	−2	−1	0	1	2	3	4
11.	−4	−3	−2	−1	0	1	2	3	4
12.	−4	−3	−2	−1	0	1	2	3	4
13.	−4	−3	−2	−1	0	1	2	3	4

APPENDIX B: NUMERICAL VALUES OF RESULTS

TABLE BI. MVEP (V) vs hedonic ratings (H) of pulse trains.

	MVEP (μV)		0.2 kHz		4.0 kHz	
dBa	$V(I)$	s.d.	$H(I)$	s.d.	$H(I)$	s.d.
55	⋯	⋯	1.0	1.2	1.1	1.2
60	⋯	⋯	0.9	1.4	0.9	1.4
65	⋯	⋯	0.8	1.4	0.9	1.4
70	⋯	⋯	0.9	1.6	0.7	1.2
75	⋯	⋯	0.5	1.5	0.7	1.0
80	⋯	⋯	0.1	1.4	0.4	1.3
85	⋯	⋯	−0.2	1.5	0.2	1.5
90	⋯	⋯	0.6	1.3	0.2	1.4
95	⋯	⋯	0.3	1.6	−0.4	1.5
100	58	31	0.1	1.7	−0.9	1.2
105	98	43	−0.2	1.9	−1.7	1.2
110	148	65	−0.9	1.9	−2.3	0.9
115	214	93	−1.6	2.3	−2.5	1.5

TABLE BII. MVEP vs hedonic ratings of pulse trains, split by subject type.

	MVEP (μV)				0.2 kHz				4.0 kHz			
dBA	$V_+(I)$	s.d.	$V_-(I)$	s.d.	$H_+(I)$	s.d.	$H_-(I)$	s.d.	$H_+(I)$	s.d.	$H_-(I)$	s.d.
55	⋯	⋯	⋯	⋯	1.4	1.0	0.5	1.2	1.1	0.9	1.0	1.2
60	⋯	⋯	⋯	⋯	1.3	1.2	0.4	1.5	0.9	1.0	0.9	1.3
65	⋯	⋯	⋯	⋯	1.3	1.4	0.3	1.4	0.8	0.9	0.9	1.4
70	⋯	⋯	⋯	⋯	1.4	1.2	0.3	1.5	0.8	0.8	0.6	1.1
75	⋯	⋯	⋯	⋯	1.2	1.4	−0.2	1.4	0.5	0.9	0.9	0.9
80	⋯	⋯	⋯	⋯	0.3	1.3	−0.1	1.7	0.4	1.3	0.3	1.0
85	⋯	⋯	⋯	⋯	0.3	1.0	−0.6	1.8	0.2	1.5	0.2	1.3
90	⋯	⋯	⋯	⋯	1.2	0.8	0.0	1.6	0.5	1.7	−0.1	1.2
95	⋯	⋯	⋯	⋯	1.1	1.0	−0.6	1.7	−0.2	1.7	−0.6	1.2
100	53	35	63	27	1.4	0.6	−1.2	1.3	−1.1	1.4	−0.7	0.9
105	93	51	104	36	1.3	1.3	−1.6	1.1	−1.1	1.3	−1.5	1.1
110	150	82	146	47	0.5	1.5	−2.2	1.3	−2.4	1.0	−2.1	0.9
115	219	114	202	72	−1.2	2.3	−2.8	1.5	−1.9	1.0	−3.0	1.2

Bickford, R., Jacobson, J., and Cody, D. (**1964**). ''Nature of averaged evoked potentials to sound and other stimuli in man,'' Ann. N.Y. Acad. Sci. **112**, 204–223.

Burian *et al.* (**1989**). ''Saccular afferent fibres to the cochlear nucleus in the guinea pig,'' Arch. Otolaryngol. **246**, 238–241.

Cazals, Y., Aran, J., and Erre, J. (**1982**). ''Frequency selectivity and selectivity of acoustically evoked potentials after complete cochlear hair-cell destruction,'' Brain Res. **231**, 197–203.

Cazals, Y., Aran, J., and Erre, J. (**1983a**). ''Intensity difference thresholds assessed with eighth nerve and auditory cortex potentials: Compared values from cochlear and saccular responses,'' Hear. Res. **10**, 263–268.

Cazals, Y., Aran, J., Erre, J., Guihaume, A., and Aurousseau, C. (**1983b**). ''Vestibular acoustic reception in the guinea pig: A saccular function?,'' Acta Oto-Laryngol. **95**, 211–217.

Colebatch, J., Halmagyi, G., and Skuse, N. (**1994**). ''Myogenic potentials generated by a click-evoked vestibulocollic reflex,'' J. Neurol. Neurosurg. Psychiatry **57**, 190–197.

Darwin, C. (**1871**). *The Descent of Man, and Selection in Relation to Sex* (Murray, London).

Depue, R. A., and Collins, P. F. (**1999**). ''Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation and extraversion,'' Behav. Brain Sci. **22**, 491–517.

Dibble, K. (**1995**). ''Hearing loss and music,'' J. Audio Eng. Soc. **43**(4), 251–266.

Di Lazzaro, V., Quartarone, A., Higuchi, K., and Rothwell, J. (**1995**). ''Short-latency trigemino-cervical reflexes in man,'' Exp. Brain Res. **102**, 474–482.

Ervin, F. R., and Martin, J. (**1986**). ''Neurophysiological bases of the primary emotions,'' in *Emotion-Theory, Research and Experience, Vol. 3, Biological Foundations of Emotion*, edited by R. Plutchik and H. Kellerman (Academic, Orlando), pp. 145–170.

Eysenk, H. J. (**1990**). ''Biological dimensions of personality,'' in *Handbook of Personality: Theory and Research*, edited by L. A. Pervin (Guildford, New York), pp. 244–276.

Ferber-Viart, C., Dubreuil, C., and Duclaux, R. (**1999**). ''Vestibular evoked myogenic potentials in humans: A review,'' Acta Otolaryngol. Suppl. (Stockh.) **119**, 6–15.

Fernandez, C., and Goldberg, J. M. (**1976**). ''Physiology of the peripheral neurones innervating the otolith organs of the squirrel monkey. Parts 1, 2 and 3,'' J. Neurophysiol. **39**, 970–1008.

Florentine, M., Hunter, W., Robinson, M., Ballou, M., and Buus, S. (**1999**). ''On the behavioural characteristics of loud music listening,'' Ear Hear. **19**(6), 420–428.

Geissmann, T. (**2000**). ''Gibbon songs and human music from an evolutionary perspective,'' in *The Origins of Music*, edited by N. L. Wallin, B. Merker, and S. Brown (MIT, Cambridge, MA), pp. 103–134.

Gray, J. A. (**1987**). *The Psychology of Fear and Stress* (Cambridge University Press, Cambridge, UK).

Guilford, J. (**1954**). ''System in the relationship of affective value to frequency and intensity of auditory stimuli,'' Am. J. Psychol. **67**, 691–695.

Laird, D., and Cove, K. (**1929**). ''Psychological measurements of annoyance as related to pitch and loudness,'' J. Acoust. Soc. Am. **1**, 158–163.
Lewis, E. R., and Lombard, R. E. (**1988**). ''The amphibian inner ear,'' in *The Evolution of the Amphibian Auditory System*, edited by B. Fritzsch, M. J. Ryan, W. Wilcynski, T. E. Hetherington, and W. Walkowiak (Wiley, New York), pp. 93–123.
Lewis, E. R., and Narins, P. M. (**1985**). ''Do frogs communicate with seismic signals?'' Science **227**, 187–189.
Lewis, E. R., Leverenz, E., and Bialek, W. (**1985**). *The Vertebrate Inner Ear* (CRC Boca Raton, FL).
Katafuchi, T., Puthraya, K., Yoshimatsu, H., and Oomara, Y. (**1987**). ''Response of rat lateral hypothalamic neuron to vestibular nuclei stimulation,'' Brain Res. **400**, 62–69.
Kraus, N., and McGee, T. (**1992**). ''Electrophysiology of the auditory system,'' in *The Mammalian Auditory Pathway: Neurophysiology*, edited by A. N. Popper and R. R. Fay (Springer-Verlag, New York), pp. 335–404.
McCormick, C. A. (**1988**). ''Evolution of auditory pathways in amphibia,'' in *The Evolution of the Amphibian Auditory System*, B. Fritzsch, M. J. Ryan, W. Wilcynski, T. E. Hetherington, and W. Walkowiak (Wiley, New York), pp. 587–612.
McCormick, C. A. (**1992**). ''Evolution of central auditory pathways in anamniotes.'' in *The Evolutionary Biology of Hearing*, edited by D. B. Webster, R. R. Fay, and A. N. Popper (Springer-Verlag, New York), pp. 323–350.
McCue, M., and Guinan, J. (**1994**). ''Acoustically responsive fibres in the vestibular nerve of the cat,'' J. Neurosci. **14**(10), 6058–6070.
McCue, M., and Guinan, J. (**1995**). ''Spontaneous activity and frequency selectivity of acoustically responsive vestibular afferents in the cat,'' J. Neurophysiol. **74**(4), 1563–1572.
Merker, B. (**2000**). ''Synchronous chorusing and human orgins,'' in *The Origins of Music*, edited by N. L. Wallin, B. Merker, and S. Brown (MIT, Cambridge, MA), pp. 315–328.
Miller, G. (**2000**). ''Evolution of human music through sexual selection,'' in *The Origins of Music*, edited by N. L. Wallin, B. Merker, and S. Brown (MIT, Cambridge, MA), pp. 329–360.
Moffat, A., and Capranica, R. (**1976**). ''Auditory sensitivity of the sacculus in the American toad (*Bufo americanus*),'' J. Comp. Physiol. **105**, 1.
Moore, B. C. J. (**1989**). *An Introduction to the Psychology of Hearing*, 3rd ed. (Academic, London).
Neary, T. J. (**1988**). ''Forebrain auditory pathways in Ranid frogs,'' in *The Evolution of the Amphibian Auditory System*, edited by B. Fritzsch, M. J. Ryan, W. Wilcynski, T. E. Hetherington, and W. Walkowiak (Wiley, New York), pp. 233–352.
Pfaffmann, C. (**1960**). ''The pleasures of sensation,'' Psychol. Rev. **67**, 253–268.
Pickering, A.D., and Gray, J. A. (**1999**). ''The neuroscience of personality,'' in *Handbook of Personality: Theory and Research*, edited by L. A. Pervin and O. P. John (Guilford, London), pp. 277–299.
Pitcher, T. J. (**1986**). *Behavior of Teleost Fish* (Chapman and Hall, London).
Pollack, I. (**1952**). ''Comfortable listening levels for pure tones in quiet and noise,'' J. Acoust. Soc. Am. **24**(2), 158–162.
Popper, A., Platt, C., and Saidal, W. (**1982**). ''Acoustic functions in the fish ear,'' Trends Neurosci. **August**: 276–280.
Rolls, E. T. (**1975**). *The Brain and Reward* (Pergamon, Oxford).
Silverman, S. (**1947**). Ann. Otol. Rhinol. Laryngol. **56**, 658.
Singer, W., and Young, P. (**1941a**). ''Studies in affective reaction I: A new affective rating scale,'' J. Gen. Psychol. **24**, 281–302.
Singer, W., and Young, P. (**1941b**). ''Studies in affective reaction II: Dependence of affective ratings upon the stimulus-stimulus situation,'' J. Gen. Psychol. **24**, 303–325.
Todd, N. P. M. (**1993**). ''Vestibular feedback in music performance,'' Music Perception **10**(3), 379–382.
Todd, N. P. M., and Cody, F. (**2000**). ''Vestibular responses to loud dance music: A physiological basis for the 'rock and roll threshold?'' J. Acoust. Soc. Am. **107**(1), 496–500.
Todd, N. P. M., Cody, F., and Banks, J. (**2000**). ''A saccular origin of frequency tuning in myogenic vestibular evoked potentials?: Implications for human responses to loud sounds,'' Hear. Res. **141**, 180–188.
Todd, N. P. M., Cody, F., and Banks, J. (**2001**). ''Bipolar vs monopolar recordings of myogenic vestibular evoked potentials,'' Hear. Res. **152**, 175–176.
Townsend, G., and Cody, D. (**1971**). ''The average inion response evoked by acoustic stimulation: its relation to the sacculus,'' Ann. Otol. Rhinol. Laryngol. **80**, 121–131.
Uchino, Y., Sato, H., Sasaki, M., Imagawa, M., Ikegami, H., Isu, N., and Graf, W. (**1997**). ''Sacculocollic reflex arcs in cats,'' J. Neurophysiol. **77**, 3003–3012.
Vitz, P. C. (**1973**). ''Preference for tones as a function of frequency and intensity,'' Percept. Psychophys. **11**, 84–88.
Wit, H., Bleeker, J., and Mulder, H. (**1984**). ''Response of pigeon vestibular nerve fibres to sound and vibration with audio frequencies,'' J. Acoust. Soc. Am. **75**, 202–208.
Wundt, W. (**1910**). *Grundzüge der Physiologischen Psychologie. Zweiter Band (sechsten Auflage)* (Engelmann, Leipzig).
Yates, B. J. (**1996**). ''Vestibular influences on the autonomic nervous system,'' in *New Directions in Vestibular Research*, edited by S. M. Highstein, B. Cohen, and J. A. Buttner-Ennever, Ann. N.Y. Acad. Sci. **781**, 458–473.
Young, E., Fernandez, C., and Goldberg, J. (**1977**). ''Responses of squirrel monkey vestibular neurons to audio-frequency sound and head vibration,'' Acta Oto-Laryngol. **84**, 352–360.

Effects of asynchrony and ear of presentation on the pitch of mistuned partials in harmonic and frequency-shifted complex tones

Jeffrey M. Brunstrom[a] and Brian Roberts[b]
School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, England

(Received 3 January 2001; revised 10 April 2001; accepted 16 April 2001)

When a partial of a periodic complex is mistuned, its change in pitch is greater than expected. Two experiments examined whether these partial-pitch shifts are related to the computation of global pitch. In experiment 1, stimuli were either harmonic or frequency-shifted (25% of $F0$) complexes. One partial was mistuned by $\pm 4\%$ and played with leading and lagging portions of 500 ms each, relative to the other components (1 s), in both monaural and dichotic contexts. Subjects indicated whether the mistuned partial was higher or lower in pitch when concurrent with the other components. Responses were positively correlated with the direction of mistuning in all conditions. In experiment 2, stimuli from each condition were compared with synchronous equivalents. Subjects matched a pure tone to the pitch of the mistuned partial (component 4). The results showed that partial-pitch shifts are not reduced in size by asynchrony. Similar asynchronies are known to produce a near-exclusion of a mistuned partial from the global-pitch computation. This mismatch indicates that global and partial pitch are derived from different processes. The similarity of the partial-pitch shifts observed for harmonic and frequency-shifted stimuli suggests that they arise from a grouping mechanism that is sensitive to spectral regularity. [DOI: 10.1121/1.1379079]

PACS numbers: 43.66.Ba, 43.66.Fe, 43.66.Hg, 43.66.Rq [SPB]

I. INTRODUCTION

When a partial in a harmonic complex tone is mistuned, it tends to stand out from the other components as a separate perceptual entity (e.g., Moore *et al.*, 1986; Hartmann *et al.*, 1990). Mistuning also produces a *partial-pitch shift*—typically, the pitch of the component is displaced from its expected value in the same direction as the mistuning (Hartmann *et al.*, 1990; Hartmann and Doty, 1996; Lin and Hartmann, 1998; Roberts and Brunstrom, 1998). Proposals have been made concerning the origin of these pitch shifts, but these have either been shown to be flawed or have yet to receive thorough empirical evaluation.

It has long been known that the pitch of a partial in a complex tone may not correspond exactly with its pitch when heard in isolation (Terhardt, 1971). These pitch shifts have been interpreted as a peripheral phenomenon, resulting from asymmetric masking interactions between partials in the excitation pattern evoked by the complex tone. However, Hartmann *et al.* (1990) noted that, unlike the association between direction of mistuning and sign of pitch shift that they observed, an implementation of Terhardt *et al.*'s (1982) pitch model predicted positive pitch shifts for downward as well as for upward mistunings. This led Hartmann and Doty (1996) to develop a more elaborate peripheral model, incorporating both place and time interactions between neighboring partials. Specifically, both partial-masking interactions in a simulated excitation pattern and peak locations in a simulated interspike-interval histogram were used to predict the pitch of a mistuned harmonic. This hybrid place-and-time model was successful in predicting the overall relationship between direction of mistuning and sign of pitch shift, but was still unable to account for some key aspects of the empirical data. In particular, it could account for neither the pitch shifts resulting from mistuning the fundamental component, nor the general tendency of pitch shifts to saturate around $\pm 4\%$ mistuning and to decrease beyond.

de Cheveigné (1997) first suggested that the partial-pitch shifts resulting from mistuning might not have a peripheral origin. Instead, he proposed that they arise from a central interaction between an internal representation of the mistuned partial and a harmonic ''sieve'' or ''template'' (Gerson and Goldstein, 1978; Duifhuis *et al.*, 1982) that has been activated by a set of in-tune partials. Lin and Hartmann (1998) provided support for the idea of a central template mechanism by demonstrating that the characteristic association between direction of mistuning and sign of pitch shift can be found even for components isolated from the rest of the complex tone by gaps of one or more harmonic positions.

The concept of a harmonic template originally arose in the context of pattern-recognition theories of pitch perception (Gerson and Goldstein, 1978). Duifhuis *et al.* (1982) envisaged the template as an adjustable mechanism, with regularly spaced slots in the frequency domain. These slots restrict the set of frequency components contributing to global pitch to those with a harmonic or near-harmonic relationship to one another. This concept has also been used to account for the *global-pitch shifts* that result when a partial in a harmonic complex tone is mistuned (Moore *et al.*, 1985;

[a]Current address: Department of Human Sciences, Loughborough University, Loughborough, Leicestershire LE11 3TU, England. Electronic mail: J.M.Brunstrom@Lboro.ac.uk
[b]Contact for reprints. Electronic mail: B.Roberts@Bham.ac.uk

Darwin *et al.*, 1994). Moore *et al.* (1985) mistuned individually each of the low-numbered harmonics (1–6) in a 12-harmonic complex tone and measured the corresponding changes in global pitch. Although each component differed in its weighted contribution to global pitch, mistuning a component produced a reliable change in the same direction as the mistuning. The pitch-shift function associated with each component was almost linear for mistunings of up to ±3%, was greatest for mistunings around ±4%, and declined towards zero as the mistuning was increased to ±8%. This suggests that the template works not on an all-or-none basis, but rather that a harmonic makes a progressively reduced contribution to the global pitch of the complex as its mistuning is increased beyond ±3%.

de Cheveigné (1997) has argued that both global and partial pitches arise from the same mechanism, and has proposed a model of the interaction between a mistuned component and a hypothetical harmonic template to account for the pattern of partial-pitch shifts reported by Hartmann and Doty (1996). He reasoned that the internal representation of a component has associated noise, leading to a distribution of estimates of the component's frequency. For a fixed degree of mistuning, some estimates will fall closer to the center frequency of a template slot than others. Estimates closer to the center frequency will lead to greater perceptual fusion, thereby reducing the salience of the partial pitch, whereas estimates further away will be heard as more salient. These changes in salience will bias the distribution of listeners' reported pitch estimates away from the slot center frequency, resulting in a measured partial-pitch shift.

More recently, de Cheveigné (1999) has elaborated on the mechanism that is responsible for harmonic fusion and partial-pitch shifts using the concept of ''harmonic cancellation'' (Lea, 1992; Summerfield and Culling, 1992). Harmonic cancellation is the process by which in-tune components can be suppressed on the basis of their common periodicity, thereby enhancing the perception of a mistuned target. de Cheveigné (1993) proposed a model of harmonic cancellation based on time-domain processing using a ''neural cancellation filter,'' which he subsequently developed into a model of global-pitch perception (de Cheveigné, 1998). de Cheveigné's (1999) proposal conceptualizes the neural cancellation filter as a gating neuron with three input pathways—one direct and excitatory and the others delayed and inhibitory. The output from this circuit is gated ''on'' unless an excitatory spike on the direct input pathway is coincident with an inhibitory spike on either of the delayed input pathways. Hence coincident inputs are removed from the output of the circuit. One of the inputs has a delay tuned to the period of the background complex. This suppresses the response of the circuit to in-tune components, making the mistuned component relatively more salient. The other delay is variable, and the value that produces a minimum output provides an estimate of the period of the mistuned partial. de Cheveigné (1999) proposes that partial-pitch shifts result from the stochastic nature of neural firing and its interaction with the two delayed pathways. Moreover, he demonstrates that it is possible to model Hartmann and Doty's (1996) empirical data with an appropriate choice of parameters.

de Cheveigné's (1999) model elegantly incorporates the three major aspects of harmonic complex-tone perception: global pitch, partial pitch, and the perceptual segregation of a mistuned component. Nevertheless, there are empirical data that are difficult to reconcile with the notion of a common mechanism governing both global- and partial-pitch perception. First, the effects of small degrees of mistuning on the salience of a component and on its contribution to global pitch are different. Hartmann *et al.* (1990) found that mistuning a harmonic by as little as ±0.5% can be sufficient to increase significantly its salience. This contrasts with the weighted contribution that a component makes to global pitch, which is undiminished for mistunings up to ±3% (Moore *et al.*, 1985). Second, Moore *et al.* (1985) restricted the test set of components to harmonics 1–6, because pilot work had shown that higher-numbered harmonics did not make a significant contribution to global pitch. In contrast, Hartmann and Doty (1996) observed clear partial-pitch shifts for mistuned harmonics up to the ninth. Third, Roberts and Brunstrom (1998) found that subjects can successfully identify a mistuned component when the complex tone is regular but inharmonic in spectral structure. These stimuli were derived from harmonic complex tones either by frequency shifting [addition of a fixed increment to the frequency of each partial (de Boer, 1976; Patterson, 1973)] or by spectral stretch (addition of a cumulative increment to intercomponent spacing with increasing component number). Roberts and Brunstrom (1998) argued that a single harmonic template could not fit their stimuli well enough to account for their results. Supplementary to the measure of component salience, the design of their study also enabled an estimation of partial-pitch shifts for the mistuned components. For both shifted and stretched stimuli, pitch shifts were broadly comparable to those observed in harmonic stimuli. This suggests that partial-pitch shifts are not reliant solely on the detection of a common cross-channel periodicity, which is a prerequisite of time-domain models of global-pitch perception (e.g., Meddis and O'Mard, 1997; de Cheveigné, 1998).

In the light of this evidence, the experiments reported here were designed to explore further the relationship between the perceptual estimation of global pitch and the occurrence of partial-pitch shifts. The approach taken was to discover whether or not factors that are known to influence the contribution of a given component to global pitch also produce predictable changes in the pattern of partial-pitch shifts associated with mistuning that component.

II. EXPERIMENT 1

Onset asynchrony is arguably the most important factor influencing the auditory grouping of concurrent acoustic elements (see Bregman, 1990). Darwin and Ciocca (1992) measured the effects of onset asynchrony on the contribution to global pitch made by the mistuned fourth harmonic of a periodic complex tone with a fundamental ($F0$) frequency of 155 Hz. The mistuned component always ended in synchrony with the rest of the complex. In their second experiment, the degree of mistuning was fixed at ±3% and the onset asynchrony was varied systematically in steps between 0 and 640 ms. Darwin and Ciocca (1992) found that the

contribution of the mistuned harmonic to the pitch of the complex was reduced when it began at least 80 ms before the other components, and was negligible when the onset asynchrony was increased to 320 ms. Ciocca and Darwin (1993) confirmed that this effect was primarily a consequence of auditory grouping rather than peripheral adaptation. Darwin and Ciocca (1992) also explored the effects on global pitch of presenting the mistuned fourth component and the rest of the complex to opposite ears. They found that, in the absence of asynchrony, dichotic presentation of the mistuned component reduced its contribution to global pitch, but this effect was less than that brought about by onset asynchrony. Neither kind of manipulation has been applied before in the context of measures of the pitch of the mistuned partial itself.

Experiment 1 sought to determine whether or not partial-pitch shifts are preserved when an onset and offset asynchrony, greater than 300 ms, is applied to a mistuned component. Evidence for substantial partial-pitch shifts would indicate that partial- and global-pitch shifts are not governed by a common mechanism. In this experiment, the effects of asynchrony were assessed using a combination of two factors: (a) type of complex (harmonic or frequency shifted) and (b) ear of presentation (monaural or dichotic) of the mistuned component relative to the others (the *spectral frame*). Evidence for a similar pattern of partial-pitch shifts in harmonic and in frequency-shifted complexes would suggest that partial pitches are not reliant on cross-channel comparison of periodicity estimates of the components in the spectral frame. Evidence of pitch shifts in the dichotic context would indicate that they are not produced solely by peripheral interactions between neighboring partials.

A. Method

1. Overview

Subjects heard stimuli in which one component was mistuned, relative to the components of a spectral frame, and lengthened so that it began before and ended after the frame. Subjects judged the pitch of the portion of the mistuned component concurrent with the spectral frame, with respect to that of the leading and lagging portions. A ''higher'' judgment was required when subjects heard a rise followed by a fall in the pitch of the mistuned component, and vice versa for a ''lower'' judgment. A greater proportion of ''higher'' or ''lower'' responses was taken as evidence for a positive or a negative pitch shift, respectively.

2. Subjects

Eight subjects participated, all of whom reported having normal hearing. One subject was the first author and the others were research students at the University of Birmingham. Of these, four had previous experience of auditory perception experiments. All subjects successfully passed a screening procedure designed to ensure that they were able both to detect small changes in pitch and to identify the direction of the change (higher or lower). Each trial comprised two 1-s pure tones with a 500-ms silent interval between them. The frequency of the first tone was selected from a range corresponding to that of the components tested in the main experiment, and the frequency of the second tone was either 1% higher or lower than that of the first. The task was to judge which tone was higher in pitch, and the performance criterion was an overall accuracy of 75% or greater correct.

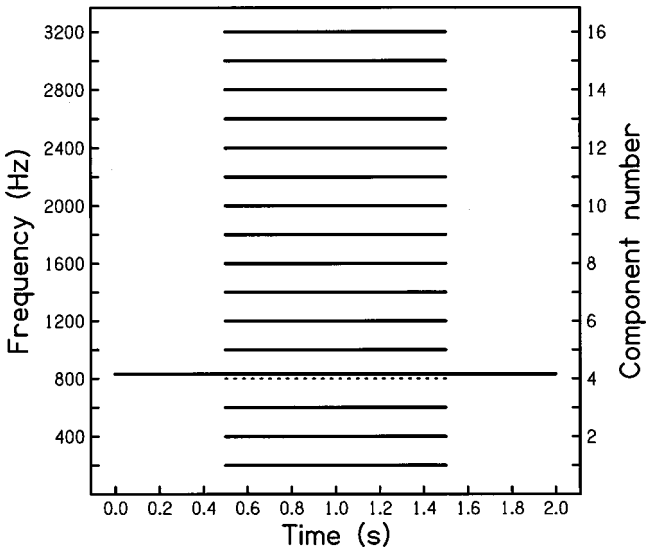


FIG. 1. Schematic of the stimuli used in the harmonic and monaural (HM) condition of experiment 1. In this example, component 4 has been mistuned by +4%. The dotted line corresponds to the in-tune frequency for component 4.

3. Stimuli and conditions

Each subject was tested in four conditions; harmonic and monaural (HM), harmonic and dichotic (HD), frequency shifted and monaural (SM), and frequency shifted and dichotic (SD). Figure 1 shows a schematic of a stimulus presented in the HM condition. In this condition, the stimuli comprised the first 16 partials of an $F0$ frequency of 200 Hz. The subset of partials tested was restricted to components 1–12, because Hartmann *et al.* (1990) found that listeners could not reliably match mistuned harmonics above the twelfth.

On any given trial, one of the test partials was mistuned by ±4% from its original value and was lengthened to make it asynchronous with the other components (the spectral frame). The mistuned component was heard in isolation for the first 500 ms, was then accompanied by the spectral frame for 1 s, and was again heard in isolation for the last 500 ms. Listeners were informed that the change in pitch to be judged was small. They were instructed to pay particular attention to the points at which the mistuned component entered and left the spectral frame, because pilot work had identified the pitch transitions at these points to be the most salient cue.

Both the mistuned component and the spectral frame were presented to the left ear only in the HM condition. In the HD condition, the spectral frame was presented to the right ear instead. The stimuli for the frequency-shifted conditions (SM and SD) were identical to their harmonic counterparts except that an increment of 50 Hz (i.e., 25% of $F0$) was added to the frequency of each harmonic. This value represents the maximum deviation from harmonicity that is

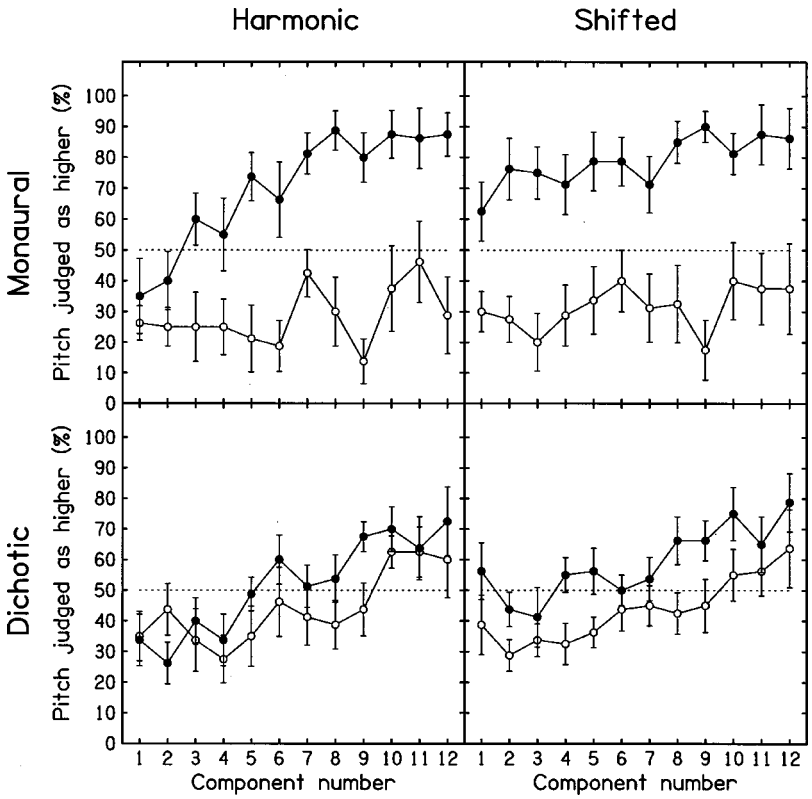


FIG. 2. Results for experiment 1. Mean responses are shown for eight subjects, with intersubject standard errors, indicating the percentage of occasions on which the concurrent portion of the mistuned component was judged as higher in pitch than the leading and lagging portions. The data for each condition are presented in a separate panel, and data for positive and negative mistunings are indicated by filled and open circles, respectively. Mean responses predicted by chance are indicated with a horizontal dotted line.

possible through frequency shifting (50% of $F0$ would give rise to an odd-harmonic complex tone). The mistunings of $\pm 4\%$ were applied after this modification, and were chosen because they typically produce maximum pitch shifts (Hartmann and Doty, 1996). Both the degree of mistuning and degree of frequency shift employed are also consistent with our previous work (Roberts and Brunstrom, 1998; Brunstrom and Roberts, 2000).

The durations of the mistuned partial (2 s) and the spectral frame (1 s) both included linear onset and offset ramps of 20 ms each. Each component, including the mistuned target, was set to 60 dB SPL, and the components of the spectral frame all began in sine phase. All stimuli were generated using MITSYN software (Henke, 1990). Stimuli were synthesized at a sampling rate of 16 kHz and played back via a 16-bit digital-to-analog converter (Data Translation DT2823). They were low-pass filtered (Fern Developments EF16X module; corner frequency=5.2 kHz, roll-off=100 dB/oct) and presented binaurally over Sennheiser HD 480-13II earphones. The levels of the stimuli were set using programmable attenuators (0.25-dB steps), and were calibrated with a sound-level meter (Brüel and Kjaer, type 2209, linear weighting) connected to the earphones by an artificial ear (type 4153). Stimuli were played to the listeners in a double-walled sound-attenuating chamber (Industrial Acoustics 1201A).

4. Procedure

Subjects made their responses using a keyboard and received no feedback of any kind. They were tested in two sessions of about 45–60 min each. In one session, stimuli from the monaural conditions were tested, and, in the other, stimuli from the dichotic conditions were tested. The order of these sessions was counterbalanced across subjects. In each session, harmonic and frequency-shifted stimuli were presented together in ten blocks, each in a new randomized order. This provided ten responses for each stimulus, giving a total of 480 trials per session (10 sets $\times$ 12 components $\times$ 2 directions of mistuning $\times$ 2 conditions). From these responses, the proportion of ''higher'' judgements was calculated separately for each component, mistuning direction, and condition. Subjects were given a practice set of one block of trials (48 trials) at the beginning of each session. These data were not included in the analysis.

B. Results

Figure 2 shows the mean effect of direction of mistuning and of component number on the proportion of responses selected as ''higher'' for each condition separately. The dashed lines represent the proportion that one would expect if the responses were made randomly. A separate two-way repeated-measures ANOVA was applied to the data for each condition. In each case, the direction of mistuning was found to be a significant factor {[HM: $F(1,7)=11.33$, $p<0.025$]; [SM: $F(1,7)=14.52$, $p<0.01$]; [HD: $F(1,7)=6.43$, $p<0.05$]; [SD: $F(1,7)=12.45$, $p<0.01$]}. Indeed, positive mistuning produced a greater proportion of ''higher'' responses than negative mistuning in all but two cases (components 1 and 2 in the HD condition).

Response bias was assessed by computing the overall proportion of ''higher'' responses for each condition, collapsed across component number and direction of mistuning. The mean proportions, with intersubject standard errors, were as follows: HM=0.49±0.021, SM=0.55±0.017, HD =0.48±0.029, SD=0.51±0.013. All the means were close

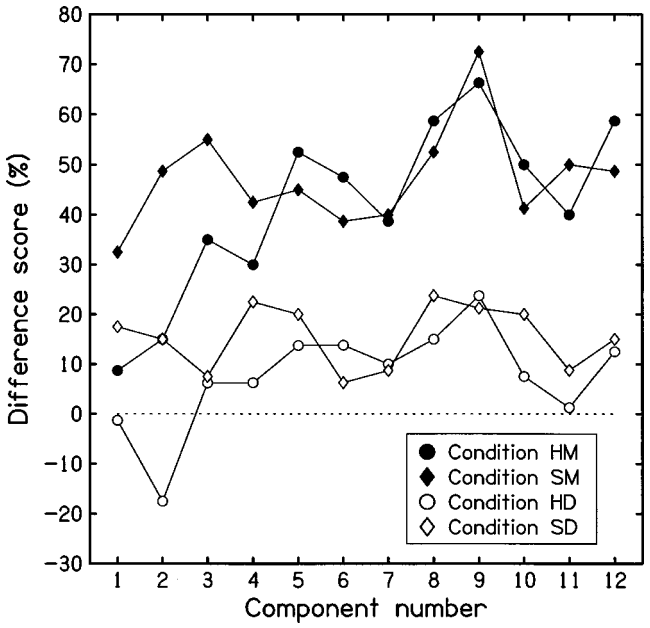


FIG. 3. Results for experiment 1. Mean difference scores are shown to illustrate further the effect of direction of mistuning on pitch judgments in the four conditions. These values are calculated by subtracting the proportion of ''higher'' judgments associated with negative mistuning from the proportion associated with positive mistuning. Symbols corresponding to the different conditions are given in the figure inset.

to 0.5, which suggests little overall response bias. However, it is evident from Fig. 2 that the proportion of ''higher'' responses tends to rise with increasing component number in all conditions. This tendency was quantified by collapsing the data across direction of mistuning and using linear regression (percentage ''higher'' responses versus component number). The slopes of these functions (HM=2.92, SM =1.32, HD=3.29, SD=2.65) were all positive ($r>0.8$, $p<0.001$ in all cases). Based solely on the monaural data, one might appeal by way of explanation to pitch shifts produced by partial masking (Terhardt, 1971). This is because the masking of the lower and higher skirts of the excitation pattern evoked by a given component becomes increasingly asymmetric for higher component numbers. However, the regression slopes showed a similar pattern in the dichotic conditions, for which peripheral interactions between the target component and spectral frame were absent. Therefore, the origin of this tendency in the data is unclear.

To facilitate a cross-condition comparison of the effect of mistuning on partial pitch, a set of difference scores was calculated for each subject, component, and condition. Each difference score was derived by subtracting the percentage of ''higher'' responses associated with negative mistuning from the percentage of ''higher'' responses associated with positive mistuning. These scores provide a measure of the effect of mistuning from which any offset towards ''higher'' or ''lower'' responses has been removed (see above). Mean across-subject difference scores are shown in Fig. 3. These were analyzed using a three-way repeated-measures ANOVA, which revealed three significant main effects. First, the magnitude of the difference scores varied with component number [$F(11,77)=2.86$, $p<0.005$]. It can be seen that the difference scores are somewhat smaller for the lower-numbered components, and that they appear to peak around component 9. Second, the difference scores were much greater in the monaural conditions than in the dichotic conditions [$F(1,7)=8.70$, $p<0.025$]. This effect is very apparent in Figs. 2 and 3. Third, the analysis revealed an effect of complex type [harmonic versus shifted: $F(1,7)=10.50$, $p<0.025$]. Difference scores were larger for the frequency-shifted conditions (across-component means and intersubject standard errors: SM=47.3±10.2, SD=15.5±4.4) than for their harmonic counterparts (HM=41.7±12.4, HD=7.6 ±3.0). This effect may be due to the relatively smaller difference scores associated with the low-numbered components in the harmonic conditions. However, it should be noted that the interaction between complex type and component number did not reach significance [$F(11,77)=1.47$, $p=0.158$]. No other interaction term approached significance.

C. Discussion

Subjective reports of partial-pitch shifts were clearly influenced by the direction of mistuning applied to the asynchronous component. Although reduced relative to the monaural conditions, this effect was also present in the dichotic conditions, indicating that partial-pitch shifts are not dependent solely on peripheral processing. Furthermore, partial-pitch shifts were reported in the frequency-shifted as well as the harmonic conditions. This implies that partial-pitch shifts can occur despite a lack of common periodicity amongst the components of the spectral frame. The possible basis for partial-pitch shifts in frequency-shifted complexes is considered further in the General Discussion (Sec. IV B).

Perhaps the most striking result from experiment 1 is that partial-pitch shifts appear to be robust in the presence of an onset and offset asynchrony of 500 ms each between the mistuned partial and the spectral frame. This evidence for a mismatch between the effect of asynchrony on global-pitch shifts (Darwin and Ciocca, 1992) and on partial-pitch shifts suggests that these phenomena are not governed by a common mechanism. There is, however, an alternative explanation for this inconsistency that is difficult to rule out on the basis of the categorical data collected in experiment 1. It might be that onset and offset asynchronies of 500 ms greatly reduce the size of partial-pitch shifts, but do not eliminate them completely. Subjects in experiment 1 were required to identify the *presence* of partial-pitch shifts rather than to estimate their magnitude. Perhaps, therefore, the results reflect partial-pitch shifts which, though detectable, were very small.[1] In this regard, it is noteworthy that Darwin and Ciocca's (1992) first experiment provides some evidence for a residual global-pitch shift even when the mistuned partial has an onset asynchrony of 300 ms. To resolve this issue, in experiment 2 we explored the possibility that subjects had been listening and responding to residual but detectable pitch shifts.

III. EXPERIMENT 2

To determine whether or not residual pitch shifts might account for the results of experiment 1, the magnitude of the partial-pitch shifts was measured explicitly. In each case, for

comparison, an estimate was also made of the partial-pitch shift on a mistuned component when it was presented in synchrony with the onset and offset of the spectral frame. Pitch shifts were measured using a pitch-matching task of the kind used previously by Lin and Hartmann (1998). For reasons of economy, mistunings were imposed only on component 4 of the complex tones tested. This was the component chosen by Darwin and Ciocca (1992) to explore the effects of mistuning on global pitch. It is reasonably well resolved from its neighbors and falls within the dominant region for global pitch calculation (component range 2–5: Ritsma, 1967; Moore *et al.*, 1985).

A. Method

1. Overview

Subjects listened to a 16-component complex tone in which the fourth partial was mistuned by ±4% from its original position. Eight variants of this complex tone were tested. Each was derived from an orthogonal combination of three binary-state variables: (a) whether the mistuned component was synchronous or asynchronous with the spectral frame, (b) whether the mistuned component and spectral frame were presented monaurally or dichotically, and (c) whether the complex tone was harmonic or frequency shifted. During each trial, subjects heard one of these stimuli, followed by a repeating, adjustable, pure tone. On occasions when the mistuned component and spectral frame were synchronous, subjects were instructed to match the adjustable tone to the pitch of a pure tone embedded in the complex. For the asynchronous stimuli, subjects were instructed to match the adjustable tone to the pitch of the concurrent portion of a pure tone running through the complex.

2. Subjects

The same eight subjects who participated in experiment 1 were also tested in this experiment.

3. Stimuli

The asynchronous stimuli were identical to their counterparts in experiment 1. The synchronous stimuli were identical to the corresponding asynchronous stimuli, except that the leading and lagging steady-state portions of the mistuned component were deleted and replaced with 500-ms silences. The interval between the end of the spectral frame and the first adjustable pure tone was always 1 s in duration. Each adjustable tone had a duration of 310 ms (including 20-ms linear onset and offset ramps), and was played at the same level as each component (60 dB SPL). Adjustable tones were played consecutively with an intertone interval of 500 ms. During each trial, subjects were allowed to listen to the test complex as many times as desired to refresh their memory of the pitch of the mistuned component. When this option was selected, the sequence of pure tones stopped and a single instance of the test complex was played. The interval between the last adjustable tone and the spectral frame was 1 s long, and the adjustable tones began again 1 s after the spectral frame ended. At the beginning of a trial, the frequency of the adjustable pure tone always corresponded to the true frequency of the mistuned fourth component, and this tone was always played in the same ear as the mistuned component to be matched.

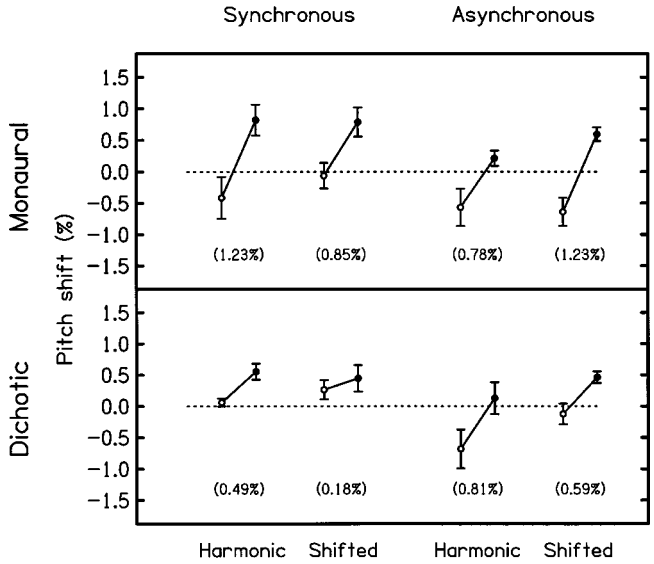


FIG. 4. Results for experiment 2. Mean pitch shifts for each stimulus are shown across eight subjects, with intersubject standard errors. The pitch-shift gradient for each stimulus is represented by a straight line joining the values resulting from negative and positive mistunings, which are indicated by open and filled circles, respectively. The corresponding numerical value (given below in brackets) is the difference between these values. The horizontal dotted lines represent veridical matches to the frequency of the mistuned partial.

The matching range of the adjustable pure tones was limited to ±4% of the mistuned target, quantized in steps of 1 Hz. This range was chosen because it was broad enough to accommodate the magnitude of any partial-pitch shift (see Hartmann *et al.*, 1990; Hartmann and Doty, 1996; Roberts and Brunstrom, 1998), yet narrow enough to prevent subjects from matching the pure tone to an adjacent component in the spectral frame. The long time interval between the end of the mistuned component and the sequence of adjustable tones was chosen to exclude the possibility of pitch shifts on the adjustable tones resulting from prior stimulation in the same frequency region (Rakowski and Hirsh, 1980).

4. Procedure

Subjects adjusted the matching tone using a trackball mouse. If the upper or lower limit of the range was reached during an adjustment, the message ''top of range'' or ''bottom of range'' was displayed, as appropriate. The buttons on the mouse were used to begin and end each trial and to replay the complex tone. Before recording their response, subjects were instructed to use the replay option to check that they were satisfied with the accuracy of their chosen match.

There were four sessions (monaural and synchronous, monaural and asynchronous, dichotic and synchronous, dichotic and asynchronous), within each of which harmonic and shifted stimuli were randomized together. Each session consisted of eight blocks of four different trials [2 types of complex (harmonic or shifted) × 2 mistunings (positive or negative)], giving a total of 32 trials. A new randomized

order was used in each block. Subjects were given a practice set of two blocks (eight trials) at the beginning of each session. These data were not included in the final analysis. Each session lasted approximately 45–60 min. The order of the sessions was counterbalanced across subjects using a Latin-square design.

B. Results

Following earlier studies (Hartmann and Doty, 1996; Lin and Hartmann, 1998), mean matches and intersubject standard errors are expressed in terms of percentage deviation from the true frequency of the mistuned fourth component. The results are displayed in Fig. 4. For each condition, a straight line links the points corresponding to the mean partial-pitch shifts produced by negative and positive mistunings (left-hand and right-hand sides, respectively). Each line has a positive gradient, which indicates that the partial-pitch shift produced by upward mistuning was more positive (or less negative) than that produced by downward mistuning. A four-way repeated-measures ANOVA confirmed that the main effect of direction of mistuning on the magnitude of the pitch shifts was highly significant [$F(1,7)=27.09$, $p<0.0025$]. To explore the influence of the three binary-state variables (state of synchrony, ear of presentation, type of complex), each pair of pitch shifts was replaced by a single value. These values were calculated for each subject by taking the difference between the percentage shift produced by corresponding positive and negative mistunings. Lin and Hartmann referred to these difference scores as *pitch-shift gradients*. These gradients were then analyzed using a three-way repeated-measures ANOVA.

The analysis revealed that the pitch-shift gradient was significantly larger for the monaural stimuli (mean and intersubject standard error: $1.02\%\pm0.23$) than for the dichotic stimuli ($0.52\%\pm0.08$) [$F(1,7)=7.60$, $p<0.05$]. Although the mean gradient was numerically larger for the asynchronous stimuli ($0.85\%\pm0.17$) than for the synchronous stimuli ($0.69\%\pm0.17$), this difference was not significant [$F(1,7)=0.91$, $p>0.25$]. Neither was the effect of the type of complex [$F(1,7)=0.53$, $p>0.25$]. The mean pitch-shift gradient for the asynchronous dichotic stimuli was greater than that for the synchronous dichotic stimuli, but the interaction between ear of presentation and state of synchrony fell short of significance [$F(1,7)=4.12$, $p=0.082$]. No other interaction term approached significance.

C. Discussion

Pitch-shift gradients were found to be comparable to those reported previously by Lin and Hartmann (1998), and to be of similar magnitude irrespective of whether or not the mistuned partial was presented in synchrony with the spectral frame. Two key aspects of the results from this experiment correspond closely with those from experiment 1. First, partial-pitch shift gradients were largely unaffected by frequency shifting the background components. The same was true for the categorical data collected in experiment 1. Second, gradients were much larger for the monaural stimuli than for the dichotic stimuli. In experiment 1, the association between the category of response (''higher'' or ''lower'') and the direction of mistuning (positive or negative) was also greater in the monaural conditions. This correspondence between the results from both experiments validates the method used in experiment 1 to compare the partial-pitch shifts produced by different types of stimuli. An advantage of the categorical method is the speed with which data can be gathered from the subjects.

IV. GENERAL DISCUSSION

The experiments reported here reveal how ear of presentation (monaural or dichotic), type of complex tone (harmonic or frequency shifted), and asynchrony influence partial-pitch shifts, and allow comparison with earlier work on global-pitch shifts (Darwin and Ciocca, 1992). The first of these variables has similar effects on partial and global pitch, but the other two do not. Overall, these findings indicate that partial- and global-pitch shifts are probably not governed by a common mechanism. The influences of the three variables, and their implications for the origins of partial pitch, are considered in detail below.

A. Ear of presentation

Dichotic presentation of the mistuned component reduces, but does not abolish, the associated partial-pitch shift. From experiment 2, the magnitude of this reduction can be estimated to be about 50%. Darwin and Ciocca (1992) found that dichotic presentation led to a reduction of around 25% in the effect of a mistuned component on the global pitch of an otherwise harmonic complex tone. The finding of significant partial-pitch shifts with dichotic presentation indicates that partial-pitch shifts, like their global-pitch counterparts, are primarily a central phenomenon. The moderate reduction observed for both kinds of pitch shift with dichotic presentation may reflect the loss of a contribution related to peripheral encoding, or may instead reflect the influence of auditory lateralization as a weak cue for simultaneous grouping (see Darwin, 1997). The finding that partial-pitch shifts arise centrally complements the observation of Brunstrom and Roberts (2000, experiment 3) that the effects of spectral pattern on the perceptual fusion of components in a frequency-shifted complex cannot be explained in terms of peripheral combination tones.

B. Type of complex

The application of a frequency shift equal to 25% of $F0$ to a harmonic complex tone produces a complex with a global pitch that is weak and ambiguous in comparison with the original sound (Patterson, 1973). The pitches heard are consistent with the $F0$ frequencies of the best-fitting harmonic templates for the frequency-shifted complex, assuming a greater weighting for components in the dominant region (Ritsma, 1967; Moore *et al.*, 1985). However, the *perceptual fusion* of partials in a complex tone cannot be explained simply by reference to harmonic relations.

Roberts and his colleagues have shown that a single even harmonic, in an otherwise odd-harmonic complex, is typically more salient than its odd neighbors (Roberts and

Bregman, 1991; Roberts and Bailey, 1993a, b, 1996a, b; Roberts, 1998). They proposed that spectral regularity, in a broader sense than harmonics relations, can be used for auditory grouping. Roberts and Brunstrom (1998) have since shown that auditory grouping processes are sensitive to the regular but inharmonic spectral structure of frequency-shifted and spectrally stretched complex tones. The dissociation between global pitch and perceptual fusion suggested by their findings has recently been explored further. Ciocca (1999) presented evidence that grouping by spectral regularity does not influence the perception of global pitch. Brunstrom and Roberts (2000) measured the perceptual salience of a pure-tone probe inserted at one of a set of positions within a three-component spectral gap in a frequency-shifted complex. They found that the probe became less salient when it coincided with one of the missing components, but not when it fell close to a harmonic multiple of one of the global pitches evoked by the complex. This finding was taken as evidence for an auditory-grouping template that is sensitive to regular but inharmonic spectral pattern, and hence distinct from the global-pitch template. The exact means by which an inharmonic template might be instantiated remains unclear (see Roberts and Brunstrom, 1998, p. 2336, for a discussion).

In the current study, the overall similarity in the pattern of pitch shifts observed for harmonic and shifted conditions extends the results of Roberts and Brunstrom (1998). In particular, this similarity suggests that partial-pitch shifts are related to the process of perceptual fusion rather than to the computation of global pitch. This interpretation is consistent with the idea that an integrated complex-tone percept relies on the perceptual suppression of the individual components that comprise it (Brunstrom and Roberts, 1998, 2000). However, de Cheveigné (1999) has argued that Roberts and Brunstrom's (1998) data for frequency-shifted stimuli merely reflect the combined outputs of more than one harmonic template. Specifically, each of these harmonic templates is envisaged as passing a local region of consecutive components, within which the deviation from harmonicity is not great. While this idea might plausibly account for the pattern of pitch shifts resulting from the mistuning of some of the components in our frequency-shifted complex, it cannot explain the partial-pitch shifts associated with the lowest set of components tested in experiment 1, which deviate most from harmonicity. This can be illustrated by analogy with the three-component frequency-shifted stimuli employed by Schouten *et al.* (1962) in their study of residue (i.e., global) pitch.

Consider a harmonic template applied locally to components 1–3 of the frequency-shifted complexes that we employed. If the second slot of this template is centered around component 2 (450 Hz), then it will have a slot spacing of 225 Hz (450/2). This means that slots 1 and 3 would require a tolerance around their center frequencies of at least 11.1% and 3.8%, respectively, for components 1 and 3 to pass through. In both cases, this exceeds considerably the mistuning required for a harmonic to be heard out from a periodic complex (1.3% to 2.1%; Moore *et al.*, 1986). For component 1, the required tolerance even exceeds the level of mistuning beyond which a harmonic makes little contribution to global pitch (>8%; Moore *et al.*, 1985). Furthermore, the frequency difference between a mistuned component (±4% in experiment 1) and the center frequency of the nearest slot in our hypothetical template will vary greatly with the direction of mistuning. For example, mistunings of +4% and −4% on component 3 would result in mistunings from the center frequency of the third template slot of −7.56% and +0.15%, respectively. Such differences might be expected to affect the magnitude and sign of the measured partial-pitch shifts for the frequency-shifted stimuli.

The discrepancies previously described cannot easily be reconciled with the fact that partial-pitch shifts around the low-numbered components are strongly correlated with the direction of mistuning in our frequency-shifted stimuli. Indeed, ''higher'' and ''lower'' judgments for components 1–3 were rather more symmetrical for positive and negative mistunings in the frequency-shifted conditions than in the harmonic conditions. Therefore, we propose that partial-pitch shifts arising from mistuning are a consequence of grouping by means of spectral regularity.

C. Asynchrony

The principal evidence for a dissociation between the computation of global pitch and partial pitch is provided by the asynchrony data. Both experiments reported here have shown that onset and offset asynchronies of 500 ms each had no observable effect on the magnitude of partial-pitch shifts, despite the evidence from previous work that an onset asynchrony around 300 ms can almost completely eliminate the contribution of a mistuned component from the calculation of global pitch (Darwin and Ciocca, 1992). The apparent robustness of partial-pitch shifts in the presence of asynchronies as great as 500 ms suggests that no asynchrony, however large, can abolish this effect of mistuning. This insensitivity to asynchrony is, we believe, the first instance of its kind reported in the literature, and it poses a conundrum.

It has been suggested that partial-pitch shifts arise from the operation of a template mediating the perceptual fusion of components (Roberts and Brunstrom, 1998). Furthermore, it is known that asynchrony is a powerful grouping cue (see Bregman, 1990). Indeed, even an in-tune harmonic can be heard out if it begins before and ends after the rest of the complex (e.g., Darwin, 1984a, p. 205). Why, therefore, is the perceptual segregation brought about by asynchrony unable to prevent the interaction of the mistuned component and the hypothetical template that leads to a partial-pitch shift? Two points underpin the explanation that is proposed in Sec. IV D. First, the time constant for the effect of asynchrony varies with the nature of the perceptual task. Second, exclusion of an acoustic element from a perceptual classification on the basis of asynchrony, or other grouping cues, is not always complete. These points are considered in turn.

Relatively little attention has been paid to the fact that the time constant associated with the effects of an onset asynchrony between one component and the rest of a complex tone varies with perceptual task. A notable exception is Hukin and Darwin (1995). These authors pointed out that auditory-streaming judgments (Dannenbring and Bregman,

1978; Bregman and Pinker, 1978), vowel-quality judgments (Darwin, 1984a, b; Darwin and Sutherland, 1984; Roberts and Moore, 1991), and profile-analysis judgments (Green and Dai, 1992) all indicate that an onset asynchrony of 30–50 ms is usually sufficient to affect performance substantially. A similar effect of asynchrony has since been observed on judgments of perceived lateralization (Hill and Darwin, 1996). In contrast, an onset asynchrony of 80 ms or more is needed before any appreciable effect occurs on global-pitch judgments, and an effect approaching full exclusion requires an onset asynchrony of about 300 ms (Darwin and Ciocca, 1992; Ciocca and Darwin, 1993). Hukin and Darwin (1995) confirmed that this discrepancy was not a consequence of differences in stimulus structure between these studies (such as differences in duration, or the use of flat-spectrum complex tones rather than voiced vowels). They compared the effects of onset asynchrony on the contribution of a mistuned component to the global pitch and phonemic identity of the same vowel-like stimuli, and found similar differences in time constant to those reported earlier. Hukin and Darwin (1995) argued that the discrepancy in time constants for pitch and timbre judgments is incompatible with the notion that auditory-grouping processes are all-or-none and independent of subsequent perceptual classifications. The long (perhaps infinite) time constant for partial-pitch shifts may be considered as a further example of this kind of discrepancy. Certainly, it is difficult to reconcile this observation with the idea of a common origin for global- and partial-pitch shifts.

The effect of onset asynchrony on the perceptual contribution of a component to a complex tone varies not only in time constant, but also in the maximum extent of exclusion. Our data indicate that partial-pitch shifts are not reduced by asynchrony, whereas asynchrony can produce a complete exclusion of a mistuned component from the computation of global pitch (Darwin and Ciocca, 1992). Other studies have shown intermediate degrees of perceptual exclusion. For example, Roberts and Moore (1991) observed that the contribution of additional frequency components to perceived vowel quality was only reduced by about two-thirds at asymptote (asynchronies greater than 40 ms). Similarly, Hill and Darwin (1996) found that asynchrony reduced, but did not abolish, the influence of a spectrally flanking complex on the perceived lateral position of a harmonic. We should not be surprised that perceptual exclusion can be incomplete, given that the allocation of acoustic elements to separate auditory objects cannot eliminate all interactions between them (e.g., beats, masking). What is needed is an account of why partial pitch and global pitch might fall at opposite ends of this continuum.

D. A proposal

We propose that the differential sensitivity to asynchrony of partial- and global-pitch reflects differences in the way the perceptual properties of pure tones and complex tones are computed. A single component is an indivisible acoustic element. Once it has been mistuned sufficiently to be heard out as a separate tone, it has only to be allocated a pitch. The error in this allocation is small, as indicated by measured pitch shifts which rarely exceed $\pm 3\%$ and are often $\pm 1\%$ or less. These pitch shifts are unlikely to be important biologically, and so there is no imperative for central processes to correct them. Hence, partial-pitch shifts will arise whenever a partial falls close to a slot in a grouping template that has been activated by a set of appropriately patterned components, whether or not that partial has undergone prior allocation to a separate perceptual stream on the basis of asynchrony. In contrast, several components typically contribute to the global pitch of a complex tone. Therefore, a failure to use onset asynchrony to exclude components from the cross-channel process that computes global pitch could lead to significant deviation from a veridical representation.

Relevant to our proposal is the finding by Darwin *et al.* (1990) that a component can be allocated to one auditory stream and yet still contribute to the perceptual properties of another. These authors measured the effect of mistuning a single component on the global pitch of two concurrent harmonic complexes with different $F0$ frequencies, under binaural presentation. They found that a component which was exactly in-tune with one harmonic series, and hence fused with it perceptually, could still contribute to the pitch of the other complex. Darwin *et al.* (1992) presented the concurrent complex tones dichotically with a single mistuned component presented only to one ear. They observed that the mistuned component could even affect the pitch of the complex in the other ear when it was exactly in-tune with the complex in the same ear. Darwin *et al.* (1990, 1992) noted that these findings represent a violation of the Gestalt principle of exclusive allocation (see Bregman, 1990), and may be considered to be a form of duplex perception (Rand, 1974). Moore (1987) came to a similar conclusion regarding the observation that a harmonic mistuned by $\pm 3\%$ makes a full contribution to global pitch (Moore *et al.*, 1985), yet is clearly audible as a separate tone (Moore *et al.*, 1986). Although our results do not indicate a violation of exclusive allocation, there is an important analogy with the findings of Darwin *et al.* (1990, 1992). Specifically, the pitch of the mistuned component is affected by the activation of an internal template that mediates the perceptual fusion of a set of components from which it has *already been segregated* on the basis of asynchrony.

E. Concluding remarks

While de Cheveigné's (1998) harmonic cancellation approach to global-pitch perception has much to commend it, it is clear that partial-pitch shifts cannot easily be subsumed within this framework (as proposed by de Cheveigné, 1999). Instead, we propose that partial-pitch shifts are associated with auditory-grouping processes rather than with the computation of global pitch (see Brunstrom and Roberts, 2000). From this perspective, de Cheveigné's (1997) earlier model of partial-pitch shifts has the advantage that the spacing of the template slots is defined arbitrarily as harmonic, and so could easily be modified to accommodate the grouping of frequency-shifted components. However, the insensitivity of partial-pitch shifts to asynchrony shows that this model would require further modification. This is because a key

assumption is that pitch shifts depend on an increase in the salience of a partial as it is mistuned away from the center frequency of a template slot. If asynchrony has already made a partial highly salient, then mistuning will make little further difference. Therefore, the magnitude of the pitch shifts should fall, rather than remain constant, under asynchronous presentation.

In conclusion, the insensitivity of partial-pitch shifts to asynchrony contrasts with the effectiveness of this cue at excluding acoustic elements from the computation of global pitch and timbre. This contrast probably reflects differences in the perceptual computations required to retrieve the properties of multi-component and single-component entities.

ACKNOWLEDGMENTS

This research was supported by Research Grant 6/S04782 from the Biotechnology and Biological Sciences Research Council (UK) to the second author. We are grateful to Valter Ciocca and Nicolas Grimault for their helpful comments on an earlier draft of this manuscript.

[1]An estimate of the limits of detectability for partial-pitch shifts can be derived from frequency discrimination thresholds (DLFs) for pure tones. Moore and Glasberg (1986) predicted DLFs using an equation based on a summary of frequency-discrimination data presented by Nelson *et al.* (1983). This equation indicates that the DLF at 1 kHz for pure tones at a level similar to that of our components is around 0.35%. Furthermore, the DLF function is fairly flat between 500 Hz and 3 kHz, which encompasses most of the test components. Given that psychometric functions are not infinitely steep, it is possible that partial-pitch shifts even smaller than the DLF might influence listeners' responses.

Bregman, A. S. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge, MA).

Bregman, A. S., and Pinker, S. (**1978**). ''Auditory streaming and the building of timbre,'' Can. J. Psychol. **32**, 19–31.

Brunstrom, J. M., and Roberts, B. (**1998**). ''Profiling the perceptual suppression of partials in periodic complex tones: Further evidence for a harmonic template,'' J. Acoust. Soc. Am. **104**, 3511–3519.

Brunstrom, J. M., and Roberts, B. (**2000**). ''Separate mechanisms govern the selection of spectral components for perceptual fusion and for the computation of global pitch,'' J. Acoust. Soc. Am. **107**, 1566–1577.

Ciocca, V. (**1999**). ''Evidence against an effect of grouping by spectral regularity on the perception of virtual pitch,'' J. Acoust. Soc. Am. **106**, 2746–2751.

Ciocca, V., and Darwin, C. J. (**1993**). ''Effects of onset asynchrony on pitch perception: Adaptation or grouping?'' J. Acoust. Soc. Am. **93**, 2870–2878.

Dannenbring, G. L., and Bregman, A. S. (**1978**). ''Streaming vs. fusion of sinusoidal components of complex tones,'' Percept. Psychophys. **24**, 369–376.

Darwin, C. J. (**1984a**). ''Auditory processing and speech perception,'' in *Attention and Performance X: Control of Language Processes*, edited by H. Bouma and D. G. Bouwhuis (Erlbaum, Hillsdale, NJ), pp. 197–210.

Darwin, C. J. (**1984b**). ''Perceiving vowels in the presence of another sound: Constraints on formant perception,'' J. Acoust. Soc. Am. **76**, 1636–1647.

Darwin, C. J. (**1997**). ''Auditory grouping,'' Trends Cognitive Sci. **1**, 327–333.

Darwin, C. J., and Ciocca, V. (**1992**). ''Grouping in pitch perception: Effects of onset asynchrony and ear of presentation of a mistuned component,'' J. Acoust. Soc. Am. **91**, 3381–3390.

Darwin, C. J., and Sutherland, N. S. (**1984**). ''Grouping frequency components of vowels: When is a harmonic not a harmonic?'' Q. J. Exp. Psychol. A **36**, 193–208.

Darwin, C. J., Buffa, A., and Smits, R. L. H. M. (**1990**). ''Pitch of simultaneous complex sounds with a single mistuned component,'' Proc. Inst. Acoust. **12**, 499–506.

Darwin, C. J., Ciocca, V., and Sandell, G. J. (**1994**). ''Effects of frequency and amplitude modulation on the pitch of a complex tone with a mistuned harmonic,'' J. Acoust. Soc. Am. **95**, 2631–2636.

Darwin, C. J., Buffa, A., Williams, D., and Ciocca, V. (**1992**). ''Pitch of dichotic complex tones with a mistuned frequency component,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford), pp. 223–229.

de Boer, E. (**1976**). ''On the 'residue' and auditory pitch perception,'' in *Handbook of Sensory Physiology, Vol. 5*, edited by W. D. Keidel and W. D. Neff (Springer-Verlag, Berlin), pp. 479–583.

de Cheveigné, A. (**1993**). ''Separation of concurrent harmonic sounds: Fundamental frequency estimation and a time-domain cancellation model of auditory processing,'' J. Acoust. Soc. Am. **93**, 3271–3290.

de Cheveigné, A. (**1997**). ''Harmonic fusion and pitch shifts of mistuned partials,'' J. Acoust. Soc. Am. **102**, 1083–1087.

de Cheveigné, A. (**1998**). ''Cancellation model of pitch perception,'' J. Acoust. Soc. Am. **103**, 1261–1271.

de Cheveigné, A. (**1999**). ''Pitch shifts of mistuned partials: A time-domain model,'' J. Acoust. Soc. Am. **106**, 887–897.

Duifhuis, H., Willems, L. F., and Sluyter, R. J. (**1982**). ''Measurement of pitch in speech: An implementation of Goldstein's theory of pitch perception,'' J. Acoust. Soc. Am. **71**, 1568–1580.

Gerson, A., and Goldstein, J. L. (**1978**). ''Evidence for a general template in central optimal processing for pitch of complex tones,'' J. Acoust. Soc. Am. **63**, 498–510.

Green, D. M., and Dai, H. (**1992**). ''Temporal relations in profile comparisons,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford), pp. 471–478.

Hartmann, W. M., and Doty, S. L. (**1996**). ''On the pitches of the components of a complex tone,'' J. Acoust. Soc. Am. **99**, 567–578.

Hartmann, W. M., McAdams, S., and Smith, B. K. (**1990**). ''Hearing a mistuned harmonic in an otherwise periodic complex tone,'' J. Acoust. Soc. Am. **88**, 1712–1724.

Henke, W. L. (**1990**). MITSYN: A coherent family of high-level languages for time signal processing, software package (Belmont, MA).

Hill, N. I., and Darwin, C. J. (**1996**). ''Lateralization of a perturbed harmonic: Effects of onset asynchrony and mistuning,'' J. Acoust. Soc. Am. **100**, 2352–2364.

Hukin, R. W., and Darwin, C. J. (**1995**). ''Comparison of the effect of onset asynchrony on auditory grouping in pitch matching and vowel identification,'' Percept. Psychophys. **57**, 191–196.

Lea, A. (**1992**). ''Auditory models of vowel perception.'' Doctoral dissertation, University of Nottingham, UK.

Lin, J.-Y., and Hartmann, W. M. (**1998**). ''The pitch of a mistuned harmonic: Evidence for a template model,'' J. Acoust. Soc. Am. **103**, 2608–2617.

Meddis, R., and O'Mard, L. (**1997**). ''A unitary model of pitch perception,'' J. Acoust. Soc. Am. **102**, 1811–1820.

Moore, B. C. J. (**1987**). ''The perception of inharmonic complex tones,'' in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 180–189.

Moore, B. C. J., and Glasberg, B. R. (**1986**). ''The role of frequency selectivity in the perception of loudness, pitch and time,'' in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London), pp. 251–308.

Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (**1985**). ''Relative dominance of individual partials in determining the pitch of complex tones,'' J. Acoust. Soc. Am. **77**, 1853–1860.

Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (**1986**). ''Thresholds for hearing mistuned partials as separate tones in harmonic complexes,'' J. Acoust. Soc. Am. **80**, 479–483.

Nelson, D. A., Stanton, M. E., and Freyman, R. L. (**1983**). ''A general equation describing frequency discrimination as a function of frequency and sensation level,'' J. Acoust. Soc. Am. **73**, 2117–2123.

Patterson, R. D. (**1973**). ''The effects of relative phase and the number of components on residue pitch,'' J. Acoust. Soc. Am. **53**, 1565–1572.

Rakowski, A., and Hirsh, I. J. (**1980**). ''Poststimulatory pitch shifts for pure tones,'' J. Acoust. Soc. Am. **68**, 467–474.

Rand, T. C. (**1974**). ''Dichotic release from masking for speech,'' J. Acoust. Soc. Am. **55**, 678–680.

Ritsma, R. J. (**1967**). ''Frequencies dominant in the perception of the pitch of complex sounds,'' J. Acoust. Soc. Am. **42**, 191–198.

Roberts, B. (**1998**). ''Effects of spectral pattern on the perceptual salience of

partials in harmonic and frequency-shifted complex tones: A performance measure,'' J. Acoust. Soc. Am. **103**, 3588–3596.
Roberts, B., and Bailey, P. J. (**1993a**). ''Spectral pattern and the perceptual fusion of harmonics. I. The role of temporal factors,'' J. Acoust. Soc. Am. **94**, 3153–3164.
Roberts, B., and Bailey, P. J. (**1993b**). ''Spectral pattern and the perceptual fusion of harmonics. II. A special status for added components?'' J. Acoust. Soc. Am. **94**, 3165–3177.
Roberts, B., and Bailey, P. J. (**1996a**). ''Regularity of spectral pattern and its effects on the perceptual fusion of harmonics,'' Percept. Psychophys. **58**, 289–299.
Roberts, B., and Bailey, P. J. (**1996b**). ''Spectral regularity as a factor distinct from harmonic relations in auditory grouping,'' J. Exp. Psychol. **22**, 604–614.
Roberts, B., and Bregman, A. S. (**1991**). ''Effects of the pattern of spectral spacing on the perceptual fusion of harmonics,'' J. Acoust. Soc. Am. **90**, 3050–3060.
Roberts, B., and Brunstrom, J. M. (**1998**). ''Perceptual segregation and pitch shifts of mistuned components in harmonic complexes and in regular inharmonic complexes,'' J. Acoust. Soc. Am. **104**, 2326–2338.
Roberts, B., and Moore, B. C. J. (**1991**). ''The influence of extraneous sounds on the perceptual estimation of first-formant frequency in vowels under conditions of asynchrony,'' J. Acoust. Soc. Am. **89**, 2922–2932.
Schouten, J. F., Ritsma, R. J., and Cardozo, B. L. (**1962**). ''Pitch of the residue,'' J. Acoust. Soc. Am. **34**, 1418–1424.
Summerfield, Q., and Culling, J. F. (**1992**). ''Periodicity of maskers not targets determines ease of perceptual segregation using differences in fundamental frequency,'' J. Acoust. Soc. Am. **92**, 2317(A).
Terhardt, E. (**1971**). ''Pitch shifts of harmonics, an explanation of the octave enlargement phenomenon,'' in *Proceedings of the 7th International Congress on Acoustics, Budapest, Vol. 3*, pp. 621–624.
Terhardt, E., Stoll, G., and Seewann, M. (**1982**). ''Algorithm for extraction of pitch and pitch salience from complex tonal signals,'' J. Acoust. Soc. Am. **71**, 679–688.

Modulation detection interference: Effects of concurrent and sequential streaming

Andrew J. Oxenham[a)]
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Torsten Dau[b)]
Hearing Research Center, Department of Biomedical Engineering, Boston University, 44 Cummington Street, Boston, Massachusetts 02215

(Received 14 September 2000; revised 19 March 2001; accepted 29 March 2001)

The presence of amplitude fluctuations in one frequency region can interfere with our ability to detect similar fluctuations in another (remote) frequency region. This effect is known as modulation detection interference (MDI). Gating the interfering and target sounds asynchronously is known to lead to a reduction in MDI, presumably because the two sounds become perceptually segregated. The first experiment examined the relative effects of carrier and modulator gating asynchrony in producing a release from MDI. The target carrier was a 900-ms, 4.3-kHz sinusoid, modulated in amplitude by a 500-ms, 16-Hz sinusoid, with 200-ms unmodulated fringes preceding and following the modulation. The interferer (masker) was a 1-kHz sinusoid, modulated by a narrowband noise with a 16-Hz bandwidth, centered around 16 Hz. Extending the masker carrier for 200 ms before and after the signal carrier reduced MDI, regardless of whether the target and masker modulators were gated synchronously or were gated with onset and offset asynchronies of 200 ms. Similarly, when the carriers were gated synchronously, asynchronous gating of the modulators did not produce a release from MDI. The second experiment measured MDI with a synchronous target and masker and investigated the effect of adding a series of precursor tones, which were designed to promote the forming of a perceptual stream with the masker, thereby leaving the target perceptually isolated. Four modulated or unmodulated precursor tones presented at the masker frequency were sufficient to completely eliminate MDI. The results support the idea that MDI is due to a perceptual grouping of the masker and target, and show that conditions promoting sufficient perceptual segregation of the masker and target can lead to a total elimination of MDI. [DOI: 10.1121/1.1373443]

PACS numbers: 43.66.Dc, 43.66.Mk [SPB]

I. INTRODUCTION

There is mounting evidence from experimental and theoretical studies that many aspects of the perception of amplitude modulation (AM) can be modeled in terms of a bank of linear, overlapping bandpass filters in the modulation domain (Bacon and Grantham, 1989; Houtgast, 1989; Dau *et al.*, 1997a, 1997b; Lorenzi *et al.*, 1997; Dau *et al.*, 1999; Ewert and Dau, 2000). So far, models of this process have assumed that the output of each peripheral filter is processed independently so that, in effect, there is a separate and independent modulation filterbank at the output of each peripheral auditory filter (Dau *et al.*, 1997a, 1997b). In contrast, there is substantial evidence that AM within each peripheral channel is not always perceived independently and that instead significant interference, or crosstalk, can occur. This effect, known as modulation detection (or discrimination) interference (MDI), is fairly robust and occurs over very large carrier frequency separations, even when the masker and signal carriers are presented to opposite ears (Yost and Sheft, 1989; Yost *et al.*, 1989; Bacon and Opie, 1994).

The dependence of signal modulation thresholds on masker modulation frequency and bandwidth in MDI is similar to that found when the masker and signal both modulate a single carrier (Verhey and Dau, 2000). This single-carrier case is known as modulation masking (Bacon and Grantham, 1989; Houtgast, 1989). That fact that the same modulation frequency-specific tuning is observed in both MDI and modulation masking suggests that at some level, both effects are reflections of similar underlying mechanisms. In terms of modulation filters, this could be implemented in at least two ways: either the modulation filters receive inputs from several peripheral channels, and are therefore broadly tuned with respect to carrier frequency, or the outputs from modulation filters tuned to different carrier frequencies, but to similar modulation frequencies, converge at a higher level of processing (Yost *et al.*, 1989). Recordings from cat inferior colliculus have indicated sharp tuning to carrier frequency in modulation-selective neurons (Langner and Schreiner, 1988), whereas modulation-sensitive neurons in the gerbil auditory cortex have been found that are broadly tuned with respect to carrier frequency (Schulze and Langner, 1997, 1999). This

[a)]Electronic mail: oxenham@mit.edu
[b)]Present address: Arbeitsgruppe Medizinische Physik, Universität Oldenburg, 26111 Oldenburg, Germany; electronic mail: torsten.dau@medi.physik.uni-oldenburg.de

could be interpreted in terms of convergence across carrier frequencies at higher levels of modulation analysis.

Psychophysical data from Hall and Grose (1991) argue against a simple ''hard-wired'' implementation of modulation filterbanks with broad tuning to carrier frequency. They found that the effect of MDI was greatly reduced if the masker was gated on before (and gated off after) the target. As pointed out by Hall and Grose (1991), this finding supports the idea that MDI occurs when the two carriers are grouped to form one perceptual object. Asynchronous gating, by as little as 100 ms, can lead to a perceptual segregation of the two carriers, thereby facilitating separate analysis of the modulation of each carrier. The reduction in MDI due to asynchronous gating has since been confirmed in many studies, although all have found some residual MDI, even when the masker is presented continuously (Moore and Jorasz, 1992; Moore and Shailer, 1992; Mendoza *et al.*, 1995a, 1995b).

So far, studies measuring the effect of masker asynchrony on MDI have used stimuli that were modulated throughout their duration (Hall and Grose, 1991; Moore and Shailer, 1992; Mendoza *et al.*, 1995a, 1995b). Because of this, it is not clear whether the release from MDI is due to asynchronous gating of the carriers, the modulators, or a combination of both.[1] It is possible, for instance, that gating on the masker modulation before the target modulation could facilitate segregation even if the two carriers are gated synchronously. Conversely, gating the masker and target modulation on synchronously could conceivably lead to MDI even when the carriers are gated asynchronously. The paradigm used in the original MDI experiment of Yost and Sheft (1989) employed synchronously gated 1000-ms carriers and 500-ms target modulation, temporally centered within 1000-ms masker modulation. The fact that Yost and Sheft observed significant MDI suggests that modulation gating asynchrony is not sufficient to eliminate MDI. However, it is not known whether greater MDI would have been observed in their listeners if the modulators had been gated synchronously. To our knowledge, no one has examined the effect of gating the carriers asynchronously while gating the modulators synchronously. Our first experiment attempts to evaluate the roles of carrier and modulator gating asynchrony in the release from MDI by independently manipulating onset/offset relationships between the carriers and the modulators.

Auditory grouping mechanisms are known to extend over considerable time periods, such that nonsimultaneous sounds can be combined to form perceptual streams (such as melodies or running speech) based on parameters such as spectral similarity, spatial location, or fundamental frequency (Bregman and Campbell, 1971; van Noorden, 1975; Hartmann and Johnson, 1991; Vliegen and Oxenham, 1999). These sequential cues can sometimes override the grouping cues associated with concurrent sounds, such as common onset, spectral proximity, and harmonicity. A striking example of this is when a component within a harmonic complex is ''captured'' by a preceding sequence of repeated tones at the component frequency: not only is the component heard separately from the remaining components in the harmonic complex, but its contribution to the overall pitch and timbre of the complex is also greatly reduced (Darwin *et al.*, 1989; Darwin *et al.*, 1995), suggesting that at least part of the segregation occurs before the formation of higher-level percepts, such as pitch and timbre. Our second experiment investigates whether such sequential perceptual streaming can also lead to a release from MDI.

II. EFFECTS OF CARRIER AND MODULATOR ASYNCHRONY ON MDI

A. Stimuli

The target carrier was a 4.3-kHz sinusoid, presented at a level of 65 dB SPL. Both the carriers and the modulators were gated on and off with 50-ms raised-cosine ramps. The signal was 16-Hz sinusoidal amplitude modulation imposed on the target carrier, with a random starting phase in each trial and a total duration of 500 ms. The duration of the target carrier was 900 ms, providing 200-ms unmodulated fringes before and after the signal modulation. The masker carrier was a 1-kHz sinusoid, presented at a level of 65 dB SPL. The masker modulation was a narrowband Gaussian noise, set at an rms level of -10 dB (*re* unity), with a center frequency of 16 Hz and cutoff frequencies of 8 and 24 Hz. We used a random-noise modulator, as it has been shown that this generally produces more MDI than a sinusoidal modulator (Mendoza *et al.*, 1995a), and because our pilot impressions indicated that practice effects may be less pronounced than when using sinusoidal masker modulation.[2] A new sample of noise was generated for each interval of the experiment. In conditions where the carriers were gated synchronously, the masker carrier duration was also 900 ms. In conditions where the carriers were gated asynchronously, the masker carrier duration was increased by 400 ms, so that the masker carrier extended beyond the signal carrier by 200 ms at either end. The same procedure was used for the masker modulation, so that it was either synchronous with the signal modulation or extended beyond the signal modulation by 200 ms at either end. The conditions tested in this experiment are illustrated in Fig. 1. In the first condition, the target was presented alone (TA). In the following four conditions, the carriers (C) and the modulators (M) of the target and masker were gated either synchronously (s) or asynchronously (a). In the final condition, the masker carrier was gated synchronously with the target, but was unmodulated (CsMu). Each interval in a trial was separated by an interstimulus interval (ISI) of 300 ms. The stimuli were generated digitally and presented at a sampling rate of 32 kHz via a LynxStudio LynxOne DAC. The output of the DAC was passed through a programmable attenuator (TDT PA4) and headphone buffer (TDT HB6), and was delivered diotically to Sennheiser HD 580 headphones.

B. Procedure

An adaptive three-interval 3AFC procedure was used in conjunction with a 2-down 1-up tracking rule to estimate the 70.7% correct point on the psychometric function. The intervals were marked on a computer monitor and feedback was provided after each trial. Listeners responded via the computer keyboard or mouse. At the start of a run, the signal

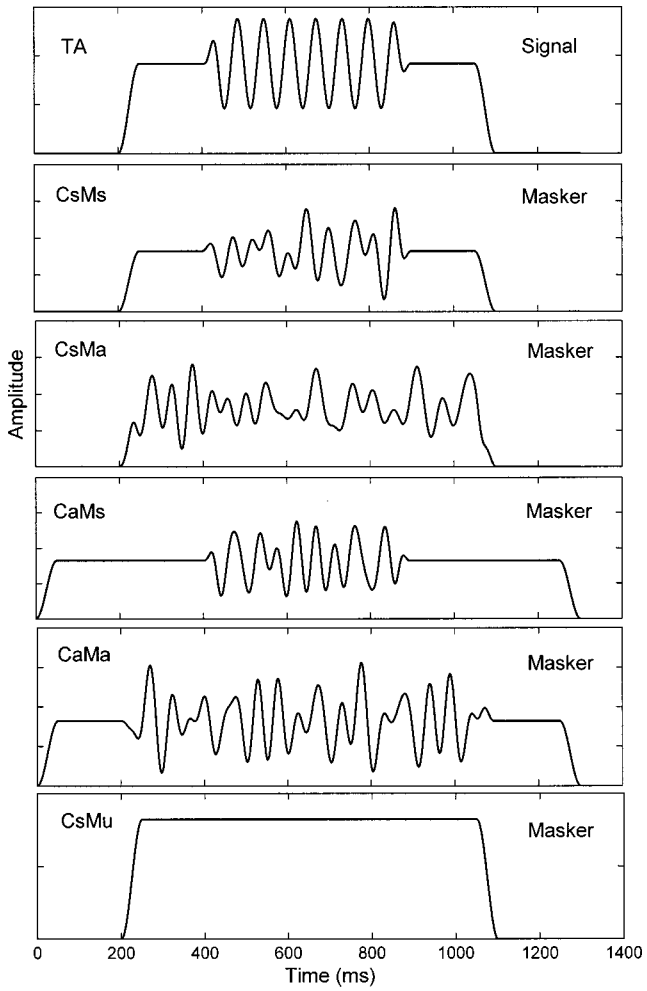


FIG. 1. Schematic diagram of the stimulus envelopes used in experiment 1. The upper panel shows the envelope of the target (TA). The remaining five panels show the envelope of the masker for the four combinations of carrier (C) and modulator (M) gated synchronously (s) or asynchronously (a) with the target, and for the unmodulated masker (CsMu). The ordinate scale is arbitrary.

modulation (m) was set to -6 dB. The initial step size was 4 dB, which was reduced to 1 dB after four reversals. The run was terminated after a further six reversals, and threshold was defined as the mean of the levels at the last six reversals. The conditions were presented in random order, such that all six conditions were tested before any were repeated. New random orders were generated for each repetition and each listener. Four threshold estimates were obtained from each listener in each condition.

C. Subjects

Eight listeners, of which five were female, participated as paid subjects in this experiment. All had thresholds at or lower than 15 dB HL at octave frequencies between 250 and 8000 Hz and none reported any history of hearing difficulties or disorders. Their ages ranged from 19 to 29 years (mean age of 23) and all but two were college students. All listeners received at least 4 h training before data were collected.

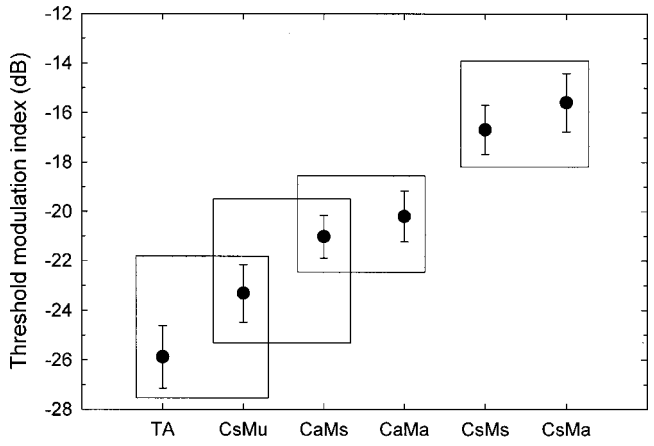


FIG. 2. Mean results from experiment 1. Thresholds are shown for the target carrier alone (TA), for unmodulated interferer (CsMu) gated synchronously with the target carrier and for the four combinations of carriers (C) and modulators (M) gated synchronously (s) or asynchronously (a). Error bars represent ± 1 standard error of the mean. Symbols that fall within the same box are not statistically different from each other ($p > 0.05$) using an all-pairs Bonferroni *post hoc* comparison.

D. Results and discussion

Figure 2 shows the mean data from the six conditions. The error bars denote ± 1 standard error of the mean. The trends observed in the mean data are representative of the individual data, although there were considerable differences in overall sensitivity to modulation. The leftmost condition represents the mean threshold for the target alone (TA). The next condition (CsMu) shows the threshold when the masker and signal carriers were gated synchronously, but the masker was unmodulated. The remaining four conditions show the combinations of having the carriers (C) or modulators (M) gated synchronously (s) or asynchronously (a). The two rightmost conditions represent classic MDI, with the carriers gated synchronously (CsMs and CsMa). The maximum amount of MDI (about 7 dB when compared with the unmodulated masker, or 10 dB when compared with the target alone) is in good agreement with data from the literature for comparable levels of noise masker modulation (Mendoza *et al.*, 1995a; Moore *et al.*, 1995).

The effects of modulator and carrier gating synchrony were investigated with a two-way within-subjects ANOVA, using only the four conditions from Fig. 2 in which the masker was modulated. The effect of carrier gating synchrony was found to be highly significant ($F_{1,7} = 27.8$, $p = 0.001$), while the effect of modulator gating synchrony was not ($F_{1,7} = 3.34$, $p = 0.1$). There was no significant interaction between the two factors ($F_{1,7} < 1$, NS). Bonferroni *post hoc* paired comparisons were also made across all six conditions. The results of these comparisons are shown in Fig. 2 as boxes. Conditions that lie within the same box are not significantly different at the level of $p = 0.05$. Another, less conservative test, Fisher's least significant difference (LSD) test, as used in a recent study of frequency MDI (FMDI; Gockel and Carlyon, 2000), produces more significant differences, with only the CaMa/CaMs and CsMa/CsMs differences remaining not significant at the level of $p = 0.05$. However, all tests, including the ANOVA, confirm that there

was no significant effect of modulator gating synchrony.

Gating the carriers asynchronously produces a release from MDI, regardless of the modulator gating synchrony. However, some interference is still present, at least when compared with thresholds for the target alone (TA). This is consistent with other studies investigating the role of asynchronous gating (Moore and Jorasz, 1992; Moore and Shailer, 1992; Mendoza *et al.*, 1995a). This remaining interference was the topic of a recent investigation of FMDI by Gockel and Carlyon (2000). They found that the interference was significant according to Fisher's LSD test, and that it persisted even when the interfering carrier was switched off during the presentation of the target carrier, suggesting that something other than simultaneous modulation was responsible for the residual interference. Their findings are addressed further in the discussion of experiment 2.

Modulator gating asynchrony had no significant effect on thresholds, regardless of whether the carriers were synchronous or not. It seems that listeners cannot use a modulation onset/offset difference of 200 ms to help segregate the two carriers in order to perform modulation detection. Conversely, simultaneous onset of masker and signal modulation does not lead to more MDI if the carriers are gated asynchronously. Although we found no effect of modulator asynchrony, we cannot rule out the possibility that modulation fringes longer than the 200 ms employed in the present study may lead to some release from MDI.

III. RELEASE FROM MDI IN CONDITIONS PROMOTING SEQUENTIAL STREAMING

The second experiment examines the effect of sequential streaming on MDI. The question was whether such streaming, which must be based on longer-term analysis, is sufficient to produce a release from MDI, allowing independent analysis of the modulation on the two carriers. The paradigm used in this experiment is similar to that employed by Darwin and colleagues (Darwin *et al.*, 1989, 1995; Darwin and Hukin, 1997). In some conditions, the test stimuli (masker and target carrier) were preceded by a sequence of four tones at the masker carrier frequency. The assumption was that the preceding tones would form a perceptual stream with the masker carrier, leaving the target carrier perceptually isolated.

A. Method

The target carrier was a 4.3-kHz sinusoid, with a total duration of 187.5 ms, gated with 20-ms raised cosine ramps. The masker carrier was a 1-kHz sinusoid which, when present, was always gated synchronously with the target carrier. The signal modulation was a 16-Hz sinusoid with a random starting phase, which was present throughout the target carrier. The masker modulation was again a Gaussian noise with a center frequency of 16 Hz, a bandwidth of 16 Hz, and an rms level of −10 dB (re. unity), which was present throughout the masker carrier. The four precursor tones were all the same duration as the target and masker and were separated by gaps of 62.5 ms, giving an overall repetition period of 250 ms. Six conditions were tested, as illustrated in Fig. 3: (1) Target alone (TA); (2) target with unmodulated masker (Mu); (3) target with modulated masker (Mm; classical MDI condition); (4) target and modulated masker preceded by four precursors at the masker frequency and modulated with the same Gaussian statistics as the masker (PSm); (5) target and modulated masker preceded by four unmodulated precursors at the masker frequency (PSu); (6) target and modulated masker preceded by four modulated precursors with a carrier frequency of 1750 Hz (PDm).

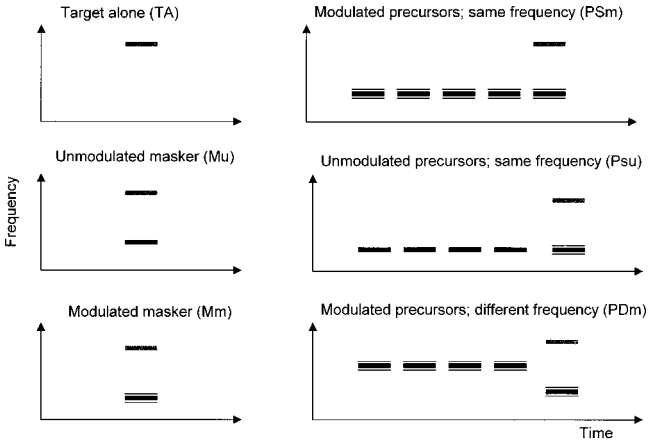


FIG. 3. Schematic diagram of the conditions tested in experiment 2, drawn in the form of a spectrogram. The masker carrier was at 1 kHz, while the signal carrier was at 4.3 kHz.

Condition 5 (PSu) was run in order to examine the effect of modulating the precursors. It may be that the sequential binding between the masker and the precursors is reduced if only the masker is modulated. In condition 6 (PDm), the precursor frequency was sufficiently remote from both the masker and the target to make any sequential binding unlikely. Thus, this condition served as a control for any general effect of preceding the target stimulus with modulated precursors.

The procedures and the listeners were the same as in experiment 1. As all listeners had completed experiment 1 before embarking on this experiment, they already had substantial practice in MDI tasks.

B. Results and discussion

The mean results from this experiment are shown in Fig. 4. The mean pattern of results is representative of the individual data, but again the differences between subjects in absolute values were substantial. Thresholds are generally higher than in experiment 1, which is due to the shorter duration of the target in this experiment (Sheft and Yost, 1990). As in experiment 1, there is a difference of about 3 dB in thresholds between the target alone (TA) and the unmodulated masker (Mu) conditions. As expected, gating a modulated masker with the target (Mm) results in substantial MDI. However, MDI is totally eliminated by introducing precursors at the masker frequency, whether they are modulated (PSm) or not (PSu). When the precursors are at a different carrier frequency (PDm), MDI is substantially restored. A Bonferroni *post hoc* test in an all-pairs comparison confirmed these impressions: the Mm and PDm conditions were not significantly different from each other, but they

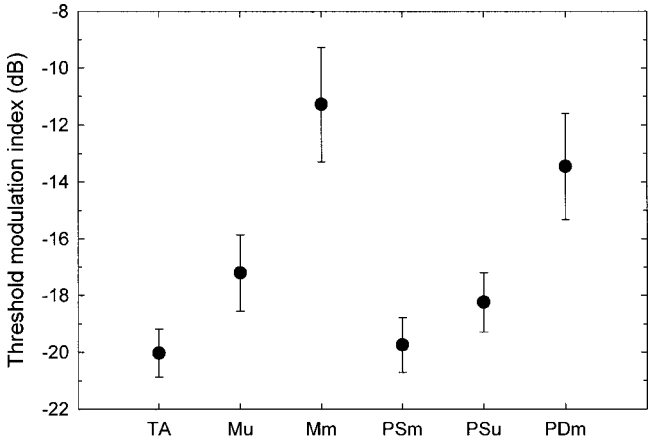


FIG. 4. Mean results from experiment 2. From left to right, thresholds are shown for the target alone (TA), the target with the unmodulated masker (Mu), the target with the modulated masker (Mm), the target and modulated masker with modulated precursors at the same frequency as the masker frequency (PSm), the target and modulated masker with unmodulated precursors at the same frequency as the masker frequency (PSu), and target and modulated masker with modulated precursors at a different frequency (PDm). Error bars represent ±1 standard error of the mean.

were both significantly different from all the other conditions at the level of $p=0.05$. There were no other significant differences, meaning that the conditions with the precursors at the same frequency as the masker (PSm and PSu) were not significantly different from the target alone condition (TA). In this experiment a reanalysis using Fisher's LSD test resulted in essentially the same conclusions: the Mm and PDm conditions were not significantly different from each other but were significantly different from all other conditions, and the PSm and PSu conditions were not significantly different from the TA condition at the level of $p=0.05$. One difference between the tests was that the LSD test detected a significant difference between the TA and the Mu conditions, as has also been reported in previous studies (e.g., Moore and Shailer, 1992).

To our knowledge, this is the first study to explicitly examine the effects of sequential streaming on MDI. However, one condition studied by Moore and Shailer (1992) has some similarities to those studied here. They used a preceding unmodulated ''cue'' tone at the target carrier frequency to help listeners concentrate on the target and ignore the masker. This manipulation had very little effect on MDI. The main differences between their experiment and ours are the frequency of the cue (masker frequency in ours and target frequency in theirs); the gap between the cue and the test stimulus (62.5 ms in ours and 300 ms in theirs); and the number of cue tones (four in ours and one in theirs). We speculate that the most important parameter may be the number of cue tones. Increasing the number of repetitions generally leads to an increased tendency for a separate perceptual stream to be formed. The single cue tone used by Moore and Shailer may not have perceptually separated the target from the masker sufficiently to reduce MDI.

If the results from a frequency MDI experiment can be compared with those from an amplitude MDI experiment, our results appear to be in conflict with those of Gockel and Carlyon (2000).[3] They found that nonsimultaneous modulated maskers (one presented before and one after the target), centered around 2300 Hz, could interfere with FM detection for a 1000-Hz target, whereas our experiment 2 shows that the addition of nonsimultaneous modulated maskers can eliminate MDI in the presence of a simultaneous masker. In the framework of the present study, the results of Gockel and Carlyon can be understood in terms of the maskers and target forming one perceptual stream and the masker modulation within that stream interfering with the detection of the target modulation, even though they do not physically overlap in time. Gockel and Carlyon found that increasing the gaps between the maskers and the target had little or no effect on the amount of FMDI measured. They argued that the longest gaps should have resulted in the three tones being clearly perceived as separate auditory events, thereby presumably ruling out effects of concurrent, or quasi-concurrent, grouping in assigning masker modulation to the target. However, the longest gaps they employed were only 200 ms. While three auditory events (masker–target–masker) would have been clearly perceived, it is possible that they were perceived as belonging to one auditory stream: the segregation of streams based on frequency differences is known to take time to build up (Bregman, 1978), such that the first few repetitions of a tone sequence are often perceived as a single stream, even for large frequency differences. It is therefore possible that the FMDI observed by Gockel and Carlyon was a result of the maskers and target forming a single auditory stream, thereby making the task of analyzing the target modulation separately from masker modulation more difficult. The present results suggest that the interference observed by Gockel and Carlyon (2000) might be eliminated simply by adding more masker bursts prior to the target, thereby encouraging the perceptual segregation of the masker bursts from the target.

The finding of Gockel and Carlyon (2000) that nonsimultaneous maskers can produce some MDI, in combination with our finding of a total release from MDI in a sequential streaming paradigm, suggests that modulation analysis may be performed not only across concurrent sounds, forming auditory objects, but also across sequential sounds, forming auditory streams. This intriguing possibility has yet to be investigated. However, the explanation of interference based on perceptual stream formation is different from the one offered by Gockel and Carlyon (2000). Their explanation relates to the perceptual salience of modulated versus unmodulated tones, rather than to issues of perceptual grouping. For this reason, their explanation would not predict a release from MDI due to additional sequential masking tones.

The results of the present study show that stimuli designed to encourage sequential streaming can lead to a total elimination of MDI. The reduction in MDI found in this experiment therefore seems greater than that produced by introducing carrier gating asynchronies (experiment 1) where, consistent with previous results, some residual MDI remained. This could be interpreted as evidence that the perceptual segregation produced by a preceding sequence of tones is more complete than that produced by an onset/offset gating asynchrony. The results also argue against the conclusion of Moore and Shailer (1992) that factors other than per-

ceptual grouping may be partly responsible for MDI. If their hypothesis were correct in all circumstances, some residual MDI would have been expected in our sequential tasks also. However, there are some circumstances, such as when the carriers are close in frequency, in which MDI can be attributed in part to within-channel processing (Bacon and Konrad, 1993). In these cases, it is unlikely that even total stream segregation would eliminate MDI.

While the present results emphasize the role of perceptual grouping mechanisms in MDI, it is important to remember that MDI exhibits the same modulation-frequency specificity as modulation masking on a single carrier (Verhey and Dau, 2000). This suggests that analysis of both modulation masking and MDI in terms of the same underlying mechanisms, such as modulation filters, may still be appropriate. Nevertheless, it is clear that such analysis takes place only after some degree of perceptual grouping has occurred and that MDI cannot be modeled in terms of ''hard-wired'' modulation crosstalk across peripheral frequency channels.

IV. SUMMARY

Two experiments investigated the role of perceptual organization on modulation detection interference (MDI). The first experiment independently manipulated the onset/offset asynchrony of the carriers and the modulators. An onset/offset asynchrony between the masker and target carriers led to a reduction in MDI; this reduction was the same whether the onset and offset of the modulators were synchronous or asynchronous. Similarly, when the carriers were gated synchronously, gating the modulators asynchronously produced no release from MDI.

The second experiment measured MDI for a synchronously gated masker and target in the presence of precursor tones, which were designed to form a perceptual stream with the masker and thus promote the perceptual segregation of the target. Introducing a sequence of four modulated or unmodulated precursor tones at the masker frequency led to a complete elimination of MDI. In contrast, presenting the precursors at a frequency remote from both the target and the masker produced no significant reduction in MDI. The results support the idea that MDI results from an inability to separate the target modulation from that of the masker because the two carriers are to some extent perceptually grouped. The finding that precursor tones can totally eliminate MDI, while an onset/offset carrier asynchrony produces a smaller reduction in MDI, suggests that segregation may be graded and that even a large onset/offset asynchrony may not result in the same degree of perceptual segregation (at least in terms of modulation analysis) as that produced by a sequence of precursor tones.

ACKNOWLEDGMENTS

This work was supported by NIH/NIDCD grant R01 DC 03909 (A.J.O.) and by the Max Kade Foundation and the Deutsche Forschungsgemeinschaft (DFG) (T.D.). Stephan Ewert provided help with programming. Sid Bacon, Hedwig Gockel, and Lee Mendoza provided helpful comments on a previous version of this paper.

[1]Moore and Jorasz (1992) tested AM depth discrimination in the presence of an interfering modulator and gated the increment or decrement in modulation depth separately from the interfering modulation. However, as both the reference and interfering modulation were still gated on and off together, this manipulation would have been unlikely to lead to segregation of the target modulation.

[2]Due to its random nature there is a finite probability of overmodulation (i.e., an instantaneous value of the modulator exceeding unity) when using a noise modulator. However, at −10 dB rms, the average probability is small (≈0.0015) and is therefore unlikely to affect the results. Furthermore, Mendoza *et al.* (1995a) found that much higher levels of overmodulation produced MDI results that were very similar to those produced with little or no overmodulation, suggesting that no qualitative differences result from overmodulation.

[3]Gockel and Carlyon (2000) used a modulation frequency of 15 Hz. Based on the results of Moore and Sek (1995), it is likely that at this modulation frequency listeners are detecting FM based on AM within individual auditory filters, thereby supporting a direct comparison between amplitude and frequency MDI.

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.

Bacon, S. P., and Opie, J. M. (**1994**) ''Monotic and dichotic modulation detection interference in practiced and unpracticed subjects,'' J. Acoust. Soc. Am. **95**, 2637–2641.

Bregman, A. S. (**1978**) ''Auditory streaming is cumulative,'' J. Exp. Psychol. **4**, 380–387.

Bregman, A. S., and Campbell, J. (**1971**) ''Primary auditory stream segregation and perception of order in rapid sequences of tones,'' J. Exp. Psychol. Hum. Percept. Perform. **89**, 244–249.

Darwin, C. J., and Hukin, R. W. (**1997**) ''Perceptual segregation of a harmonic from a vowel by interaural time difference and frequency proximity,'' J. Acoust. Soc. Am. **102**, 2316–2324.

Darwin, C. J., Hukin, R. W., and al-Khatib, B. Y. (**1995**) ''Grouping in pitch perception: Evidence for sequential constraints,'' J. Acoust. Soc. Am. **98**, 880–885.

Darwin, C. J., Pattison, H., and Gardner, R. B. (**1989**) ''Vowel quality changes produced by surrounding tone sequences,'' Percept. Psychophys. **45**, 333–342.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997a**) ''Modeling auditory processing of amplitude modulation. I. Detection and masking with narrowband carriers,'' J. Acoust. Soc. Am. **102**, 2892–2905.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997b**) ''Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration,'' J. Acoust. Soc. Am. **102**, 2906–2919.

Dau, T., Verhey, J., and Kohlrausch, A. (**1999**) ''Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers,'' J. Acoust. Soc. Am. **106**, 2752–2760.

Ewert, S. D., and Dau, T. (**2000**) ''Characterizing frequency selectivity for envelope fluctuations,'' J. Acoust. Soc. Am. **108**, 1181–1196.

Gockel, H., and Carlyon, R. P. (**2000**) ''Frequency modulation detection interference produced by asynchronous and nonsimultaneous interferers,'' J. Acoust. Soc. Am. **108**, 2329–2336.

Hall, 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.

Hartmann, W. M., and Johnson, D. (**1991**) ''Stream segregation and peripheral channeling,'' Music Percept. **9**, 155–184.

Houtgast, T. (**1989**) ''Frequency selectivity in amplitude-modulation detection,'' J. Acoust. Soc. Am. **85**, 1676–1680.

Langner, G., and Schreiner, C. E. (**1988**) ''Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms,'' J. Neurophysiol. **60**, 1799–1822.

Lorenzi, C., Micheyl, C., Berthommier, F., and Portalier, S. (**1997**) ''Modulation masking in listeners with sensorineural hearing loss,'' J. Speech Lang. Hear. Res. **40**, 200–207.
Mendoza, L., Hall, J. W., and Grose, J. H. (**1995a**) ''Modulation detection interference using random and sinusoidal amplitude modulation,'' J. Acoust. Soc. Am. **97**, 2487–2497.
Mendoza, L., Hall, J. W., and Grose, J. H. (**1995b**) ''Within- and across-channel processes in modulation detection interference,'' J. Acoust. Soc. Am. **97**, 3072–3079.
Moore, B. C. J., and Jorasz, U. (**1992**) ''Detection of changes in modulation depth of a target sound in the presence of other modulated sounds,'' J. Acoust. Soc. Am. **91**, 1051–1061.
Moore, B. C. J., and Sek, A. (**1995**) ''Effects of carrier frequency, modulation rate, and modulation waveform on the detection of modulation and the discrimination of modulation type (amplitude modulation versus frequency modulation),'' J. Acoust. Soc. Am. **97**, 2468–2478.
Moore, B. C. J., Sek, A., and Shailer, M. J. (**1995**) ''Modulation discrimination interference for narrow-band noise modulators,'' J. Acoust. Soc. Am. **97**, 2493–2497.
Moore, B. C. J., and Shailer, M. J. (**1992**) ''Modulation discrimination interference and auditory grouping,'' Philos. Trans. R. Soc. London, Ser. B **336**, 339–346.
Schulze, H., and Langner, G. (**1997**) ''Periodicity coding in the primary auditory cortex of the Mongolian gerbil (Meriones unguiculatus): Two different coding strategies for pitch and rhythm,'' J. Comp. Physiol. A **181**, 651–663.
Schulze, H., and Langner, G. (**1999**) ''Auditory cortical responses to amplitude modulations with spectra above frequency receptive fields: Evidence for wide spectral integration,'' J. Comp. Physiol. A **185**, 493–508.
Sheft, S., and Yost, W. A. (**1990**) ''Temporal integration in amplitude modulation detection,'' J. Acoust. Soc. Am. **88**, 796–805.
van Noorden, L. P. A. S., (**1975**) ''Temporal coherence in the perception of tone sequences,'' Ph.D. thesis, Eindhoven University of Technology.
Verhey, J., and Dau, T. (**2000**) ''Modulation detection interference with narrow-band noise interferers,'' J. Acoust. Soc. Am. **107**, 2915.
Vliegen, J., and Oxenham, A. J. (**1999**) ''Sequential stream segregation in the absence of spectral cues,'' J. Acoust. Soc. Am. **105**, 339–346.
Yost, W. A., and Sheft, S. (**1989**) ''Across-critical-band processing of amplitude-modulated tones,'' J. Acoust. Soc. Am. **85**, 848–857.
Yost, W. A., Sheft, S., and Opie, J. (**1989**) ''Modulation interference in detection and discrimination of amplitude modulation,'' J. Acoust. Soc. Am. **86**, 2138–2147.

Formant-frequency matching between sounds with different bandwidths and on different fundamental frequencies

Pascal Dissard and C. J. Darwin[a)]
Experimental Psychology, University of Sussex, Brighton BN1 9QG, United Kingdom

(Received 28 April 2000; accepted for publication 10 April 2001)

The two experiments described here use a formant-matching task to investigate what abstract representations of sound are available to listeners. The first experiment examines how veridically and reliably listeners can adjust the formant frequency of a single-formant sound to match the timbre of a target single-formant sound that has a different bandwidth and either the same or a different fundamental frequency ($F0$). Comparison with previous results [Dissard and Darwin, J. Acoust. Soc. Am. **106**, 960–969 (2000)] shows that (i) for sounds on the same $F0$, introducing a difference in bandwidth increases the variability of matches regardless of whether the harmonics close to the formant are resolved or unresolved; (ii) for sounds on different $F0$'s, introducing a difference in bandwidth only increases variability for sounds that have unresolved harmonics close to the formant. The second experiment shows that match variability for sounds differing in $F0$, but with the same bandwidth and with resolved harmonics near the formant peak, is not influenced by the harmonic spacing or by the alignment of harmonics with the formant peak. Overall, these results indicate that match variability increases when the match cannot be made on the basis of the excitation pattern, but match variability does not appear to depend on whether ideal matching performance requires simply interpolation of a spectral envelope or also the extraction of the envelope's peak frequency. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1379085]

PACS numbers: 43.66.Jh, 43.71.Es [RVS]

I. INTRODUCTION

In the experiments reported here, listeners match the timbre of two single-formant sounds by adjusting the formant frequency of one of the sounds. We ask whether, by varying the other dimensions along which the sounds differ, we can force listeners to make judgments based on different levels of abstraction of sound.

In an earlier paper (Expt. 2, Dissard and Darwin, 2000) we examined how reliably (as measured by the standard deviation of an individual listener's matches) listeners could match the timbre of a single-formant target sound by adjusting the formant frequency of a second (''match'') sound. We showed (unsurprisingly) that when the target and the adjustable sound had the same fundamental frequency ($F0$), matches were more reliable than when they differed in fundamental frequency. However, we also showed that this difference was substantially larger for sounds on high $F0$'s than it was for sounds on low $F0$'s.

When the $F0$ of both target and match sounds was the same, matches were both veridical (in that the match formant frequency was close to the target formant frequency) and reliable (across trials for a particular listener). Listeners here are performing a match that is based on making the two sounds identical, and so the match could be made on the basis of making either the total neural activity from the two sounds identical, or some subset of the activity such as that corresponding to the excitation pattern (Moore and Glasberg, 1983).

When the $F0$'s of the target and match sounds were different, however, the strategies open to listeners differed depending on whether the harmonics in the formant region were resolved by the auditory system or unresolved. With unresolved harmonics in the formant region, listeners can make a reliable and veridical match on the basis of identity of the auditory excitation pattern. With resolved harmonics, however, harmonic ripple in the excitation pattern disrupts formant matches, encouraging listeners to make the match on the basis of a more abstract property such as an interpolated envelope of the excitation pattern, which would smooth out harmonic ripple (for excitation patterns, see Fig. 1). We found that giving target and match sounds a different $F0$ changed the reliability of matches in a way that reflected these two different strategies. For sounds with unresolved harmonics in the formant region, matches were slightly less reliable than when $F0$ was the same on both sounds, but for sounds with resolved harmonics in the formant region, matches were much less reliable than when the $F0$ was the same on both sounds.

We argued against the possibility that this latter result was due to listeners trying to match excitation patterns from sounds with different $F0$'s. We modeled performance that minimized the rms error between excitation patterns and showed that this strategy predicted a much larger increase in variability than we actually found. We concluded that the increased variability of matches on different $F0$'s for sounds with resolved harmonics in the region of their single formant reflected the perceptual cost of interpolating an envelope across harmonic peaks in the excitation pattern. This claim—that matches made at more abstract levels are more variable—forms the basis for the present paper.

In the present paper we present new experiments in

a)Author to whom correspondence should be addressed.

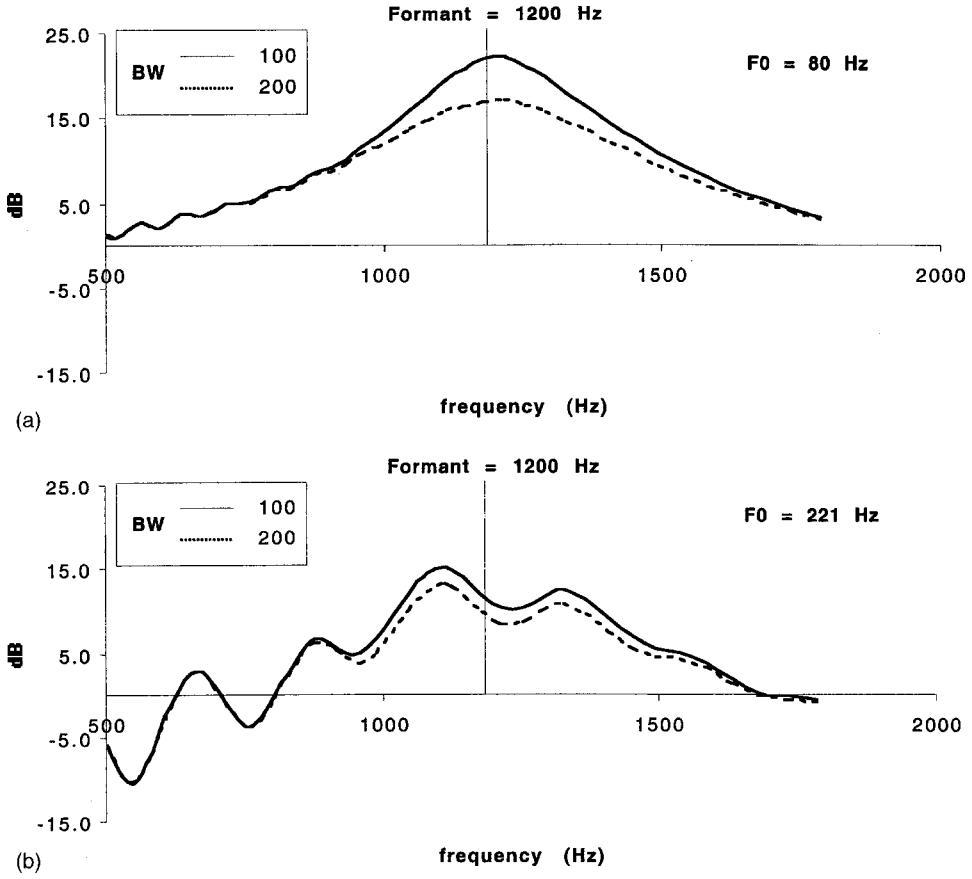


FIG. 1. Excitation patterns of single-formant (1200 Hz) sounds with different formant bandwidths. The upper panel shows sounds with different bandwidths on a fundamental frequency of 80 Hz. The lower panel shows similar sounds with a fundamental frequency of 212 Hz.

which we also manipulate formant bandwidth (BW). The point of this additional manipulation is that it could force listeners to use a more abstract representation than the interpolated spectral envelope—such as the formant frequency—even in the absence of a difference in $F0$.[1]

The rationale behind the manipulations made in this and our previous experiments can be understood by reference to Table I. Here, we show the stimulus conditions used in the present experiment 1 and in our previous paper (Expt. 2, Dissard and Darwin, 2000). The left-hand column provides acronyms for the particular stimulus conditions that we will refer to later (the letters refer to the contents of the next three columns, that is whether the target and match sounds are the same or different in harmonic resolution, $F0$ and BW, respectively). The second column indicates whether the harmonics near to the formant peak is resolved or unresolved by the normal ear. This property is illustrated in Fig. 1, which shows excitation patterns (Moore and Glasberg, 1983) for two single-formant sounds differing in bandwidth, with both sounds on either a low (80 Hz) or a high (212 Hz) $F0$. For sounds on a high $F0$, the excitation pattern resolves individual harmonic components, whereas for sounds on a low $F0$ this ripple is absent, indicating that the individual harmonics are not resolved. The third and fourth columns indicate whether the two sounds in a trial (the target and the match sound) share the same or have different $F0$'s and bandwidths. Finally, the fifth column describes the least abstract property that listeners could use to ensure a reliable and veridical (i.e., same formant frequency) match. The table aims to provide a hierarchy of abstraction for matches, without necessarily committing the reader to specific abstractions.[2]

When target and match sounds have the same $F0$ and bandwidth (USS, RSS), the formant-matching task can be

TABLE I. Conditions used in experiment 1 and in a previous experiment.

	Harmonics near formant	$F0$	BW	Match property	Old Expt. 2	Expt. 1
USS	Unresolved	Same	Same	Identity	x	
RSS	Resolved	Same	Same	Identity	x	
UDS	Unresolved	Different	Same	Excitation	x	
RDS	Resolved	Different	Same	Envelope	x	
USD	Unresolved	Same	Different	Excitation peak		x
RSD	Resolved	Same	Different	Envelope peak		x
UDD	Unresolved	Different	Different	Excitation peak		x
RDD	Resolved	Different	Different	Envelope peak		x

performed veridically by listeners making the sounds identical—the least abstract level of matching.

Introducing a difference in $F0$ requires a more abstract level of matching which differs depending on the resolution of the harmonics near to the formant peak. With unresolved harmonics (UDS), listeners can achieve a veridical match by making the excitation patterns identical (which will not show any harmonic ripple) in the vicinity of the formant peak. With resolved harmonics (RDS), this strategy will also fail in general, so that listeners will need to perform a more abstract match that generalizes across different frequencies of harmonic ripple; one such abstraction is the envelope of the excitation pattern.

If we now consider sounds that differ in bandwidth, the level of abstraction of all matches is increased to similarly abstract levels whether the sounds have the same or different $F0$'s. For sounds with unresolved harmonics near the formant frequency (USD, UDD), the peak in the excitation pattern provides the least abstract criterion for a veridical match. For resolved harmonics (RSD, RDD) the match must be made at a more abstract level, such as by first interpolating an envelope for the excitation pattern and then determining the peak of this envelope.

Although a task may theoretically require a particular level of abstraction in order to perform a veridical match, listeners may adopt a suboptimal strategy that uses a less abstract representation. In the previous paper we tested whether listeners were using a strategy of minimizing the rms error between excitation patterns as a matching strategy for sounds on different $F0$'s. A simulation of this strategy predicted a much higher variability of subjects' performance in matching sounds on different, high fundamentals than was actually found in the experiment. We will make similar arguments in this paper that the strategy of minimizing excitation-pattern rms error also does not explain subjects' performance in the present experiments.

While a difference in bandwidth can in principle substantially alter the criteria that listeners use in a matching experiment, differences in formant bandwidth have generally had rather little effect on listeners' phonetic judgments of vowel-like speech sounds. Carlson *et al.* (1979) found that manipulation of formant bandwidth had less of an influence on listeners' dissimilarity ratings than did changes to formant frequency. In addition, using similar sounds, Klatt (1979) showed that when making phonetic comparisons listeners pay far less attention to changes in formant bandwidths than they do when making psychophysical judgments. The exception to this pattern is that a reduction in the prominence of the first formant peak has been proposed as the primary perceptual correlate of vowel nasality (Hawkins and Stevens, 1985). However, in the experiments described here single-formant sounds are used where the formant frequency is well above the first-formant range, so a percept of changing nasality is unlikely to influence listeners' judgments.

The first experiment asks how reliably listeners can match single-formant sounds that differ in bandwidth (USD, RSD) or in both $F0$ and bandwidth (UDD, RDD). In particular, we are interested to discover whether the reliability of listeners' matches mirrors the hierarchy of levels of abstraction that we introduced in Table I, thereby producing perceptual evidence for different levels of abstraction for the processing of complex sounds.

II. EXPERIMENT 1

A. Stimuli and procedure

The general procedure for the experiment was similar to that used in our previous paper. Listeners had to adjust, by moving a roller-ball up or down, the formant frequency of a 500-ms periodically excited single-formant complex sound to match the timbre of a similar sound with a formant frequency of either 1100 or 1200 Hz. The pair of sounds could be repeated as often as necessary on each trial by pressing the roller-ball's button. The two sounds could have the same or different $F0$'s. Three factors were varied orthogonally across 8 blocks of 20 trials: whether both sounds in a trial were from the low (80 and 90.4 Hz) or the high (221.2 and 250 Hz) $F0$ range, whether the target sound had a narrow (100 Hz) or a wide (200 Hz) bandwidth, and whether target and match sounds had the same or different $F0$'s. These $F0$ manipulations, coupled with the difference in bandwidth, allow us to separate direct effects on matching reliability of a difference in $F0$ (such as distracting listeners from making the match) from the level of abstraction needed to make the match. In the previous paper (Dissard and Darwin, 2000; experiments 2, 3, and 4) matches which were nonidentical always had a difference in $F0$. For each pair of sounds, the direction of pitch was always upward (i.e., the target's $F0$ was lower than the adjustable sound's $F0$), with a frequency ratio of 1:1.13. Within a block the target formant frequency was randomly set to either 1100 or 1200 Hz. Each target sound was matched ten times in a quasirandom order. Each block took about 30 min to complete and the order of experimental blocks was randomized across subjects.

Sounds were synthesized in real time at 22.05 kHz using the parallel branch of SENSYN PPC™ (Sensimetrics, Cambridge, MA) incorporated into custom software. Voice source parameters were set to their default values, which are the same as described in Klatt (1979). Sounds were output through a Digidesign Protools board and presented through Sennheiser HD414 headphones in an IAC booth. An Apple Power Macintosh 7100 computer controlled the experiment; overall output level was around 60 dB SPL. At the beginning of each trial the formant frequency of the adjustable sound was chosen at random from the permitted range (850 to 1500 Hz). As subjects moved the roller-ball to adjust the formant frequency of the comparison sound a screen cursor also moved. The cursor was recentered after each button press so that subjects could not base their adjustment on the cursor's position. Moving the cursor by half a screen led to a maximum change of about 33 Hz (fine adjustment, minimum change: 0.1 Hz) or 100 Hz (coarse adjustment, minimum change: 0.3 Hz). Subjects could toggle between the coarse and fine adjustments. If the formant frequency was adjusted outside the permitted range, it was reset to a random value within the range and a warning sound played.

Nine subjects (including the first author) participated in the experiment. Subjects were university students or staff

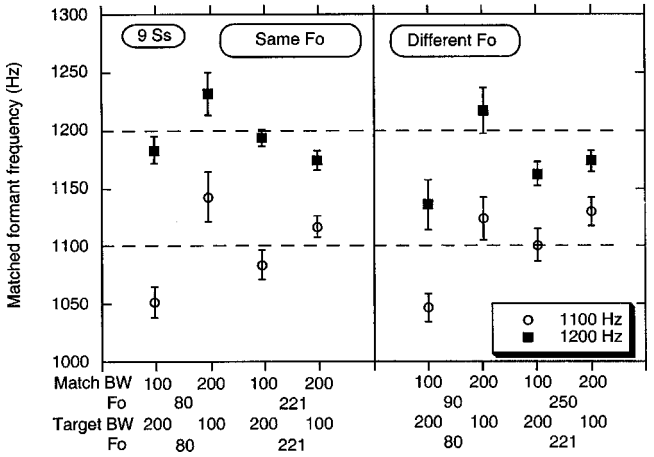


FIG. 2. Mean matched formant frequencies in experiment 1 for target formant frequencies of 1100 and 1200 Hz. Error bars are standard errors of the mean over nine subjects. The left panel shows matches made with stimuli on different $F0$; the right panel shows matches with stimuli on the same $F0$.

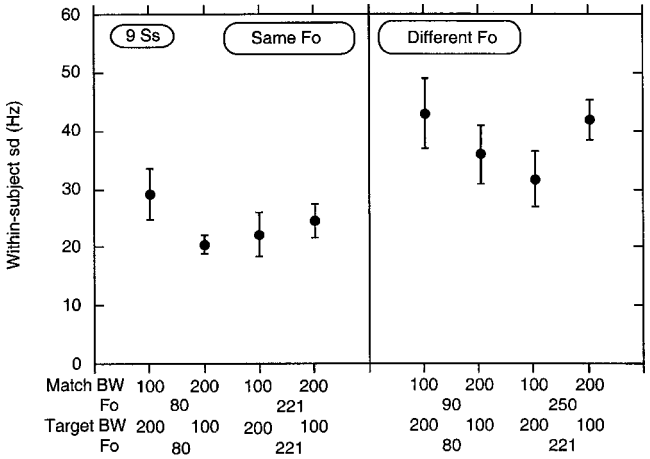


FIG. 3. Average within-subject standard deviations of matches in experiment 1 with their standard errors over nine subjects. The left panel shows matches made with stimuli on different $F0$; the right panel shows matches with stimuli on the same $F0$.

and were paid for their services. All had pure-tone thresholds within the normal range at octave frequencies between 250 Hz and 4 kHz. All the subjects had participated in previous matching experiments and were experienced in formant-matching tasks.

B. Results

1. Mean matches (veridicality)

Mean matches to the 1100- and 1200-Hz targets (Fig. 2) are well-separated and the standard errors across listeners of these mean matches are generally small regardless of whether the sounds differ in $F0$.

Nevertheless, there are some systematic deviations from veridicality. For unresolved harmonics, matches are higher in frequency when the target has a narrow bandwidth (and the match a wide one) than when it has a wide bandwidth (and the match a narrow one), regardless of whether the sounds are on the same $F0$ or not. The direction of this effect is equivalent to the narrow bandwidth sound being heard as having a higher formant frequency, and may be due to the duller sound of a wider-bandwidth adjustable sound being compensated for by it being adjusted to a higher formant frequency (and vice versa). The resolved-harmonic sounds show a similar though weaker tendency.

2. Match variability (reliability)

The mean within-subject standard deviations of matches across the different conditions are shown in Fig. 3.

The main results from this experiment, where subjects made matches across sounds on different bandwidths, are (i) that matches are more variable for sounds that have different $F0$'s than for those with the same $F0$ [$F(1,8)=9.95$, $p<0.02$], but (ii) that this difference does not depend on whether the sounds have resolved or unresolved harmonics near the formant peak [$F(1,8)=0.1$, $p>0.5$].

If we contrast these present results with those from experiment 2 in Dissard and Darwin (2000), which used similar conditions but with sounds that always had the same bandwidth, an interesting pattern then emerges which is illustrated in Fig. 4 for the eight listeners who took part in both experiments. For sounds with the same $F0$ (left-hand panel of Fig. 4), making the bandwidths different increased the match variability equally for sounds with unresolved and resolved harmonics [main effect of bandwidth: $F(1,7)=36.9$, $p=0.0005$; bandwidth$\times F0$ interaction: $F(1,7)<1$]. But, when the sounds had a different $F0$ (right-hand panel of Fig. 4), making the bandwidths different increased the match variability more for sounds with unresolved than with resolved harmonics [bandwidth$\times F0$ interaction: $F(1,7)=13.4$, $p=0.008$]. The three-way interaction reflecting the different patterns in the two panels of Fig. 4 is marginally significant [$F(1,7)=4.1$, $p=0.08$].

This pattern of results can be interpreted as follows. Matches show low variability if they can be made directly from the excitation pattern (i.e., same-bandwidth matches for unresolved harmonics on either the same or different $F0$'s, or same-bandwidth matches for resolved harmonics provided that they are on the same $F0$). If listeners are prevented from making a match directly from the excitation pattern either by putting resolved harmonics on a different $F0$, or by giving the sounds a different bandwidth, then variability increases;

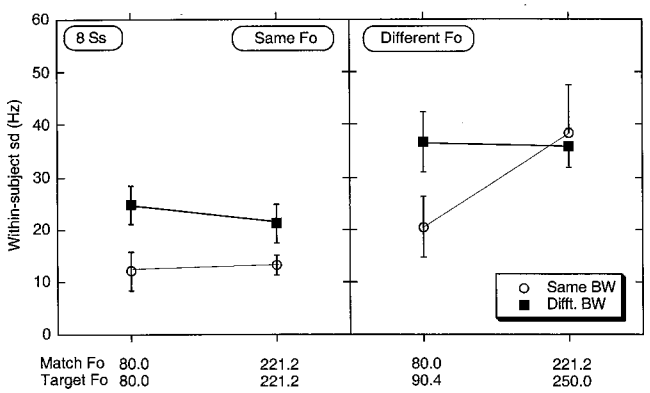


FIG. 4. Comparison of within-subject standard deviations between the present experiment 1, in which target and match sounds had the different bandwidths and experiment 2 of Dissard and Darwin (2000), in which they had the same bandwidth.

TABLE II. Comparison of experimental matching accuracy in experiment 1 with predictions of a model based on best-fitting excitation patterns.

Condition	Match BW/ target BW	Match criterion	Model formant range (Hz)	±1 within-subject s.d. (Hz)	Ratio col.4/col.5
USD	100/200	Excit peak	156	56	2.8
USD	200/100	Excit peak	163	40	4.1
RSD	100/200	Env peak	84	44	1.9
RSD	200/100	Env peak	146	52	2.8
	Average		137	48	2.9
UDD	100/200	Excit peak	145	86	1.7
UDD	200/100	Excit peak	172	72	2.4
RDD	100/200	Env peak	182	64	2.8
RDD	200/100	Env peak	313	84	3.7
	Average		203	76.5	2.7
	Grand Av		170	62	2.8

but having a difference both in $F0$ and in bandwidth gives no further increase in variability over having a difference in either one.

An additional point: sounds which have the same $F0$ but differ in bandwidth are matched with about the same reliability whether they have resolved or unresolved harmonics near the formant peak. This similarity is interesting since the simplest way that the matches could be made differs between the resolved and unresolved cases. With unresolved harmonics, the match can be made through the peak in the explicit excitation pattern. But, with resolved harmonics the match can only be made via the peak in the interpolated envelope.

Taken together, these observations suggest that the metric of match reliability that we have used is not sensitive to differences in the level of abstraction of a match beyond the simple distinction between making a match at the level of the excitation pattern, and making a match at a more abstract level.

Before adopting this somewhat negative conclusion, we should make sure that listeners are not just making the matches via the raw excitation pattern. In Dissard and Darwin (2000) we addressed this problem by simulating a strategy of minimizing the error between the target and the match excitation patterns. In the earlier paper we found that when listeners were matching identical stimuli (same $F0$, same bandwidth) the range of formant frequencies, for which the mean-square error difference between the excitation patterns was 1 dB above its minimum, was about 1.4 times the experimentally determined within-subject standard deviations. Following this approach we use the same model as in the earlier paper to compare the expected variability of their matches from the model with those from the experimental data. The fourth column of Table II shows the range of formant frequencies for which the mean-square error difference between the excitation patterns of target and matched sounds was 1 dB above its minimum. The fifth column of Table II shows ±1 within-subject standard deviation from our experimental data, and the sixth column shows the ratio of these two values. The average ratio of the two measures is 2.9, double the expectd ratio of 1.4. Listeners are thus making the matches twice as reliably as predicted from a simple excitation-pattern-based model. It is thus unlikely that listeners are simply finding the best match of the raw excitation patterns.

III. EXPERIMENT 2

In our previous paper (Dissard and Darwin, 2000) we showed that the perceptual cost (as measured by match variability) of interpolating a spectral envelope in one-formant complex sounds was higher for stimuli with resolved harmonics than for stimuli with unresolved harmonics.

The goal of experiment 2 was to assess whether this difference is in fact a continuous function which decreases smoothly as the number of harmonics that sample the formant envelope increases, or a discrete one that reflects the difference between resolved and unresolved harmonics. We test these two possibilities by increasing the frequency of $F0$ within the resolved-harmonic range. If match variability increases gradually, then the effect is due simply to the density of harmonic sampling. If there is no increase in match variability within the resolved-harmonic range, then the difference that we obtained previously is likely to be due to the discrete difference between resolved and unresolved harmonics.

A complicating factor in designing the experiment was that formant frequency difference limens are generally smaller when the formant peak is located symmetrically between two harmonics than when it lies on a harmonic frequency (Lyzenga and Horst, 1997). We control for any effect of the position of harmonics under the formant peak on the variability of matches by selecting $F0$ values so that both the target and the adjustable sound on a particular trial had the formant frequency either at a harmonic frequency, or midway between two harmonic frequencies.

A. Stimuli and procedure

As in the previous experiment, subjects matched single-formant (1100-Hz) sounds that differed in $F0$. Within a condition, the $F0$ of the target and the match sound were constructed so that both sounds either had a harmonic at the formant frequency, or had the formant frequency symmetrically between the harmonics. Half the stimuli had either the third, fourth, fifth, or sixth harmonic aligned to the formant

TABLE III. $F0$ values used in experiment 2.

Block /‖\		Block /\	
Harmonic number	$F0$ (Hz)	Harmonic number	$F0$ (Hz)
3.5	314	3	366.7
4.5	244	4	275
5.5	200	5	220
6.5	169.2	6	183.3

peak of the target sound (/|\); the other half had a formant located halfway between two harmonics (/‖\). The exact $F0$ values are given in Table III.

Because of the impossibility of keeping an exactly constant $F0$ ratio between match and target, we made sure that the average $F0$ ratio within each of the two groups was very similar (see Table IV).

Nine subjects (including the first author) participated in the experiment. Subjects were university students or staff and were paid for their services. All had pure-tone thresholds within the normal range at octave frequencies between 250 Hz and 4 kHz and were already trained in formant-matching tasks.

B. Results

1. Mean matches (veridicality)

The mean matched formant frequencies for each condition across the six subjects are shown in Fig 5. All matches are relatively accurate; no main effect of the $F0$ factor is observed [$F(5,8)=1.15$]; in addition, accuracy of matches is the same for stimuli with a formant peak between two harmonics and for those with a harmonic at the formant frequency.

2. Match variability (reliability)

Figure 6 shows the mean within-subject standard deviation of matches within each block of trials, together with the standard error of these means across the nine subjects. Variability of matches does not differ significantly either for the six $F0$ conditions [$F(5,8)=1.84$] or between the two formant-alignment conditions (/|\ vs /‖\), [$F(1,8)=1.79$].

C. Discussion

Experiment 2 has found no effect on the accuracy of formant matches of either the density or the alignment of harmonics under the formant envelope. Since there is no increase in match variability within the resolved-harmonic range, then the difference that we obtained previously between conditions with resolved and unresolved harmonics is likely to be due to the discrete difference between resolved and unresolved harmonics rather than simply to the density of harmonic sampling. Our previous conclusion, that matches have lower variability if they can be made directly from the excitation pattern rather than requiring the interpolation of a spectral envelope, is therefore justified by the results of experiment 2.

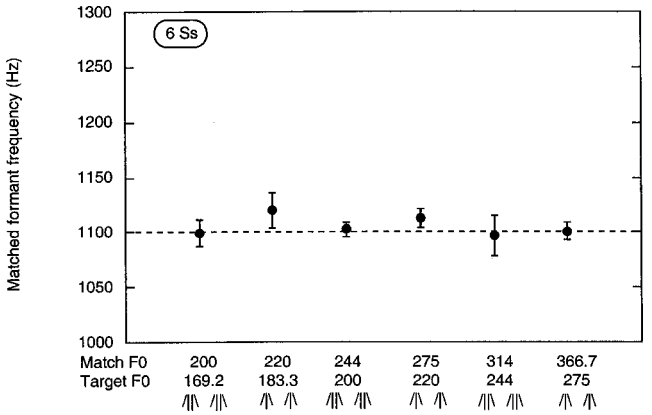


FIG. 5. Mean matched formant frequencies in experiment 2 for a target formant frequency of 1100 Hz. Error bars are standard errors of the mean over six subjects.

The experiment also found no difference in match variability between conditions where a harmonic coincided with the formant frequency (/|\) and those where the formant frequency lay midway between two harmonics (/‖\). This result contrasts with the general finding of Lyzenga and Horst (1997) that formant frequency difference limens are higher in the former condition than in the latter. The discrepancy, however, is only apparent since in the specific conditions of their experiment that most closely match the stimuli that we have used (Expt. 1, Fig. 3, Klatt envelope, $F0=200$ Hz, F_1) Lyzenga and Horst also find no reliable difference between the two harmonic alignments.

IV. GENERAL DISCUSSION

The results of these experiments confirms the conclusions from our previous paper that formant matches that can be made on the basis of the matched sounds' excitation patterns are less variable than are matches that require a more abstract representation of the sound, such as its envelope. The present experiments have varied the bandwidth as well as $F0$ across the target and match sounds and have shown that a difference in bandwidth between the target and match sounds generally increases the variability of the matches. The exception to this general rule is when the sounds differ

TABLE IV. $F0$ values and ratios for target and match sounds in experiment 2.

Block /‖\			Block /\		
Match $F0$	Target $F0$	Ratio M/T	Match $F0$	Target $F0$	Ratio M/T
200	169.2	1.18	220	183.3	1.2
244	200	1.22	275	220	1.25
314	244	1.29	366.7	275	1.33
mean=1.23			mean=1.26		

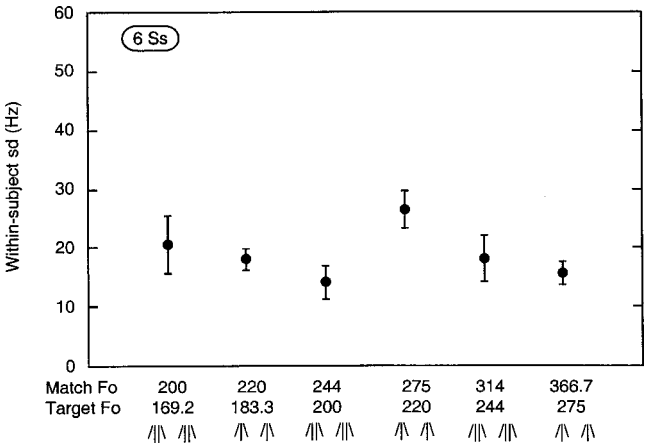


FIG. 6. Average within-subject standard deviations of matches in experiment 2 with their standard errors over six subjects.

in $F0$ and also have $F0$'s that give resolved harmonics around the formant frequency. Under these conditions, giving additionally a difference in bandwidth does not increase matching variability.

This pattern of results implies that although the matching task is sensitive to the difference between matches that can be made on the basis of the excitation pattern and those that need a more abstract representation, it is not sensitive to the difference between matches that can be made on the basis of an interpolated spectral envelope, and those that require the peak of the spectral envelope (or formant frequency) to be extracted.

This result is compatible with the possibility that listeners are basing their matches on the peak of the spectral envelope even when the spectral envelope itself can provide an adequate basis for a match. Although attractive, this conclusion needs more supporting evidence. Further work on this question could make use of the hierarchy of properties shown in Table I to explore the sensitivity of other tasks to different levels of abstractness needed to perform the task.

ACKNOWLEDGMENTS

The first author was supported by Grant No. GR/L03422 from the UK EPSRC. Paul Russell wrote the software for formant matching based on SENSYN for the PowerPC. We are grateful to Rob Hukin for technical assistance, and to Peter Assmann for helpful reviewer's comments.

[1]An alternative strategy for forcing listeners to use the formant peak in their matching is to make the matching sound a sine wave. This technique was used by Hermansky (1987) but suffers from the problem that it encourages listeners to hear out individual partials of a complex (particularly when they are close to the formant frequency) and to match the frequency of a particular partial. By keeping both the target and the match as complex sounds, we hope to dissuade listeners from hearing out the individual components of a complex.

[2]For a comparison of various other metrics for comparing speech spectra see Nocerino *et al.* (1985).

Carlson, R., Granström, B., and Klatt, D. H. (**1979**). ''Vowel perception: The relative perceptual salience of selected acoustic manipulations,'' STL-QPSR **3–4**, 73–83.

Dissard, P., and Darwin, C. J. (**2000**). ''Extracting spectral envelopes: Formant frequency matching between sounds on different and modulated fundamental frequencies,'' J. Acoust. Soc. Am. **107**, 960–969.

Hawkins, S., and Stevens, K. N. (**1985**). ''Acoustic and perceptual correlates of the non-nasal–nasal distinction for vowels,'' J. Acoust. Soc. Am. **77**, 1560–1575.

Hermansky, H. (**1987**). ''Why is the formant frequency difference limen asymmetric?,'' J. Acoust. Soc. Am. **81**, S18.

Klatt, D. H. (**1979**). ''Perceptual comparisons among a set of vowels similar to /ae/: Some differences between psychophysical distance and phonetic distance,'' J. Acoust. Soc. Am. **66**, S86.

Lyzenga, J., and Horst, J. W. (**1997**). ''Frequency discrimination of stylized synthetic vowels with a single formant,'' J. Acoust. Soc. Am. **102**, 1755–1767.

Moore, B. C. J., and Glasberg, B. R. (**1983**). ''Suggested formulae for calculating auditory-filter bandwidths and excitation patterns,'' J. Acoust. Soc. Am. **74**, 750–753.

Nocerino, N., Soong, F. K., Rabiner, L. R., and Klatt, D. H. (**1985**). ''Comparative study of several distortion measures for speech recognition'' Speech Commun. **4**, 317–331.

Systematic distortions of auditory space perception following prolonged exposure to broadband noise

Simon Carlile[a)]
Auditory Neuroscience Laboratory, Department of Physiology, and Institute for Biomedical Research, University of Sydney, Sydney NSW 2006, Australia

Stephanie Hyams and Skye Delaney
Department of Physiology, University of Sydney, Sydney NSW 2006, Australia

(Received 10 August 1999; revised 16 October 2000; accepted 3 April 2001)

Perceptual distortions referred to as aftereffects may arise following exposure to an adapting sensory stimulus. The study of aftereffects has a long and distinguished history [Köhler and Wallach, Proc. Am. Philos. Soc. **88**, 269–359 (1944)] and a range of aftereffects have been well described in sensory modalities such as the visual system [Barlow, in *Vision: Coding and Efficiency* (Cambridge University Press, Cambridge, 1990)]. In the visual system these effects have been interpreted as evidence for a population of cells or channels specific for certain features of a stimulus. However there has been relatively little work examining auditory aftereffects, particularly in respect of spatial location. In this study we have examined the effects of a stationary adapting noise stimulus on the subsequent auditory localization in the vicinity of the adapting stimulus. All human subjects in this study were trained to localize short bursts of noise in a darkened anechoic environment. Adaptation was achieved by presenting 4 min of continuous noise at the start of each block of trials and was maintained by a further 15-s noise burst between each trial. The adapting stimulus was located either directly in front of the subject or 30° to the right of the midline. Subjects were required to determine the location of noise burst stimuli (150 ms) in the proximity of the adapting stimulus following each interstimulus period of adaptation. Results demonstrated that following adaptation there was a general radial displacement of perceived sound sources *away* from the location of the adapting stimulus. These data are more consistent with a channel-based or place-based process of sound localization rather than a simple level-based adaptation model. A simple ''distribution shift'' model that assumes an array of overlapping spatial channels is advanced to explain the psychophysical data. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1375843]

PACS numbers: 43.66.Qp, 43.66.Pn, 43.66.Ba [DWG]

I. INTRODUCTION

Sound localization depends on a variety of localization cues present at each ear. Monaural cues arise from the spectral modifications of incoming sound by the external ear and binaural cues arise as a result of differences in the sound at the two ears. At the level of the cochleae, the sensory epithelia encode frequency and project into the auditory nervous system as tonotopic maps of frequency. However, auditory localization depends on higher-order computational processing that extracts the monaural and binaural location cues from this ordered frequency representation. The generation of a neural representation of space, and hence the accurate perception of auditory space, requires a complex integration of both monaural and binaural inputs (see Carlile, 1996a for review).

Physiological investigations of the neural representation of auditory space in the barn owl (Knudsen and Konishi, 1978a, b; Knudsen *et al.*, 1987) and the mammalian superior colliculus (e.g., guinea pig: King and Palmer, 1983; ferret: King and Hutchings, 1987; Carlile and King, 1994a, b; cat: Middlebrooks and Knudsen, 1984) have demonstrated neurones whose response rates vary systematically as a function of the spatial location of the sound source. The topographic arrangement of these neurons is consistent with the notion of a neural map of auditory space at least at these levels of the central nervous system. Furthermore, there is some evidence for lateral interactions between these ''space-mapped'' neurons (Knudsen and Konishi, 1978a, b; King *et al.*, 1990). In the study reported here we were interested in examining how this kind of ''space-selective'' processing might be manifest psychophysically in human subjects.

One way in which psychophysical studies have been used to examine complex sensory processes is via the phenomena of adaptation and aftereffects. Following prolonged exposure to a constant stimulus, a variety of perceptual distortions have been described. Such effects have been demonstrated in early experiments examining the perception of sound location using a range of adapting stimuli. Flugel (1921) and Bartlett and Mark (1922) described an auditory localization aftereffect following exposure to both binaural and monaural sound sources. In these early studies, stimulus levels were manipulated by adjusting the lengths of rubber tubing that delivered the sound stimuli to each ear. Under such stimulus conditions, sounds were lateralized within the head rather than localized in the free-field. The sound stimulus was adjusted by the subjects both prior to and following

[a)]Author to whom correspondence should be addressed; electronic mail: simonc@physiol.usyd.edu.au

 0001-4966/2001/110(1)/416/9/$18.00

stimulation by ''fatiguing'' tones so that it was perceived to be located in the center of the head. Following stimulation of one ear by an adapting tone, there was a systematic displacement of subsequent binaural stimuli to locations away from the ear to which the adapting tone was presented.

Lateralization experiments by Elfner and Perrott (1966) using headphone presentation of three different tone stimuli (i.e., 700, 1000, and 3000 Hz) demonstrated shifts in the auditory image following prolonged exposure to binaural intensity mismatches of up to 126 min. The results indicated that in most cases, the perceived location of the test stimulus was shifted toward the ear receiving the less intense stimulus. However, in some cases the shift was toward the ear where the adapting stimulus was more intense. These variations might be explained by the mismatch between the frequency of the adapting stimulus and the test stimulus.

Studies by Krauskopf (1954) and Taylor (1962) confirmed the initial observations by Flugel. Both these later experiments were conducted using free-field stimuli. In the first study, subjects adjusted the location of a sound stimulus so that it appeared to lie on the median plane both before and after a 2-min exposure to an adapting broadband stimulus. In the second experiment, subjects indicated the location of the test stimulus using a spatial scale. Prolonged exposure to the adapting stimulus on one side of the median plane shifted the subjects' perception of the midline away from the adapting stimulus. Taylor (1962) measured the magnitude of this perceived displacement when the separation between the test and adapting stimulus locations was varied. The magnitude of the displacement increased as a function of the increasingly lateral location of the adapting stimulus and peaked at 3° when the adapting stimulus was located at around 30° off the midline. Both free-field and headphone-based studies yield qualitatively similar results. This supports the notion that the headphone-based studies are examining similar processes as those measured in free-field studies, despite some differences in the perceptual experience under each condition. The results of Curthoys (1968), Elfner and Perrott (1966), and Thurlow and Jack (1973) all confirm the earlier findings of a displacement of perceived target location away from the adapting stimulus: a so-called ''repulsion effect.'' In a series of headphone-based studies, Thurlow and Jack (1973) demonstrated that offsets of interaural time and interaural level differences were equally successful in generating this aftereffect, although the results were less straightforward when interaural time and level difference were traded off as test and adapting stimuli, respectively.

Comparison of the absolute magnitude of the aftereffect across previous studies is complicated by the fact that the duration of the adapting stimuli varied considerably (from 10 s to a few minutes), the acoustic environments varied (headphone, free-field, anechoic) and in some free-field studies, free head movement was allowed. However, despite these differences there is a general finding of a ''repulsion'' effect. Curthoys proposed that the auditory spatial aftereffect resulted from a level-dependent adaptation of the ear closest to the adapting stimulus that was greater than the adaptation in the further ear. He suggested that this unequal adaptation would lead to a shift of subsequently perceived locations away from the more adapted ear. The changes in the magnitude of the effect with increasing displacement of the adapting stimulus from the midline observed by both Taylor and Curthoys is in fact consistent with what is known about the location-dependent changes in interaural level differences (Carlile and Pralong, 1994) and the resulting difference in the adaptation of the two ears. This overall explanation of the auditory spatial aftereffect is dependent on the assumption that the sensitivity of the auditory system to interaural level differences is sufficiently dominant to explain the effects observed. However, such an explanation only explores how adaptation to ILD cues affects perception of the target location and ignores the potential roles of the ITD and spectral cues. Indeed, the work of Thurlow and Jack (1973) also indicate a role for interaural time cues in the generation of this aftereffect. In addition, in experiments using VAS stimuli where the interaural time cues have been set in conflict with other localization cues, there is some evidence to suggest that perception of spatial location is dominated by the interaural time difference cue when low frequencies are present (Wightman and Kistler, 1992).

However, a number of predictions can be made based on the adaptation model discussed previously (Curthoys, 1968). If the adapting stimulus is located to one side of the anterior midline, all the target positions, whether they be located between the adapting stimulus location and midline or between the adapting stimulus and interaural axis, would be predicted to shift toward the anterior midline and therefore toward the location of the less adapted ear. Furthermore, an adapting stimulus located directly in front of the subject (where interaural differences are zero) would adapt both ears equally and should result in no auditory spatial aftereffect. In this study we examined these aftereffects using a head pointing localization task for stimuli presented in anechoic free space (Carlile *et al.*, 1997) where the subject had no preconceptions of the potential locations of the stimuli. In addition to testing the above-mentioned predictions, this approach also allowed an examination of the adaptation aftereffect over the dimensions of azimuth and elevation.

II. METHODS

We examined the ability of a total of fifteen human subjects to localize a sound source or a visual target before and after exposure to a free-field adapting auditory stimulus. The behavioral methods used in assessing localization accuracy have been described in detail previously (Carlile *et al.*, 1997). Briefly, all training and testing took place in a dark, anechoic chamber. Within the chamber, a robot arm carrying a stimulus speaker and a light emitting diode (LED) could be moved to almost any location on the surface of an imaginary sphere surrounding the subject. A single-pole coordinate system was employed to describe stimulus position with azimuth 0° and elevation 0° (0°/0°) directly ahead on the audio-visual horizon of the subject. Upward elevation and rightwards azimuth increased positively. The position of the head was tracked using an electromagnetic positioning system (Polhemus, Isotrak). The receiver was mounted on top of the subject's head using an adjustable head strap. A visual reference system, placed directly in front of the subject con-

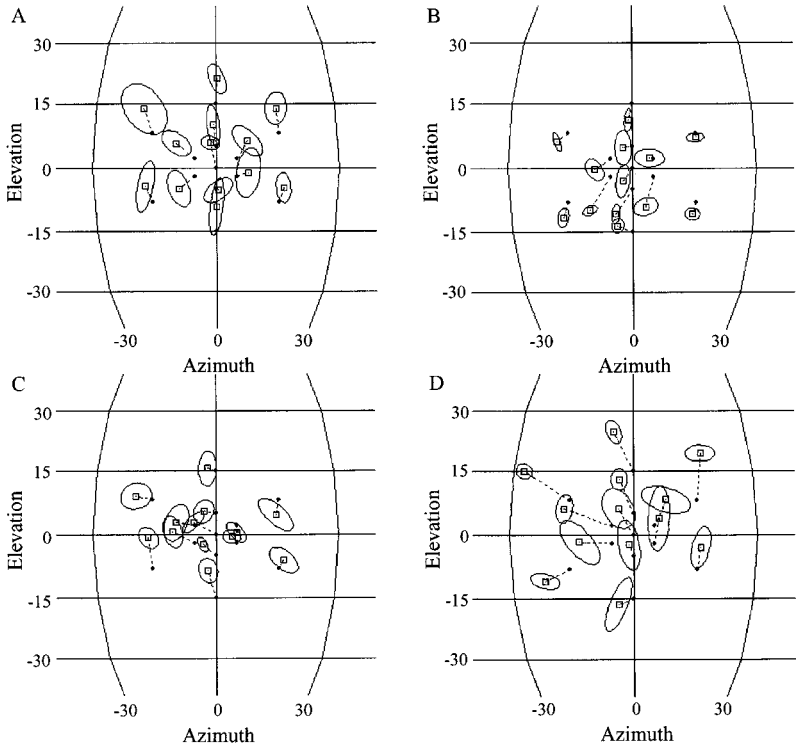


FIG. 1. The localization performance of four representative subjects prior to adaptation for the test locations around the azimuth 0°/elevation 0° location. These data have been chosen to demonstrate the range of localization biases seen in our population of fifteen subjects. Each plot represents a segment of the frontal hemisphere with the lines of azimuth spaced at 30°, the lines of elevation spaced at 15°, and the azimuth 0°/elevation 0° location directly at the center. The small closed circles indicate the target locations and the centroids of the localization estimates are indicated by the open squares. The dashed line between the closed circles and open squares indicates the magnitude and direction of the localization bias for each test location. Each centroid is calculated from nine repeat localizations for each target location, and the standard deviations of the distributions are indicated by the ellipses surrounding each centroid.

sisting of red and green colored LEDs, worked in conjunction with the head tracker and could be used to indicate to the subjects the position of his or her head relative to the stimulus coordinate system. This system was used to aid the subject in placing his or her head in the calibrated start position at the beginning of each trial. In a supplementary experiment, the visual reference system was also used to assist the subject in keeping his or her head stationary during prolonged exposure to the adapting stimulus. Subjects underwent a period of training to assist them in reliably pointing toward the perceived target location with the nose. Once satisfied he or she was correctly indicating the perceived location of the target, the subject pressed a hand-held button and the position of his or her head was recorded (for details see Carlile *et al.*, 1997).

Following the period of training, the baseline localization accuracy of each subject was determined (see Carlile *et al.*, 1997). Subjects were required to localize a 150-ms noise burst (200 Hz–14 kHz ± 3 dB; 60 dB SPL) from 76 test positions. Subjects completed 4 such tests prior to undertaking the adaptation experiments, and the localization performance of each subject was within the normal range previously reported by this laboratory (Carlile *et al.*, 1997).

Four adaptation experiments were carried out in the course of this study: two primary adaptation experiments and two supplementary experiments, the details of which are described below. In all experiments the testing procedure was divided into preadaptation and adaptation blocks of localization trials. Localization accuracy was determined over a spatial area surrounding the adapting speaker: Twelve test locations were arranged as two concentric rings centered on the location of the adapting stimulus with a spherical radius of around 6° and 17°, respectively. Stimuli were located on each ring with roughly equal circumferential spacing (see Fig. 1). In two separate experiments, adapting stimuli were located directly in front at 0°/0° (test positions: 0°/−15°; −15°/−8°; 15°/−8°; 0°/−5°; −5°/−2°; 5°/−2°; 0°/0°; −5°/2°; 5°/2°; 0°/5°; −15°/8°; 15°/8°; 0°/15°) and 30° to the right of the midline (test positions: 30°/−15°; 15°/−8°; 45°/−8°; 30°/−5°; 25°/−2°; 35°/−2°; 30°/0°; 25°/2°; 35°/2°; 30°/5°; 15°/8°; 45°/8°; 30°/15°). The location of the adapting stimulus was also included in the set of test locations. In each block, stimuli were presented 3 times at each of the 13 positions giving a total of 39 localization trials per block. Subjects completed 3 blocks for each test condition. This resulted in 9 repeat localization trials for each location and test condition, for a total of 117 localization trials per subject per test condition. Localization judgments were plotted using a spherical plotting and analysis package (SPAK: Leong and Carlile, 1998). We chose to illustrate these data using spherical means as this avoids the need to make assumptions about the type of coordinate system to employ (e.g., a single or double pole projections for azimuth and elevation). The mean (or centroid) of the cluster of localization estimates for each test location was calculated and plotted for each subject (see Fig. 1) along with the standard deviation of each distribution. Further methodological, graphical, and statistical analyses are described in the following where appropriate.

In the two primary adaptation experiments, ten subjects were tested using the adaptation stimulus directly ahead (0°/0°) and eight subjects were tested using the adaptation stimulus located at 30° to the right of the subject (30°/0°). Both positions were located on the audio-visual horizon and five of the subjects were common to both experiments. The adapting stimulus was a continuous broadband noise (similar to the target stimuli), initially presented for 4 min at the beginning of each block of trials and then for 15 s between each localization trial. During exposure to the adapting

stimulus, the subject was requested to remain in a fixed position facing directly ahead (0°/0°) during which time a LED at 0°/0° was illuminated to assist in this task. In the supplementary adaptation experiments (see the following), the head position indicator was also used to help the subject maintain head position for the duration of the adapting stimulus. Following exposure to the adapting stimulus, the localization test was repeated with the exception that the adapting stimulus was presented for 15 s between each test stimulus. Thus, subjects localized 39 positions per block and each block took between 20 and 25 min of testing. The order of the stimulus positions was randomized within each block. There was at least a 24-h break between each block of adaptation tests to minimize the effects of any prolonged adaptation aftereffect (see also Curthoys, 1968).

III. RESULTS

A. Principal adaptation experiments

Prior to any adaptation, each subject's localization performance was measured using the 13 test locations. Each subject displayed some localization bias in the preadaptation tests (Fig. 1) although the overall magnitude of the errors was within the range of localization errors seen in a larger population of subjects (see Carlile *et al.*, 1997). To illustrate the range of subject-related localization bias, example control localization results are plotted for four subjects for sound locations around 0° azimuth, 0° elevation (Fig. 1). For one subject illustrated, most targets were consistently localized as slightly too high [Fig. 1(a)] whereas for another, particular locations were systematically localized as too low [e.g., Fig. 1(b)]. In a third subject the localizations were systematically too right-ward [Fig. 1(d)] whereas in a fourth, the errors were more randomly directed [Fig. 1(c)].

For the purpose of plotting the adaptation results we were concerned that pooling localization data across subjects with different systematic biases would obscure changes produced by the experimental manipulation. Figure 2 plots the centroid of the control responses (open squares) for the subject shown in Fig. 1(b) indicating a general downward bias in this subject's preadaptation localization response. Following preadaptation testing, the subjects were then exposed to the adapting stimulus as described previously and his or her localization performance remeasured. The principal effect of the adaptation was to displace the localization away from adapting speaker location as can be seen by the radial shifts in the adaptation responses (open circles) compared to the control responses (open squares) (Fig. 2). However, in cases where the underlying control localization bias was toward the location of the adapting speaker then the extent of this effect would be underestimated if plotted in terms of the actual location of the target. In addition, combining data across subjects with the opposite directional biases would act to cancel any effect due to adaptation. To correct for this bias, all the subjects' responses were normalized in the following way. The centroids for both the preadaptation and adaptation responses were shifted by an identical amount to remove the bias in the preadaptation response at each target position. This process is equivalent to subtracting the bias in the preadaptation response from both preadaptation and adaptation response centroids at each target location, and results in the corrected preadaptation centroids falling exactly on their respective target locations [cf. Figs. 2 and 3(b)]. Thus, the deviation of the centroid of the adapted response from the actual location represents the effects of the adaptation with the underlying systematic bias removed. These data were calculated for each subject, the shifted responses were then pooled and variations across the population were calculated for each test location. In Fig. 3 the adaptation data have been plotted for the same four individual subjects illustrated in Fig. 1 to indicate again the range of the adaptation aftereffect.

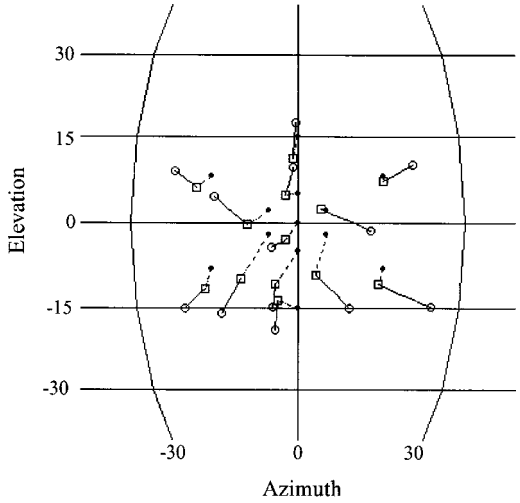


FIG. 2. The mean localization estimates for preadaptation (open squares) and adaptation (open circles) conditions for one subject. The distributions of the estimates have been omitted for clarity [see Fig. 1(b)]. All other details are as for Fig. 1.

The centroids of the normalized localization estimates are plotted for the adapting speaker located at 0°/0° (Fig. 4) pooled from all ten subjects tested at this location. The surrounding ellipses represent the standard deviation of the pooled normalized localization estimates for each location. The perception of spatial location in the vicinity of the adapting source was systematically distorted following exposure to the adapting stimulus. We have quantified this by calculating the spherical angle between the location of the adapting stimulus (at the center of the array) and each of the centroids of the location judgments for both the preadaptation and adaptation conditions. With the exception of the test location corresponding to the location of the adapting stimulus the differences between these two angles indicated that there was a shift in perceived location away from the adapting stimulus following adaptation (mean $2.6° \pm 5.4°$; mean $\pm$s.d.).

When the adapting stimulus was located at 30°/0°, a similar pattern of distortion was evident (Fig. 5). Stimulus locations between the midline and the adapting stimulus were localized toward the anterior midline, and locations to the right of the adapting stimulus were shifted mainly upward or downward. The average extent of the adaptation-induced shifts for the 30°/0° location of the adapting speaker

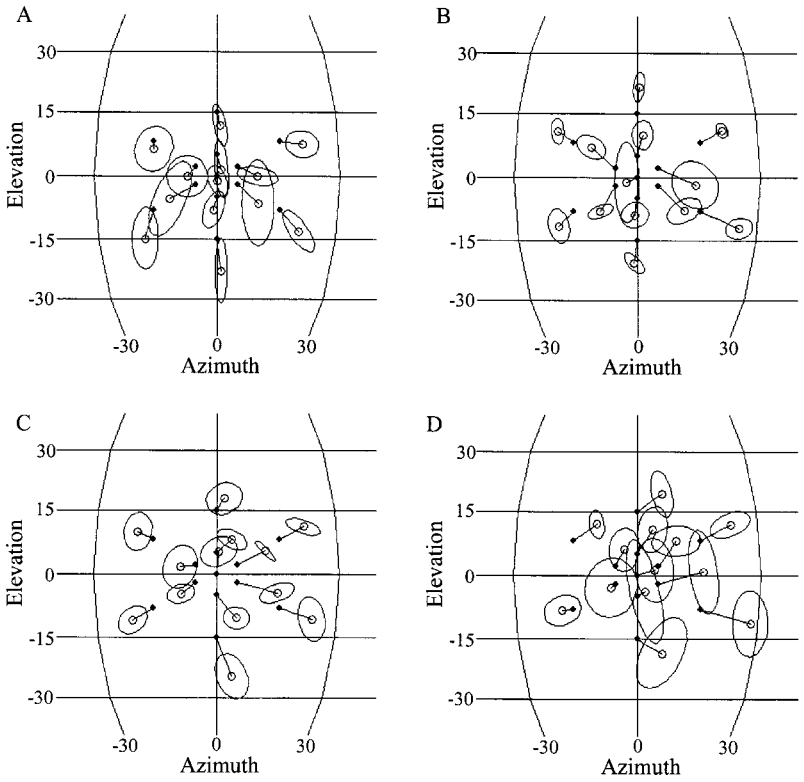


FIG. 3. The normalized adapted responses for the same four subjects as in Fig. 1. As demonstrated in Fig. 2, these data have been normalized by shifting each localization estimate so that the preadaptation estimates overlie the actual target location (see the text for details). This procedure removes the individualized preadaptation localization bias. All other details are as in Fig. 1.

was $3.7° \pm 7.9$. Compared to 0°/0°, the scatter of the data for any one test location was greater for the adapting stimulus location of 30°/0°.

To test if there was a significant radial displacement produced by exposure to the adapting stimulus, a paired t-test compared (a) the spherical angle between the preadapted localization centroid and the location of the adapting speaker with (b) the spherical angle of the adapted localization centroid and the location of the adapting speaker. This was carried out for both the adapting speaker locations 0°/0° and 30°/0° (Table I) and in both cases there was a highly-significant increase in the spherical angle as a result of the adapting stimulus ($p < 0.01$). This indicates a significant radial shift in the perceived location away from the location of the adapting stimulus.

B. Supplementary adaptation experiments

1. Supplementary experiment 1

In the two experiments reported previously the subjects were requested to keep facing forward during the course of exposure to the adapting stimulus. To aid this process subjects were provided with a visual reference point located at 0°/0°. However, we were concerned as to whether small head

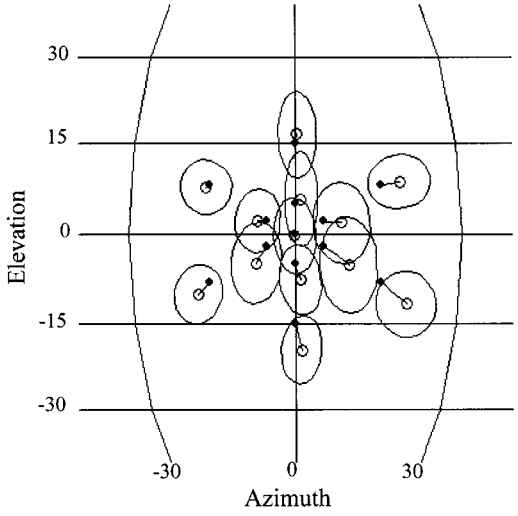


FIG. 4. The mean normalized adaptation localization responses for ten subjects for an adapting stimulus location of azimuth 0°/elevation 0°. Each point is the centroid of 90 localization judgments (small open circles). All other details are as per Fig. 1. Note that, with the exception of the location at azimuth 0°/elevation 0°, there is a clear radial shift in the location of the perceived test location.

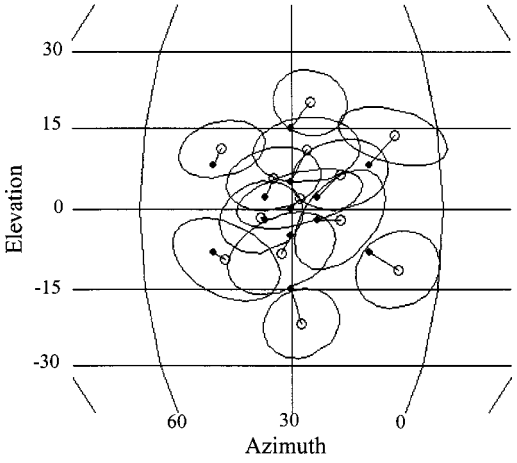


FIG. 5. The mean normalized adaptation localization responses for eight subjects for an adapting stimulus location of azimuth 30°/elevation 0°. Each small open circle indicates the centroid of 72 localization judgements. Compared to the preceding figures, the plot has been shifted so that the meridian at 30° occupies the center of the plot with the midline represented to the right. All other details are as per Fig. 1.

TABLE I. The spherical angle between the localization estimate and the adapting speaker for preadapted and adapted conditions. The mean differences are shown for each experiment together with the statistic for each comparison using a paired-t. Each condition was significant at $p<0.01$. Both supplementary experiments employed an adapting stimulus at 0°/0°.

	Expt 1: 0°/0°	Expt 2: 30°/0°	Sup 1: Auditory	Sup 2: Visual
Mean difference	2.6°±5.4°	3.7±7.9°	3.7±6.3°	0.6±3.6°
Paired-t stat.	39.3	40.87	22.4	18.9
df	1169	1040	350	350

movements may have occurred during the course of the prolonged exposure to the initial 4-min adapting stimulus. Relative changes in the location of the adapting stimulus may have the potential to ''dilute'' or reduce the location-dependent adaptive effect. To test this we devised a means of using the head tracking system to simultaneously monitor the exact position of the head at 10 Hz during the course of the adapting stimulus presentation.

Using this apparatus we studied the head movements of four additional subjects during adaptation experiments conducted at 0°/0°. Three of these subjects were also provided with feedback at 10 Hz as to the location of the head with respect to the stimulus coordinate system (see Sec. II: When the head was properly aligned a central green LED was lit; otherwise, the direction of any misalignment was indicated by the appropriately placed red LED). This monitoring indicated that there were continual small changes in the position of the head, particularly during the 4-min initial exposure to the adapting stimulus. The mean deviation recorded in the one subject who did not receive head-position feedback was 4.5° with a peak deviation of 17.7°. By contrast, mean head deviation for the three subjects who did receive feedback was in general less than 1° with peak deviations ranging from 2.0° to 2.4°.

To examine the potential effect of small head movement during exposure to the adapting stimulus we compared the adaptation of the three subjects who received head position feedback with the ten subjects tested at the same location who did not. Those subjects who received head position feedback demonstrated a mean displacement of the perceived locations of the test stimuli by 3.7°. This shift was significantly larger than the displacement observed in those subjects who received no head position feedback (2.6°; Students t-test, $p<0.01$; see Table I). These data indicate that feedback to the subject about relative head position plays an important role in head stabilization under these conditions. Consequently, the adaptation-induced displacement in the perceived location observed in the two experiments reported above might well be an underestimate of the adaptation effect. That is, the measured displacement may have been larger had head position information been provided to the subjects during adaptation.

2. Supplementary experiment 2

Two previous studies indicate the potential for a shift in the perceptual frame of reference following exposure to an adapting or concurrent auditory stimulus (Chandler, 1961) or to some posturally related stimuli (Lackner, 1974). For example Chandler (1961) demonstrated that a binaural imbalance in dichotically presented tonal stimuli could result in a misperception of the visual vertical by up to one degree. We were interested to determine if the adaptation effect observed here was the result of an auditory sensory adaptation, a disturbance of spatial reference, or a reduction in motor capacity to indicate a specific position in space. To identify the extent of the contribution of any nonauditory sensory component to the displacement of the perceived location of the auditory targets we carried out the same experiment described previously (supplementary 1) with the exception that the subjects were required to indicate the location of a *visual* target both before and after exposure to an adapting *auditory* stimulus. Four subjects were used, two of whom had contributed data to the principal adaptation experiments.

As in the first supplementary experiment, these subjects were also provided with head position information during the course of the adapting stimulus. The auditory adapting stimulus was located at 0°/0° and each subject carried out a total of 117 visual localization trials over the 13 test locations for the preadaptation condition and 117 trials for the adaptation condition. As before, the experiments were carried out in the darkened anaechoic chamber but the visual stimulus was a 150-ms flash emitted from a red LED carried by the robot arm. As would be expected, subjects were quite accurate in indicating the position of the visual stimulus. When pooled across the four subjects, visual localization following auditory adaptation was found to be radially displaced by a mean of 0.6° compared with the preadapted visual localization. A paired t-test indicated that although this difference was small, it was statistically significant ($p<0.01$) and of similar magnitude to the auditory-induced displacement of the visual vertical previously reported by Chandler (1961). These data indicate that a small displacement effect can be attributable to cross-modal or other nonsensory spatial distortion. However, when compared to the magnitude of the displacement of the auditory targets following exposure to an adapting auditory stimulus this accounts for less than 16% of the observed auditory effect measured under the same conditions.

IV. DISCUSSION

This study indicates that exposure to a broadband sound for a number of minutes results in a shift in the perceived location of subsequent sound sources located in the vicinity of the adapting stimuli. There was a general pattern of radial displacement of stimuli away from the location of the adapting stimuli. This was particularly marked for sounds located directly ahead, and less so for sources 30° to the right of the midline where there was a more pronounced shift in the perceived elevation of the test stimuli.

These findings are generally consistent with and extend the previous studies of auditory spatial aftereffects. Such studies have shown perceptual effects for stimulus locations about the anterior midline following adaptation using both free-field (e.g., Taylor, 1962; Curthoys, 1968) and headphone-based stimuli (e.g., Thurlow and Jack, 1973). A general finding was a shift of the perceived judgment of tar-

get stimuli placed around the midline as located further away from the adapting stimulus located off the median plane (Curthoys, 1968) or toward the middle of the head for the lateralized stimuli (Thurlow and Jack, 1973). Thurlow and Jack (1973) also report a small ''repulsion'' effect for a midline adapting stimulus using headphone-based stimuli. This is consistent with the results reported here for the free-field adapting stimulus at 0° azimuth. In two previous free-field studies (Taylor, 1962; Curthoys, 1968), the peak adaptation-induced shift was 3° when the adapting stimulus was located at 30° degrees off the midline. This is within the range of the average shifts noted in this study (2.6°–3.7°).

Previously such shifts had been explained in terms of a simple level-based adaptation model resulting in adaptation-induced changes in apparent interaural level differences (in particular see Curthoys, 1968). That is, if the adapting stimulus resulted in greater adaptation in the near ear compared to the far ear, the effective ILDs caused by the subsequent target stimuli would be distorted to favor the ear further from the adapting stimulus and the location of the perceived test stimulus would be shifted toward the further (less adapted) ear.

For location judgments made following adaptation at 0° azimuth/0° elevation, the simple adaptation model of Curthoys predicts that there should be no differences in the perceived location of target stimuli following adaptation as the level of the adapting stimulus was equal in both ears. However, the data in this study demonstrated a clear radial shift in perceived locations for all target stimuli in the vicinity of the adapting stimulus at azimuth 0°/elevation 0°. In addition, the distortion of perceived location encompassed both dimensions of azimuth and elevation, an entirely new observation. This upward shift in perceived location was most pronounced for adapting stimuli located at azimuth 30°/elevation 0°, particularly in the case of the targets located lateral of the adapting stimulus where the direction of the mean shift was rotated more vertically. Again, for the azimuth 30°/elevation 0° location a simple adaptation model would have predicted a straightforward shift toward the anterior midline for all target locations surrounding the adapting speaker.

Clearly, mislocalization produced by differential adaptation of each ear is insufficient to explain the pattern of mislocalization seen in the subjects in this study. In addition, a simple level-based adaptation model fails to account for cues such as interaural time difference cues (ITDs) and spectral cues. Unambiguous localization is most likely to be dependent on the integration of all these different cue types (Middlebrooks, 1992; see Carlile, 1996a, b for review). This is consistent with the neurophysiological findings in the mammalian auditory system. In the deep layers of the mammalian superior colliculus, the map of auditory space relies on neurones with relatively restricted spatial receptive fields which are shaped by the neurones selectivity for a small range of binaural and spectral cues (Palmer and King, 1982; King and Palmer, 1983; Carlile and Pettigrew, 1987; Middlebrooks, 1987; King *et al.*, 1990; Carlile and King, 1994a b). These physiological data, together with similar findings in the owl (see, e.g., Knudsen *et al.*, 1987; Olsen *et al.*, 1989),

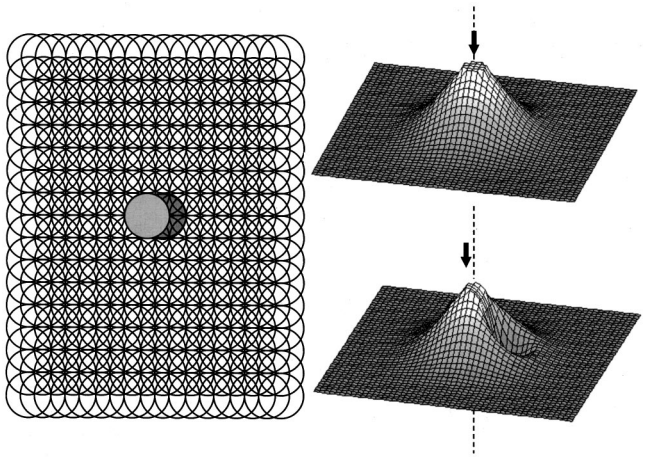


FIG. 6. Simple model of channel processing in auditory localization. Left panel: Array of overlapping channels with a zone of excitation (light gray overlying circle) and zone of previous adaptation (gray underlying circle) produced by different stimuli centered on particular spatial channels. Right upper panel: Ensemble output in the unadapted state with the mean population output indicated by the solid arrow. Right lower panel: Ensemble output where a subpopulation of channels has been adapted (bottom). Note that the locus of mean activity indicated by the solid arrow has shifted as a result of the down regulation of a subpopulation of channels by the adapting stimulus. Such a model also provides a powerful heuristic for a range of further auditory localization experiments.

indicate that these neural representations of auditory space are dependent on the integration of two or more types of cues. Such neurons may provide the basis for segregated processing of different regions of space. That is, their sensitivity to a limited range of convergent cues results in their being ''tuned'' to a specific location in space and the topographic organization of the neurons results in a neural map that is isomorphic with two-dimensional auditory space.

In the visual system there are psychophysical and neurophysiological data which suggest the presence of specific detectors or channels which account for the perception of particular features of the visual field (see, e.g., Barlow, 1990). It is not unreasonable to postulate that neurons coding for discrete regions of auditory space might also be indicative of some form of channel processing of auditory spatial perception. In other words, in the absence of a receptotopic representation of auditory space, it is likely that a computationally based representation of auditory space is comprised of a large number of neurons coding for different spatial locations (see Knudsen *et al.*, 1987). In such a scheme and particularly where the receptive fields are relatively large, the perception of auditory space might be dependent on the ensemble output of this neural array. Such a concept is illustrated in Fig. 6 where the topographic arrangement of neurones with spatially restricted receptive fields is represented by a mosaic of overlapping receptive fields. Borrowing from the visual literature, each neurone can be conceived of as a separate processing channel ''tuned'' to a particular area of space. Such a model may also provide a plausible explanation of the effects of auditory spatial adaptation. The down regulation of an individual channel by a prolonged stimulus exposure would change the resulting balance of outputs from the remaining channels, which in turn results in the sensory aftereffect observed (see Fig. 6; Mather, 1980; Wade, 1994).

Thus the down regulation of the auditory spatial channel centered on the adapting stimulus could subsequently affect the balance of outputs from the channel array for subsequent adjacent locations and result in the outward shift in the perception of nearby locations.

The assumptions of the model are straightforward and biologically plausible: namely, that localization processing is mediated through a large mosaic of overlapping spatial channels (receptive fields) and that the perception of location is derived from the output of the array. As discussed previously, examples of neurones that could form the biological substrate of such a system have been found in the mammalian and avian superior colliculus and in the avian MLD. Such models have previously been referred to as ''distribution shift'' models and provide a powerful framework for examining the characteristics of individual channels or receptive fields (see in particular Mather, 1980). The psychophysical data demonstrating the auditory spatial aftereffect are consistent with this kind of stimulus cue processing. In this model, the down regulation of the spatial channel centered on the adapting stimulus would subsequently affect the balance of outputs from the channel array and result in a shift in the perception of nearby locations away from the location of the adapting stimulus. From the current data we are unable to determine if the adaptation is likely to occur at the point of high-level convergence of the various localization cues, within the underlying processing streams, or a combination of the two. If the adaptation was occurring at the level of the processing of individual localization cues, it would also be of considerable interest to determine if each cue was equally adapted or if there was some form of differential adaptation across ITDs, ILDs, and spectral cues.

Such ''distribution'' shift models of auditory spatial aftereffects also provide a consistent basis for interpreting previous studies of auditory spatial aftereffects. It is of interest that this conception is also related to the early ideas behind the Köhler–Wallach theory of figural aftereffects when translated into a place code of auditory localization processing (see in particular Krauskopf, 1954) and are also related to the place models developed to explain localization ''funneling'' (see Thurlow *et al.*, 1965).

Such a ''channel'' model of localization processing may also be useful in interpreting experiments using multiple concurrent stimuli. In two recent brief reports, the localization of speech sounds with concurrent distracters or maskers have been described (Hawley *et al.*, 1999; Hartung and Braasch, 1999). In the case where noise has been used as a concurrent distracter, a displacement of the apparent location of the target speech away from the distracter has been reported (Hartung and Braasch, 1999). By contrast, when speech has been used as a distracter, a low incidence of mislocalization toward the distracter is reported (Hawley *et al.*, 1999). On the other hand, where nonspeech concurrent sounds were used as targets and distracters the effects were reported as being highly subject dependent (Wightman and Kistler, 1997). Of interest, the latter short report notes that one consistent effect of the concurrent distracter was on the disruption of the elevation component of the localization. Unfortunately, the significant differences in the methodology between these previous studies and the current study complicate any test of the proposed ''channel'' model.

Studies of the neural representation of auditory space in the owl (Knudsen and Konishi, 1978a, b) and ferret (King *et al.*, 1990) have also provided some neurophysiological evidence of lateral interactions between space-tuned neurons. This begs the question of the potential effects of lateral interactions between the proposed hypothetical localization channels. Such interactions could have important functional consequences by enhancing the contrast between specific spatial auditory locations. Contrast in this context might apply to separating out foreground sounds of interest (e.g., speech) from background noise. This kind of streaming of information has also been shown to be reliant on localization processing (Plomp and Mimpen, 1981).

ACKNOWLEDGMENTS

This work was supported by Australian Research Council Grant Nos. A79905421 and A49700049. The assistance of Anna Corderoy and Craig Jin with some of the statistics, data collection, and figures is warmly acknowledged.

Barlow, H. B. (**1990**). ''A theory about the functional role and synaptic mechanism of visual aftereffects,'' in *Vision: Coding and Efficiency*, edited by C. Blakemore (Cambridge University Press, Cambridge), pp. 363–375.

Bartlett, F. C., and Mark, H. (**1922**). ''A note on local fatigue in the auditory system,'' Br. J. Psych. **11**, 105–134.

Carlile, S. (**1996a**). ''Auditory space,'' in *Virtual Auditory Space: Generation and Applications*, edited by S. Carlile (Austin, TX), Chap. 1.

Carlile, S. (**1996b**). ''The physical and psychophysical basis of sound localization,'' in *Virtual Auditory Space: Generation and Applications*, edited by S. Carlile (Austin, Texas), Chap. 2.

Carlile, S., and King, A. J. (**1994a**). ''The integration of monaural and binaural cues in the generation of a neural representation of auditory space,'' Proc. Aust. Neurosci. Soc. **5**, 31.

Carlile, S., and King, A. J. (**1994b**). ''Monaural and binaural spectrum level cues in the ferret: Acoustics and the neural representation of auditory space,'' J. Neurophysiol. **71**, 785–801.

Carlile, S., Leong, P., and Hyams, S. (**1997**). ''The nature and distribution of errors in the localization of sounds by humans,'' Hear. Res. **114**, 179–196.

Carlile, S., and Pettigrew, A. G. (**1987**). ''Distribution of frequency sensitivity in the superior colliculus of the guinea pig,'' Hear. Res. **31**, 123–136.

Carlile, S., and Pralong, D. (**1994**). ''The location-dependent nature of perceptually salient features of the human head-related transfer function,'' J. Acoust. Soc. Am. **95**, 3445–3459.

Chandler, K. A. (**1961**). ''The effect of monaural and binaural tones of different intensities on the visual perception of verticality,'' Am. J. Psychol. **74**, 260–265.

Curthoys, I. S. (**1968**). ''The effect of binaural adaptation and masking on auditory localisation,'' Ph.D. thesis, Monash University (unpublished).

Elfner, L., and Perrott, D. R. (**1966**). ''Effect of prolonged exposure to a binaural intensity mismatch on the locus of a dichotically produced tonal image,'' J. Acoust. Soc. Am. **39**, 716–719.

Flugel, J. C. (**1921**). ''On local fatigue in the auditory system,'' Br. J. Psych. **11**, 105–134.

Hartung, K., and Braasch, J. (**1999**). ''Localization of distracted speech and noise in reverberant and anechoic environments,'' J. Acoust. Soc. Am. **105**, 1149.

Hawley, M. L., Litovsky, R. Y., and Colburn, H. S. (**1999**). ''Speech intelligibility and localization in a multi-source environment,'' J. Acoust. Soc. Am. **105**, 3436–3448.

King, A. J., Carlile, S., and Chevassut, T. J. T. (**1990**). ''Spatial organisation of auditory receptive fields in the ferret superior colliculus,'' Br. J. Audiol. **24**, 198.

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., and Palmer, A. R. (**1983**). ''Cells responsive to free-field auditory stimuli in guinea pig superior colliculus: Distribution and response properties,'' J. Physiol. (London) **342**, 361–381.
Knudsen, E. I., and Konishi, M. (**1978a**). ''Center-surround organization of auditory receptive fields in the owl,'' Science **202**, 778–780.
Knudsen, E. I., and Konishi, M. (**1978b**). ''A neural map of auditory space in the owl,'' Science **200**, 795–797.
Knudsen, E. I., Lac, S., and Easterly, S. D. (**1987**). ''Computational maps in the brain,'' Annu. Rev. Neurosci. **10**, 41–65.
Krauskopf, J. (**1954**). ''Figural aftereffects in auditory space,'' Am. Psychol. **67**, 278–287.
Lackner, J. R. (**1974**). ''Changes in auditory localisation during body tilt,'' Acta Oto-Laryngol. **77**, 19–28.
Leong, P. H. W., and Carlile, S. (**1998**). ''Methods for spherical data analysis and visualisation,'' J. Neurosci. Methods **80**, 191–200.
Mather, G. (**1980**). ''The movement aftereffect and a distribution-shift model for coding the direction of visual motion,'' Perception **9**, 379–392.
Middlebrooks, J. C. (**1987**). ''Binaural mechanisms of spatial tuning in the cat's superior colliculus distinguished using monaural occlusion,'' J. Neurophysiol. **57**, 688–701.
Middlebrooks, J. C. (**1992**). ''Narrow-band sound localization related to external ear acoustics,'' J. Acoust. Soc. Am. **92**, 2607–2624.
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.
Olsen, J. F., Knudsen, E. I., and Easterly, S. D. (**1989**). ''Neural maps of interaural time and intensity differences in the optic tectum of the owl,'' J. Neurosci. **9**, 2591–2605.
Palmer, A. R., and King, A. J. (**1982**). ''The representation of auditory space in the mammalian superior colliculus,'' Nature (London) **299**, 248–249.
Plomp, R., and Mimpen, A. M. (**1981**). ''Effects of the orientation of the speakers head and the azimuth of a noise source on the speech reception threshold for sentences,'' Acustica **48**, 325–328.
Taylor, M. M. (**1962**). ''The distance paradox of the figural aftereffect in auditory localization,'' Can. J. Phys. **16**, 278–282.
Thurlow, W. R., and Jack, C. E. (**1973**). ''Some determinants of localization-adaptation effects for successive auditory stimuli,'' J. Acoust. Soc. Am. **53**, 1573–1577.
Thurlow, W. R., Marten, A. E., and Bhatt, B. J. (**1965**). ''Localization aftereffect with pulse-tone and pulse-pulse stimuli,'' J. Acoust. Soc. Am. **37**, 837–842.
Wade, N. J. (**1994**). ''A selective history of the study of visual motion aftereffects,'' Perception **23**, 1111–1134.
Wightman, F. L., and Kistler, D. J. (**1992**). ''The dominant role of low-frequency interaural time-differences in sound localization,'' J. Acoust. Soc. Am. **91**, 1648–1661.
Wightman, F. L., and Kistler, D. J. (**1997**). ''Sound localization in the presence of multiple distracters,'' J. Acoust. Soc. Am. **101**, 3105.

The effects of production and presentation level on the auditory distance perception of speech

Douglas S. Brungart[a] and Kimberly R. Scott
Air Force Research Laboratory, Human Effectiveness Directorate, 2610 Seventh Street, Wright-Patterson AFB, Ohio 45433-7901

(Received 14 August 2000; revised 20 April 2001; accepted 23 April 2001)

Although both perceived vocal effort and intensity are known to influence the perceived distance of speech, little is known about the processes listeners use to integrate these two parameters into a single estimate of talker distance. In this series of experiments, listeners judged the distances of prerecorded speech samples presented over headphones in a large open field. In the first experiment, virtual synthesis techniques were used to simulate speech signals produced by a live talker at distances ranging from 0.25 to 64 m. In the second experiment, listeners judged the apparent distances of speech stimuli produced over a 60-dB range of different vocal effort levels (production levels) and presented over a 34-dB range of different intensities (presentation levels). In the third experiment, the listeners judged the distances of time-reversed speech samples. The results indicate that production level and presentation level influence distance perception differently for each of three distinct categories of speech. When the stimulus was high-level voiced speech (produced above 66 dB SPL 1 m from the talker's mouth), the distance judgments doubled with each 8-dB increase in production level and each 12-dB decrease in presentation level. When the stimulus was low-level voiced speech (produced at or below 66 dB SPL at 1 m), the distance judgments doubled with each 15-dB increase in production level but were relatively insensitive to changes in presentation level at all but the highest intensity levels tested. When the stimulus was whispered speech, the distance judgments were unaffected by changes in production level and only decreased with increasing presentation level when the intensity of the stimulus exceeded 66 dB SPL. The distance judgments obtained in these experiments were consistent across a range of different talkers, listeners, and utterances, suggesting that voice-based distance cueing could provide a robust way to control the apparent distances of speech sounds in virtual audio displays.
[DOI: 10.1121/1.1379730]

PACS numbers: 43.66.Qp, 43.71.Bp [LRB]

I. INTRODUCTION

After more than a century of research on auditory localization, relatively little is known about the ability of human listeners to judge the distances of sound sources. It is clear, however, that the mechanisms involved in auditory distance perception are complicated. Although a number of different acoustic cues could theoretically be used to determine the distance of a sound source (Coleman, 1963; Zahorik, 1998), different cues tend to dominate in different listening environments. This requires listeners to adapt their distance estimation strategies to efficiently incorporate the distance cues available in each listening situation. Auditory distance judgments therefore tend to vary widely with the characteristics of the stimulus, the characteristics of the surrounding room, and the experiences and expectations of the listener [see Zahorik (1998) for a recent review of the literature on auditory distance perception].

Under these circumstances, one would expect distance perception to be especially accurate for a stimulus that is highly familiar to the listener and has acoustic properties that vary systematically with the intensity of the source. This would allow a listener to estimate the output power of the source from its acoustic properties and determine its distance by comparing its apparent output level to the intensity of the sound at the locations of the ears.

Human speech is an example of this type of audio signal. The acoustic properties of speech change systematically with the output level of the talker, so listeners should be able to estimate the distance of a talker in the free field by comparing the apparent production level of the talker's voice to the intensity of the speech signal reaching their ears. Although little is known about the ability of listeners to determine the intensity of a talker from the acoustic properties of a speech signal, the systematic changes that occur in the human voice when the output level of the talker increases are well documented in the speech production literature. The quietest form of speech, whispering, is readily identified from its lack of voicing (Pickett, 1956). As a voiced utterance increases in intensity from conversational speech to shouted speech, the increased pressure on the vocal folds leads to an increase in the fundamental frequency $F0$ (Lienard and Benedetto, 1999), as well as an increase in the first formant frequency $F1$ and an increase in the ratio of high- to low-frequency energy in the speech spectrum (Grandstrom and Nord, 1992; Junqua, 1993; Rostolland, 1982; Shulman, 1989; Sluijter and Heuven, 1996; Traun-

[a]Author to whom correspondence should be addressed. Electronic mail: douglas.brungart@wpafb.af.mil

müller and Eriksson, 2000). These acoustic changes make it relatively easy to distinguish between low-level conversational speech and high-level shouted speech. Although there are many ways to characterize the vocal effort of a talker, for the purposes of this article we will refer to it as the production level of the speech and operationally define it as the rms power of the speech signal, in dB SPL, measured 1 m directly in front of the talker's mouth. This metric, which is assigned the variable name V, does not directly measure the stress on the vocal folds of the talker, but it does capture the changes that occur in the acoustic properties of speech when the output level of the talker increases.

The second factor that is likely to influence the perceived distance of speech is the intensity of the signal reaching the listener's ears. In a free-field listening environment, this intensity decreases 6 dB for each doubling in the distance of the source (Coleman, 1963). It therefore provides a powerful cue for determining changes in the relative distance of an unknown sound source (Mershon and Bowers, 1979). There are many ways to measure the intensity of the speech signal at the location of the listener. For the purposes of this article, we refer to this intensity level as the presentation level of the speech (P) and define it as the rms power, in dB SPL, that would occur at the location of the center of the listener's head if the listener were removed from the room. In theory, this presentation level can be compared to the apparent production level of the speech in order to estimate the distance of the talker. Speech sounds with high presentation levels but low apparent production levels (e.g., intense whispers) indicate nearby talkers, while speech sounds with low presentation levels but high apparent production levels (low-level shouts) indicate more distant talkers. Because each combination of production and presentation level corresponds to a unique talker distance, it should be possible for listeners who are able to accurately judge the absolute production and presentation levels of speech signals to make reasonably accurate judgments about the distances of the talkers.

This distance judgment ability has been verified in a number of experiments examining the apparent distances of live talkers in nonreverberant environments. Von Bekesy (1949) and Gardner (1969) examined the apparent acoustic distances of conversational-level talkers in anechoic chambers. Cochran *et al.* (1968) examined the perceived distance of a live talker relative to a 15-m reference point in a 5-acre grass field. The overall distance judgments from these studies were similar (Fig. 1), and all three indicated that listeners could make reasonably accurate estimates about the distances of live talkers in the free field. The only consistent misperception was a tendency to underestimate the ranges of distant talkers in the anechoic room.

These studies confirm that listeners are able to determine the distances of live talkers, but the considerable methodological difficulties involved in conducting experiments with live talkers make it difficult to determine which cues the listeners were using to make these distance judgments. The use of a live talker as a sound source requires either the talker or the listener to be moved between trials, introducing a host of possible confounding cues into the experiment. It also makes it difficult to prevent unintended variations in the speech resulting from the talker's knowledge of the distance to the listener. Furthermore, it is impossible to control the acoustic attributes of the speech from trial to trial with a live talker, so live talkers cannot be used to isolate the effects of presentation level and production level on the perception of speech. Because of these difficulties, most of the recent experiments examining the distance perception of speech have used prerecorded speech played over loudspeakers or headphones as the stimulus.

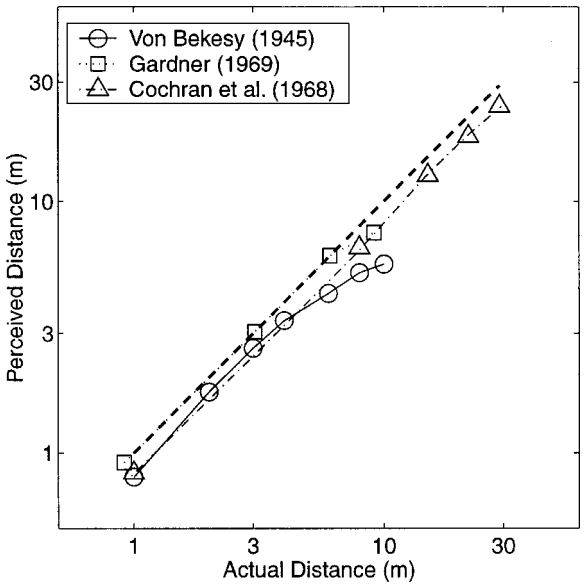


FIG. 1. Summary of results of experiments examining the apparent auditory distances of live talkers in a free-field environment. The results represent the means across 2 (Von Bekesy), 10 (Gardner), and 22 (Cochran *et al.*) listeners. The heavy dashed line represents ''correct'' responses where the distance judgments would match the actual location of the talker.

Some of these experiments have directly compared the perceived distance of a single prerecorded speech utterance to other types of prerecorded sounds. In general, these experiments have found little difference between speech and other types of signals. Nielson (1991), for example, compared the perceived distances of pop music, speech, guitar music, and white noise in a reverberant room. Zahorik (1998) compared speech to white noise in binaural simulations of a large lecture hall. Neither of these studies found any significant overall differences between speech and the other stimuli tested.[1] However, these experiments were limited because they excluded some important features of speech that may be crucial to auditory distance perception. The experiments did not vary the production level of the talker, and they both included reverberation cues that may have reduced the importance of source familiarity in the distance judgments. Thus, it would be premature to infer from these results that there is no difference between the distance perception of speech and the distance perception of other types of sounds.

A number of other experiments have used prerecorded utterances played over loudspeakers to determine the influence of presentation level on the distance perception of speech at a fixed production level. All of these experiments have shown that the perceived distance of speech systemati-

cally decreases as the stimulus level is increased (Von Bekesy, 1949; Cochran *et al.*, 1968; Gardner, 1969; Begault, 1987). The amount of increase in presentation level required to halve the perceived distance of the sound varied substantially, from 6 to 9 dB in Begault's study to 20 to 30 dB in Gardner's study.

Only a handful of studies have directly examined the impact of production level on the perceived distance of speech. Mershon and Philbeck (1991) used prerecorded whispered, conversational, and shouted speech played back over a loudspeaker to determine the influence that production level has on distance perception when the presentation level of the stimulus level is fixed. The results were dramatic. Whispered speech was perceived at less than half the distance of conversational speech, and shouted speech was perceived more than three times as far away as conversational speech. In a more recent experiment, Traunmüller and Eriksson (2000) asked listeners to determine the distance between the talker and addressee from speech recordings of talkers conversing with an experimenter in a large open field. These judgments, which were based on the listener's assessments of the appropriate conversational distance for each production level, increased from 0.47 to 31 m as the actual distance between the talker and the experimenter increased from 0.3 to 187.5 m. Both of these results indicate that the apparent distance of a speech signal increases when its production level increases at a fixed presentation level.

Overall, there is sufficient evidence to conclude that both production level and presentation level play a role in determining the perceived distance of speech. Furthermore, there is reason to believe that the acoustic changes that occur in speech as the talker increases his or her level of output provide enough information to allow listeners to make more accurate distance judgments about the location of a live talker than an unfamiliar sound source. However, the studies that have systematically examined the distance perception of speech have used stimuli with a limited range of production levels spoken by only one or two different talkers. Very little is known about the strategies listeners use to integrate perceived production level and presentation level into a single distance judgment, or how distance judgments vary across different talkers and utterances. This article describes the results of a series of three experiments that used prerecorded speech samples presented over headphones to examine the perceived distance of speech. The experiments differed from previous studies of this type both in the use of virtual synthesis techniques to precisely control the production level and presentation level of each stimulus and in the use of a wider range of production levels, talkers, and utterances in the speech stimuli. The next section describes the processing and calibration procedures used to generate the virtual speech stimuli used in the three experiments.

II. METHODS

A. Recording of virtual speech samples

1. Apparatus

The speech samples were collected in a large anechoic chamber at Wright-Patterson Air Force Base. In one corner of the chamber, a microphone (B&K 4144 1-in. pressure microphone) was mounted on an adjustable stand. The output of the microphone was connected to a variable-gain microphone power supply (B&K 5935) located in an adjacent control room, and was passed through a 10 Hz–20 kHz band-pass filter (Krohn-Hite 3100) before terminating at the input of a 16-bit, 50-kHz A/D converter (Tucker-Davis DD1). In addition, a loudspeaker connected to the output of a D/A converter was located near the center of the anechoic chamber and used to provide timing information to the talkers. The entire recording process was controlled by a Pentium II-based PC in the control room.

2. Talkers

A total of ten paid volunteers, five male and five female, ranging in age from 19 to 54, were used as talkers in the measurements.

3. Utterances

Four different utterances were recorded at each production level for each of the ten talkers:

(1) ''Don't ask me to carry an oily rag like that.''
(2) ''Over here.''
(3) ''Threat.''
(4) ''Warning.''

The first utterance was chosen because it is a calibration sentence spoken by all talkers in the Texas-Instruments/MIT (TIMIT) speech corpus. The second was believed to be an appropriate phrase for determining the distance of a talker at an unknown location. The third and fourth phrases were selected because of their potential utility in a virtual audio display system.

4. Calibration

Prior to measuring the speech samples from each subject, a 1-kHz, 94-dB calibrator was placed on the microphone and used to record a 5-s calibration tone. This calibration tone was used to determine the sound pressure levels of the subsequent measurements.

5. Procedure

Prior to each set of measurements, the microphone was adjusted to the height of the talker's mouth and the talker used a 1-m rod to position his or her chin exactly 1 m from the microphone. Then the talkers were instructed to begin speaking in their quietest whisper and to increase the level of the speech slightly on each repetition until they were unable to whisper any louder. The experimenter then left the room and began measuring the speech samples. The procedure for each measurement was as follows:

(1) The loudspeaker prompted the talker with a recording of the first sentence, ''Don't ask me to carry an oily rag like that,'' followed by a beep.
(2) At the sound of the beep, the talker repeated the phrase in the appropriate voice, and the D/A converter recorded the talker's speech for 250 000 samples (5 s).

(3) Steps 1 and 2 were repeated for each of the other three utterances, and 100 000 samples (2 s) of the talker's response were recorded for each of these utterances.
(4) A graph of the four measured speech samples was plotted on the screen of the control computer and examined by the experimenter for any signs of clipping. If clipping occurred, the experimenter adjusted the gain of the microphone power supply down by 10 dB, and asked the talker to repeat steps 1–3 at the same output level. If no clipping occurred, the speech samples were saved, and the talker was asked to increase the output level slightly.
(5) Steps 1–4 were repeated until the subject was unable to whisper any louder. Then the subjects were instructed to repeat the utterances in their quietest conversational (voiced) tone and to slightly increase the level of their speech on each repetition, and steps 1–4 were repeated until the subjects felt they could not talk any louder without shouting. Finally, the subjects were asked to repeat the utterances in their quietest shouted voice and to become progressively louder with each repetition, and steps 1–4 were repeated until the subjects were unable to shout any louder.

The number of samples recorded in the whisper, conversation, and shout modes were determined by the dynamic range of the individual talkers and the rate at which they increased their production level. The number of whispered samples ranged from 4 to 8 per subject, the number of spoken samples from 6 to 13, and the number of shouts from 5 to 22.

B. Processing of speech signals

1. Truncation and scaling of speech samples

After the data collection, the individual speech samples were visually inspected and truncated to the beginning and end of the speech utterance. Then the 94-dB calibration tone and the gains of the microphone power supply were used to scale the different utterances into absolute pressure units (pascals). Finally, the overall production level of each utterance was determined from the RMS power of the truncated speech sample. The production levels of the speech typically ranged from approximately 35 dB SPL at 1 m for low-level whispers to more than 95 dB SPL for high-level shouts.

2. HRTF processing

In order to make the headphone presentations as realistic as possible, the speech signals were processed with the individualized head-related transfer functions (HRTFs) of each listener prior to the headphone presentation. The HRTFs were measured at a fixed distance of 1 m in an anechoic chamber with a blocked-meatus Golay-code technique (Moller *et al.*, 1995; Foster, 1986). Miniature microphones (Emkay FG-OA) embedded in soft rubber swimmer's earplugs were inserted into the listener's ears, and HRTFs were measured with a loudspeaker located 1 m directly in front of the listener and 1 m directly to the left of the listener (measured from the center of the listener's head.) The headphones used in the experiment (Sennheiser HD-520) were then placed over the in-ear microphones in order to measure the transfer function from the headphone inputs to the listener's ears. Finally, the listeners were removed and the loudspeaker response at the location of the center of the head was measured with a free-field microphone. The headphone and speaker frequency responses were smoothed with a $\frac{1}{12}$-oct filter, inverted, and used to generate minimum-phase FIR filters correcting for the effects of the headphones and loudspeaker on the HRTF measurements. These inverse filters were convolved with head-related impulse responses recorded with the in-ear microphones, and the resulting filters were truncated to produce 512-point finite-impulse-response digital filters. The characteristics of the speech signals presented over headphones were therefore matched to the characteristics of the speech signals that would occur from a live talker 1 m in front or to the side of the listener.

3. Calibration

The overall levels of the headphone-presented speech stimuli were carefully calibrated to match the levels of live talkers in the free field. First, the loudspeaker in the anechoic chamber was used to generate a 1-kHz tone at 84 dB SPL at a distance of 1 m (as measured by a B&K 4144 microphone). Then the miniature microphones were inserted into the listener's ears and the listener was positioned with the loudspeaker 1 m directly in front of the center of the head. This configuration was used to measure the sound pressure produced by the 84-dB, 1-kHz tone at the location of the right in-ear microphone. The loudspeaker was then turned off, the headphones were placed over the in-ear microphones, and the voltage of a 1-kHz headphone signal was adjusted until it produced the same sound pressure at the right in-ear microphone as the 84-dB free-field sound. Prior to generating the speech stimuli, the 94-dB calibration tone measured before each set of speech measurements was processed with the 1-m HRTF and the gain was adjusted according to this 84-dB calibration voltage to ensure that the HRTF-processed speech stimuli presented over headphones matched the intensity of a live talker at the desired distance in the free field.

C. Measurement of auditory distance judgments

The speech stimuli used in each experiment were prerecorded on a portable DAT player (Sony TCD-D8) and presented to listeners who were seated in a lawn chair in the center of a large mown grass field (Fig. 2). The listeners were asked to judge the apparent distances of the speech stimuli based on the locations of nine visual markers, numbered 1–9, placed at distances of 0.5, 1, 2, 4, 8, 16, 32, 64, and 128 m from the listener. They were also instructed to use decimal numbers to respond at locations between the distance markers, and to use ''0'' to represent distances one-half as far away as the 0.5-m visual marker labeled ''1.'' The responses were written on a numbered answer sheet and later transcribed into digital form by the experimenters. The data were later compared to the stimuli to identify the locations of any skipped or missing trials. Dropped trials occurred in 15

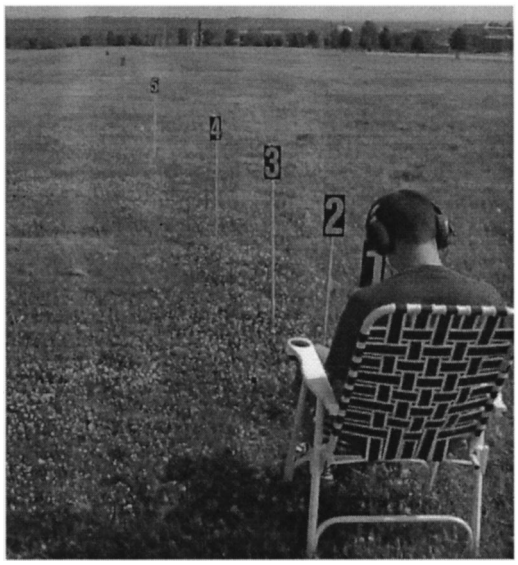


FIG. 2. Open field where the listeners judged the perceived distances of the speech sounds. Nine visual markers, numbered 1–9, were placed at locations 0.5, 1, 2, 4, 8, 16, 32, 64, and 128 m from the listener's position and used as reference points for the listeners' responses.

trials out of 7776 in the first experiment, 22 trials out of 8316 in the second experiment, and 12 trials out of 3780 in the third experiment.

III: EXPERIMENT 1: EFFECTS OF SIMULATED DISTANCE ON THE APPARENT DISTANCE OF SPEECH

In the first experiment, the production level and presentation levels of the speech stimuli were adjusted to simulate the voice characteristics of a live talker over a wide range of distances (from 0.25 to 64 m) and a relatively narrow range of presentation levels (48 d to 60 dB SPL). The results provide some insight into the abilities of listeners to judge the distances of live talkers under free-field listening conditions.

A. Experiment design

1. Participants

A total of eight paid listeners (four male, four female) were used in the experiment. All had normal hearing (15 dB HL from 500 Hz to 6 kHz) and their ages ranged from 21 to 55. Each had previous experience in psychoacoustic experiments, but none had previously participated in distance perception experiments.

2. Stimuli

The stimuli consisted of prerecorded speech samples processed to simulate live talkers at one of nine distances and presented over headphones to the listeners. The stimuli were presented in random order in a full-factorial design on five different factors: talker, utterance, direction, presentation level, and simulated distance. Each of these factors is described in more detail below.

a. Talker. Six of the ten talkers used to collect the speech samples were chosen for the experiment based on their overall dynamic range and number of different levels of production level available. Three were male talkers and three were female talkers. Three of the six talkers also participated as listeners in the experiment.

b. Utterance. Three utterances were used in the experiment: ''Over Here,'' ''Threat,'' and ''Warning.''

c. Direction. The speech sample on each trial was processed either with the HRTF measured 1 m directly in front of the listener (front) or with the HRTF measured 1 m directly to the left of the listener (side).

d. Presentation level. Three different presentation levels were used for the experiment: 48, 54, and 60 dB SPL. Each level represents the rms power of the speech utterance that would occur at the location of the center of the listener's head if the listener were removed.

e. Simulated distance. The primary experimental variable in each stimulus presentation was the simulated distance of the talker. A total of nine simulated distances were used in the experiment: 0.25, 0.5, 1, 2, 4, 8, 16, 32, and 64 m. For each stimulus presentation, the production level required by a talker at the desired distance to generate the desired presentation level at the location of the listener was determined from the inverse-squared power law of sound. For example, a desired presentation level of 60 dB SPL by a talker at 2 m would require a talker to produce speech at a level of 66 dB SPL at 1 m. A presentation of 48 dB SPL at 0.25 m would require speech at 36 dB SPL at 1 m. Once the production level was determined, the speech sample of the desired utterance by the desired talker produced at the level closest to the required level was selected and attenuated or amplified according to the simulated distance of the talker. For instance, the 60-dB presentation level speech at 2 m might be generated by selecting the speech sample produced at 65 dB SPL at 1 m and attenuating it by 6 dB. Note that this procedure precisely controls the simulated distance of the talker, but because speech samples were not available at every dB production level it inevitably results in some deviations from desired presentation level. In this example, actual presentation level was 59 dB, 1 dB lower than the nominal presentation level of 60 dB. The average presentation level was, however, within 0.1 dB of the nominal presentation level at each of the three levels tested. The trials produced at a nominal presentation level of 48 dB had an actual average level of 48.1 dB (σ=1.5), those at a nominal level of 54 dB were at an average level of 54.1 dB (σ=1.6), and those at a nominal level of 60 dB were at an average level of 60.0 dB (σ=1.5). After the speech samples were selected and scaled according to the desired stimulus distances, they were processed with the appropriate HRTF according to the desired direction of the talker before being presented to the listener over headphones.

3. Experiment design

In order to examine the importance of speaker familiarity on the results, two different conditions were tested. In the primary condition, the stimuli from the different talkers were presented in random order and the listeners were given no feedback about the distances of the talkers during the experiment. In the secondary condition, the stimuli were blocked together into groups of 100 consecutive trials from the same

talker and the listeners were exposed to the vocal characteristics of the talker across the entire range of production levels tested prior to judging the distances of each block of stimuli. The data collection for the primary condition was divided into two sessions on different days, and the data collection for the shorter secondary condition was conducted in a single session. Each session took 45 min to 1 h to complete. Half the listeners participated in the primary condition first, and half participated in the secondary condition first.

a. Primary condition. The trials in the primary condition were randomly ordered in a full-factorial design (6 talkers $\times 3$ utterances $\times 2$ directions $\times 3$ presentation levels $\times 11$ simulated distances) for a total of 1188 trials for each subject. The trials were blocked together into groups of 100 trials, and a tone and a short pause were recorded on the tape to indicate the beginning of each block. Note that the 11 distances used in the experiment design included the nine standard distances plus additional replications of the 0.25- and 0.5-m distances processed with HRTFs measured at 0.25 and 0.5 m rather than the standard 1.0-m HRTF. The results from these near-field stimuli are not presented here and have been eliminated from all data analyses.

b. Secondary condition. The trials in the secondary condition were similar to those in the primary condition, except that the same talker was used across each block of 100 trials. Also, only four of the six talkers were tested in the second condition. The total number of trials (4 talkers $\times 3$ utterances $\times 2$ directions $\times 3$ presentation levels $\times 11$ simulated distances) was increased from 792 to 800 by adding two additional trials for each talker in order to ensure that each talker was used for two full blocks of 100 trials. The results of the duplicated trials were averaged in the analysis of the data. In addition, the beep at the beginning of each block of trials in the primary condition was replaced by a series of 12 presentations of the word ''Warning'' by the talker used in that block at a simulated distance of 1 m and production levels increasing from 35 to 90 dB SPL in 5-dB steps. Thus, the listeners were provided with a simulated talker at 1 m across the entire range of production levels used in the experiment prior to judging the distances of the speech utterances. As a result of a technical error, data were collected for only seven of the eight listeners in the secondary condition.

B. Results and discussion

Before presenting the results, it is appropriate to briefly discuss how the responses obtained in the experiment should be interpreted. When making their responses, the subjects were instructed to estimate the distances of the talkers according to the logarithmically spaced visual markers in the field where the testing was conducted. In this type of task, there is always some question whether the listeners are reporting the actual locations where they perceived the sounds or they are simply mapping the variations in the most salient cue in the stimuli to the range of available responses. In addition, whenever virtual stimuli are used in the experiment there is always some concern that the virtual stimuli will be heard inside the head rather than externalized. Although none of the listeners indicated that they were having difficulty judging the distances of the stimuli because they were located ''inside the head,'' it is impossible to know precisely what strategies they were using when they made their responses. There is, however, some evidence in this experiment that the listeners were not simply mapping the stimulus attributes to the range of possible distances. The subjects generally did not use the entire range of available distances when making their responses: only about 2.5% of the responses were assigned to the two farthest visual markers (64 and 128 m), and only about 4% of the responses were assigned to the closest possible response distance (0.25 m). In addition, approximately 11% of the responses were assigned to fractional distances between the locations of the visual markers. As discussed later, the patterns of responses were also quite similar across the listeners used in the study, which probably would not occur if the listeners were simply mapping some attribute of the stimuli to a range of response values. Thus, there is reason to believe that the listeners were responding with visual locations that, in some way, matched the apparent distances of the speech stimuli. Note that the terms response distance, apparent distance, and perceived distance are used interchangeably in this discussion to represent the locations of the listeners' responses.

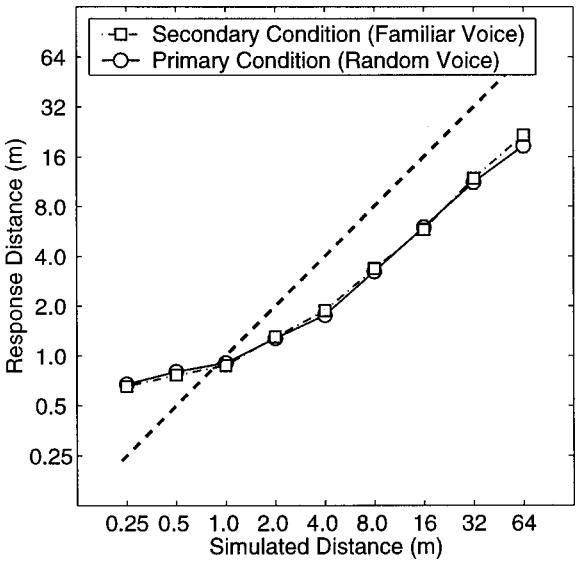


FIG. 3. Mean distance judgments in experiment 1 as a function of the simulated distances of the speech stimuli. The results are shown separately for the primary condition, where a random talker was used in each stimulus presentation, and the secondary condition, where the listeners were trained on the characteristics of the talker's voice prior to each block of trials. Note that data are shown only for the seven listeners and four talkers that were used in both conditions. Each data point represents the average response distance across all of the trials presented at that particular simulated distance with one of these seven listeners and four talkers. The data are shown on a log-log scale, and the dashed line represents ''correct'' responses where the distance judgments would match the simulated distance of the stimulus.

1. A priori exposure to talker's voice

In Fig. 3, the mean response distances at each simulated distance are shown on a log–log scale for the four talkers and seven listeners who were used in both the primary condition (solid line) and the secondary condition (dashed–dotted line) of the experiment. These curves show that the responses at each simulated distance were almost identical in the two conditions. Evidently the exposure to the entire

range of different production levels of the talker's voice provided at the beginning of each block of 100 trials in the secondary condition had no discernible effect on the perceived distances of the speech signals. A six-factor, repeated-measures ANOVA (on the five basic experimental factors described in the methods section plus the condition number) confirmed that there was no significant difference between the responses made with and without prior exposure to the talker's voice [$F(1,6) = 0.12$, $p = 0.74$]. In fact, the responses across the two conditions were similar enough to suggest that listeners' distance judgments were highly repeatable when the same stimulus was presented on different days. It is certainly puzzling that the listeners were unable to glean any useful information from exposure to the talker at a known distance prior to each block of trials. It is true that all of the listeners in the experiment had at least heard each of the talkers in the experiment speak prior to the data collection, and that some of the listeners were very familiar with the talkers (three of the listeners were also used as talkers in the experiment). Half of the listeners also heard the stimuli from the secondary condition prior to participating in the primary condition. Thus it is probably not reasonable to assume that the listeners were completely naive about the characteristics of the talkers in the primary condition. It is nonetheless surprising that no performance advantage was seen in the secondary condition where the listeners were exposed to the vocal characteristics of each talker immediately prior to data collection and had the further advantage of listening to the same talker throughout each block of 100 consecutive trials. Apparently, the short-term exposure to the talker's voice characteristics that the listeners received in the secondary condition of this experiment did not have much influence on the perceptual mechanisms that they used to determine the perceived distances of the stimuli. Because the data were nearly identical in the two conditions, the remaining discussion of experiment 1 will be restricted to the data from the primary condition.

2. Simulated distance

The data from the four talkers and seven listeners shown in Fig. 3 clearly show that the response distance increased systematically with the simulated distance of the speech stimulus in both the primary and secondary conditions of experiment 1. A more detailed analysis of all the data from the primary condition reveals that the distance judgments were substantially different for voiced and unvoiced speech (Fig. 4). Whispered speech was always perceived at a distance of approximately 0.8 m regardless of the simulated distance of the talker, while the apparent distance of voiced speech increased systematically with simulated distance across the entire range of distances tested.

Although the response distances for voiced speech varied over a relatively wide range—from 0.75 m to more than 20 m—the distance judgments did not match the simulated distances of the speech signals. The listeners tended to overestimate the range of nearby talkers (closer than 1 m) and underestimate the range of more distant talkers (farther than 1 m away). This type of compression has been reported in the responses of numerous other auditory distance experiments. Traunmüller and Eriksson (2000) reported a similar tendency to overestimate the distance of speech produced by talkers within 1.5 m of the addressee, and underestimate the distances of speech produced by talkers farther than 1.5 m away from the addressee. Zahorik (1998) analyzed the results of seven previous auditory distance perception studies and found that the range of responses was, on average, only 65% as large as the range of stimulus distances when both were plotted on a log–log scale. Thus, it is not surprising that the range of responses was somewhat smaller than the range of simulated talker distances in this experiment. It is interesting to note that the distance judgments in this experiment were substantially less accurate than those reported in earlier experiments using live talkers as a stimulus (Fig. 1). In part, this discrepancy can be explained by the inadequacies of the simulated speech signals used in this experiment. While every effort was made to make these stimuli as realistic as possible, the HRTFs used in this study, which were measured anechoically at a fixed distance from the listener, may not have captured all of the cues that were available in the live talker experiments. It is also possible that some of the potential extraneous cues that are difficult to control in live talker experiments, such as unintentional variations in the production level of the talker with the distance of the listener, contributed to the generally better performance found in the earlier live-speech studies. Differences in the response methods used in the live talker studies might also have had an effect. It would be methodologically difficult to produce live-speech stimuli with the same level of precise control of both production level and presentation level obtained in this experiment, or to use the same visual response method used in this experiment with a stimulus generated by a live talker.

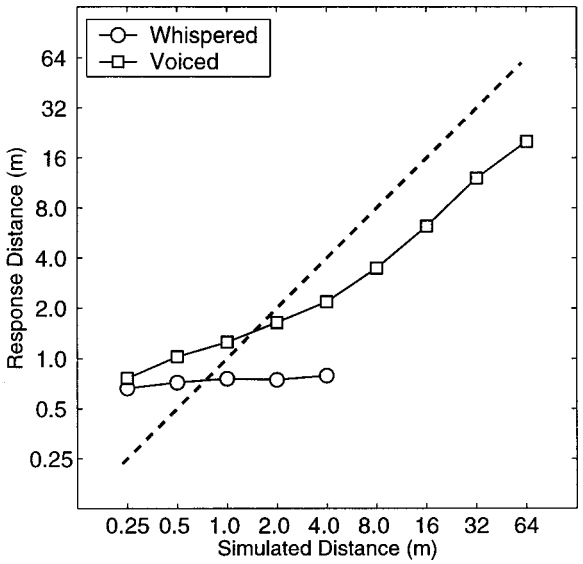


FIG. 4. Effects of voicing on the distance perception of speech in experiment 1. Data are similar to Fig. 3 except that the mean distance judgments have been plotted separately for the unvoiced trials where whispered speech was used as the stimulus and the voiced trials where conversational or shouted speech was used for the stimulus. Each data point represents the average response distance across all of the trials presented at that particular simulated distance with voiced speech or whispered speech. Note that the whispered speech was perceived at about 0.8 m across all the simulated distances.

TABLE I. Analysis of variance in primary condition of experiment 1. Note: This table shows only the significant factors and interactions ($p<0.05$) from a five-factor repeated-measures analysis of variance in the primary condition of experiment 1.

Factor	DF	SS	DFE	SSE	F	P	ω^2
D (distance)	8	21 445	56	927	162.00	0.0001	0.633
T (talker)	5	216	35	43	35.51	0.0001	0.006
L (level)	2	247	14	329	5.25	0.0199	0.006
U (utterance)	2	61	14	15	28.02	0.0001	0.002
S (side)	1	71	7	28	17.89	0.0039	0.002
T×U	10	28	70	74	2.61	0.0094	0.001
T×D	40	479	280	337	9.94	0.0001	0.013
U×D	16	83	112	114	5.06	0.0001	0.002
D×L	16	114	112	225	3.56	0.0001	0.002
T×U×D	80	123	560	524	1.64	0.0008	0.001
T×D×L	80	279	560	546	3.57	0.0001	0.006
U×D×L	32	76	224	233	2.29	0.0002	0.001
T×U×D×L	160	249	1120	1086	1.61	0.0001	0.003

It is therefore not possible to know for certain how the distance judgments for the virtual stimuli in this experiment would compare to those for equivalent live-talker stimuli. The results of this experiment should, however, provide a reasonably accurate indication of the effects that production level and presentation level have on the perception of live speech signals.

Changes in the simulated distance had a considerably larger influence on the perceived distance of relatively distant voiced speech than on the perceived distance of nearby voiced speech. In fact, the log–log slope of the stimulus–response curve for sources beyond 4.0 m (approximately 0.9) was more than twice as large as the slope of the stimulus–response curve for voiced speech closer than 1.0 m (approximately 0.35). It is interesting to note that this discontinuity roughly corresponds to the transition between conversational and shouted speech that occurred at a simulated distance of 4.0 m in the experiment: 100% of the voiced speech samples closer than 4.0 m were labeled conversational by the talkers, and 78% of the voiced speech samples farther than 4.0 m were labeled shouted by the talkers. The results of experiment 2 will show that the apparent distance of high-level voiced speech is influenced substantially more by changes in production level than the apparent distance of low-level voiced speech.

3. Other experimental factors

In order to measure the relative contributions of different experimental parameters to the distance judgments, the data were analyzed with a five-factor repeated-measures ANOVA (Table I). From this ANOVA the parameter ω^2, which is the proportion of the response variance that can be explained by each experimental variable, was used to estimate effect size for each significant factor and interaction (Cohen, 1977). The results show that the simulated distances of the speech stimuli ($\omega^2=0.663$) dominated the apparent distance judgments in this experiment, accounting for more than 100 times more of the variance than any other main effect or interaction. However, all four of the other main effects were also significant at the $p<0.02$ level. These effects are described in more detail later.

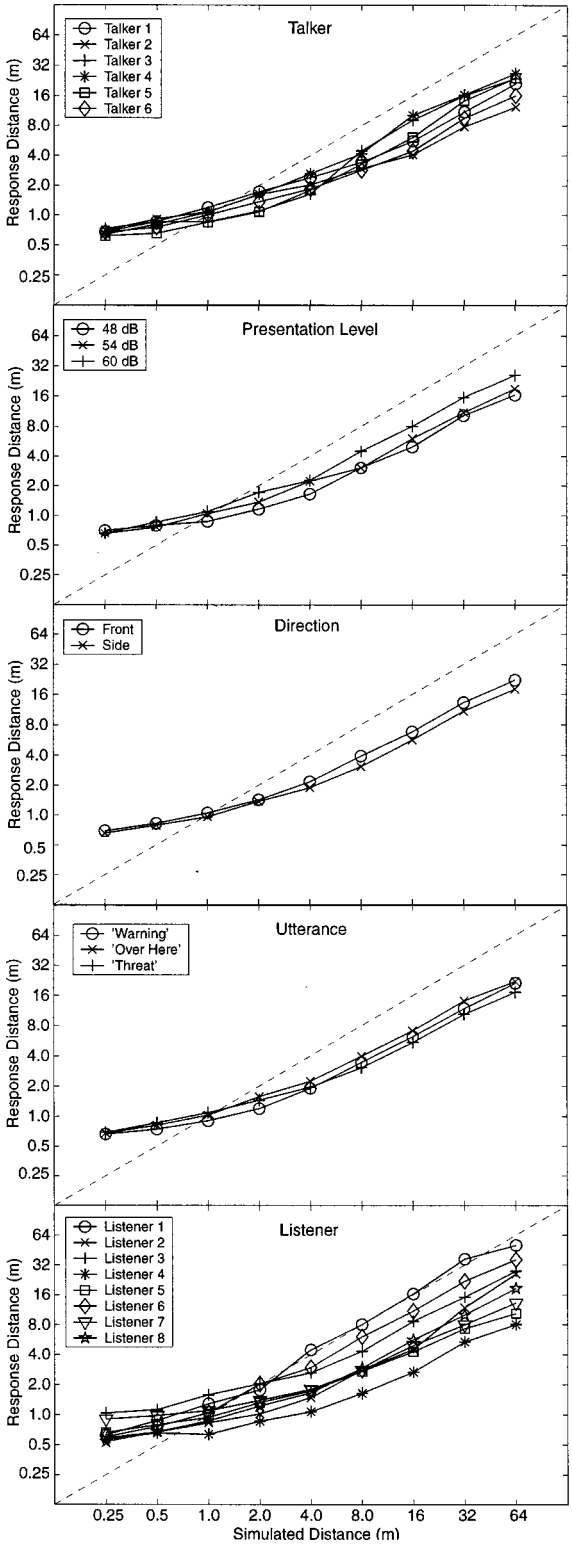


FIG. 5. Effects of presentation level, direction, utterance, and listener on the distance perception of virtual speech in experiment 1. In each panel, the response data have been averaged across all trials with the factor value shown in the legend and the simulated distance value indicated by the *x* axis. The data are shown on a log–log scale, and the dashed lines indicate ''correct'' responses where the distance judgments would match the simulated distance of the stimulus.

a. Talker. The identity of the talker was the second most influential factor in determining the perceived distances of the speech samples (Fig. 5, top panel). Talker had little impact on the mean overall response distances: the main effect

of talker accounted for only 0.6% of the variance in the responses (Table I), and the mean response distance varied less than 30% across the six talkers in the experiment. There was also little difference between mean perceived distances of male and female speech (2.83 and 2.52 m, respectively). The interaction between talker and simulated distance had a more substantial impact on the responses in the experiment than the talker alone ($\omega^2 = 0.013$). This interaction was an artifact of the differences in voicing across the talkers used in the experiment. Some talkers produced whispered speech at much higher production levels than other talkers, resulting in artificially low distance judgments at simulated distances less than 8.0 m. A more detailed explanation of this interaction is provided in the Appendix. Overall, voicing and production level appear to be far more important to the distance perception of speech than variations in the voices of different talkers.

b. Presentation level. Presentation level had a somewhat unexpected effect on the distance judgments (Fig. 5, second panel). The most intense stimulus presentations (60 dB) were consistently perceived farther away than the least intense stimulus presentations when the source distance was greater than or equal to 4.0 m. This contrasts with the results of almost all previous studies examining distance perception, which have shown that more intense stimuli are perceived closer than less intense stimuli. The results of experiment 2 will show that this increase in perceived distance with increasing presentation level occurs because listeners are influenced more by changes in the production level of speech than by changes in the presentation level of the speech.

c. Direction and utterances. Although both side and utterance were found to have significant effects on the overall perceived distances of the speech samples (Table I), neither of these effects was large enough to be considered meaningful ($\omega^2 = 0.002$), and there were no discernible differences in the response patterns with respect to the simulated distance of the source across the two conditions (Fig. 5, third and fourth panels). It appears that the effects of direction and utterance on the distance perception of speech are negligible.

d. Intersubject variability. There were some substantial differences across the responses of the eight subjects used in the experiment (Fig. 5, bottom panel). The overall mean response distances varied by a factor of about 3 across the listeners, from 1.5 m for the listener with the smallest mean distance judgments to 4.6 m for the listener with the largest. However, these intersubject variations were quite small relative to the overall range of response distances measured in the experiment, which increased by a factor of nearly 30 as the simulated distance was increased from 0.25 to 64 m. Furthermore, the overall patterns of the responses as a function of the simulated distance of the source were comparable across the eight listeners.

IV. EXPERIMENT 2: EFFECTS OF PRODUCTION AND PRESENTATION LEVELS ON THE APPARENT DISTANCE OF SPEECH

The results of the first experiment showed that the apparent distance of speech was dominated by the simulated distance of the talker. However, because the variations in production and presentation levels were constrained to those values associated with a live talker at one of the simulated distance values, it is impossible to use these results to determine the relative contributions that production and presentation levels had on the distance judgments in the experiment. In order to address this deficiency, a second experiment was conducted to determine how these two factors are combined into a single estimate of the apparent distance of a speech signal.

A. Experiment design

Six of the eight paid volunteer participants from the first experiment were used as listeners in the second experiment (three male, three female). The stimuli were nearly identical to those in the first experiment, with two exceptions. All of the stimuli in the second experiment were presented directly in front of the listener, and the variations in simulated distance in the first experiment were replaced with independent variations in the production and presentation levels of the stimuli in the second experiment. A total of 1386 speech samples, each representing a single trial, were prerecorded on a DAT tape for each listener. The speech samples were chosen randomly in a full-factorial design with the following four independent variables: six talkers (three male, three female), three utterances (''warning,'' ''threat,'' and ''over here''), 11 production levels (36, 42, 48, 54, 60, 66, 72, 78, 84, 90, and 96 dB SPL), and seven presentation levels (48, 54, 60, 66, 72, 76, and 82 dB SPL).

For each stimulus presentation, the experimental software selected the speech sample of the desired talker speaking the desired utterance that most closely matched the desired production level for that trial. The speech sample was then processed with the individualized HRTF of the listener and scaled to match the desired presentation level for that trial before being recorded on the DAT player. Note that, as in the first experiment, there were some minor differences between the desired production level and the actual production level in each trial. At each of the 11 production levels tested, the average production level of the speech samples was within 1 dB of the desired level. The overall standard deviation of the differences between the desired production levels and the actual production levels was 1.7 dB.

The samples were stored on the DAT tape with a pause of approximately 3 s between samples, and divided into blocks of 100 trials with a short tone marking the end of each block on the tape. As in the first experiment, the listeners judged the distances of the stimuli while seated outdoors in the middle of a large open field. The total experiment (1386 trials) was approximately 1 h and 45 min long for each listener, and was divided into two sessions on different days.

B. Results

1. Production Level

The effects of production level on response distance at each of the seven presentation levels used in the experiment are shown in Fig. 6. The curves in the top panel of the figure indicate that the distance judgments increased by as much as a factor of 32 when the production level of the voiced speech increased from 42 to 96 dB. This increase was not uniform

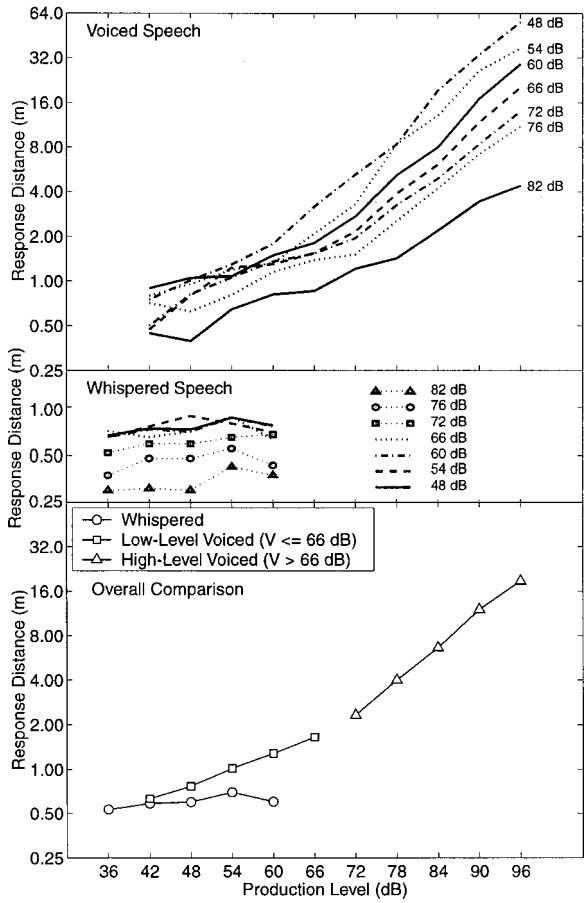


FIG. 6. Effects of production level on the perceived distance of speech in experiment 2. The top panel shows the mean distance judgments of voiced speech as a function of production level for each of the seven presentation levels used in the experiment (shown to the right of each curve). The middle panel shows the mean distance judgments of whispered speech as a function of production level at each presentation level (shown in the legend). The bottom panel shows the average of the distance judgments across all the presentation levels for three categories of speech: whispered speech, low-level voiced speech produced at or below 66 dB SPL at a distance of 1 m, and high-level voiced speech produced above 66 dB SPL at 1 m.

across all the production levels tested. The average results shown in the bottom panel of the figure indicate that the log distance judgments change nearly twice as rapidly with changes in production level for high-level voiced speech than for low-level voiced speech. The mean of the logarithmic response distance for voiced speech can be fit almost perfectly by a piecewise linear function of the production level (in dB) with different slopes for two distinct production level regions: response distance doubled with each 15-dB increase in production level for low-level voiced speech (produced at levels at or below 66 dB SPL), and doubled with each 8-dB increase in production level for high-level voiced speech (produced at levels above 66 dB SPL). Note that, in both cases, the increase in production level required to double the perceived distance of the speech was substantially larger than the 6-dB increase associated with a factor of 2 increase in the distance of a live talker in the free field. Thus, the results suggest that listeners are likely to underestimate changes in the distance of a live talker when the presentation level of the speech is fixed at the location of the listener.

As in the first experiment, the perceived distance of whispered speech did not seem to depend on the production level of the talker (Fig. 6, middle panel). The judged distances of the whispered speech samples were roughly independent of V across the presentation levels tested.

The results of a four-factor, repeated-measures ANOVA reveal that production level had a substantially larger impact on the listeners' responses than any other factor in the experiment (Table II). Production level ($\omega^2 = 0.628$) accounted for more than 60% of the overall variance in the listeners' responses, and more than seven times more of the variance than presentation level. Clearly the listeners' responses in this experiment were dominated by the apparent production level of the speech stimuli.

2. Presentation level

Production level dominated the listeners' distance judgments, but presentation level also played an important role in determining the apparent distances of the speech samples. The effects of presentation level on response distance at each of the 11 production levels used in the experiment are shown in Fig. 7. Again, there was a clear distinction between high-level voiced speech, low-level voiced speech, and whispered speech. The responses were influenced most by presentation level for high-level voiced speech ($V > 66$ dB), with a 50% decrease in response distance with every 12-dB increase in production level across the entire range of levels tested. At lower production levels, presentation level only had an effect in relatively intense stimulus presentations ($P \geqslant 72$ dB). The apparent distance of conversational and whispered speech was roughly independent of presentation level when P was less than 72 dB. In general, it appears that much larger

TABLE II. Analysis of variance from experiment 2. Note: This table shows only the significant factors and interactions ($p < 0.05$) from a four-factor, repeated-measures analysis of variance of the data from experiment 2.

Factor	DF	SS	DFE	SSE	F	P	ω^2
V (production level)	10	24 809	50	369	336.32	0.0001	0.628
P (presentation level)	6	3383	30	128	131.92	0.0001	0.085
T (talker)	5	148	25	45	16.44	0.0001	0.004
U (utterance)	2	99	10	40	12.36	0.0020	0.002
V×P	60	864	300	578	7.48	0.0001	0.019
V×T	50	612	250	281	10.90	0.0001	0.014
V×U	20	82	100	100	4.11	0.0001	0.002
V×T×U	100	203	500	433	2.35	0.0001	0.003

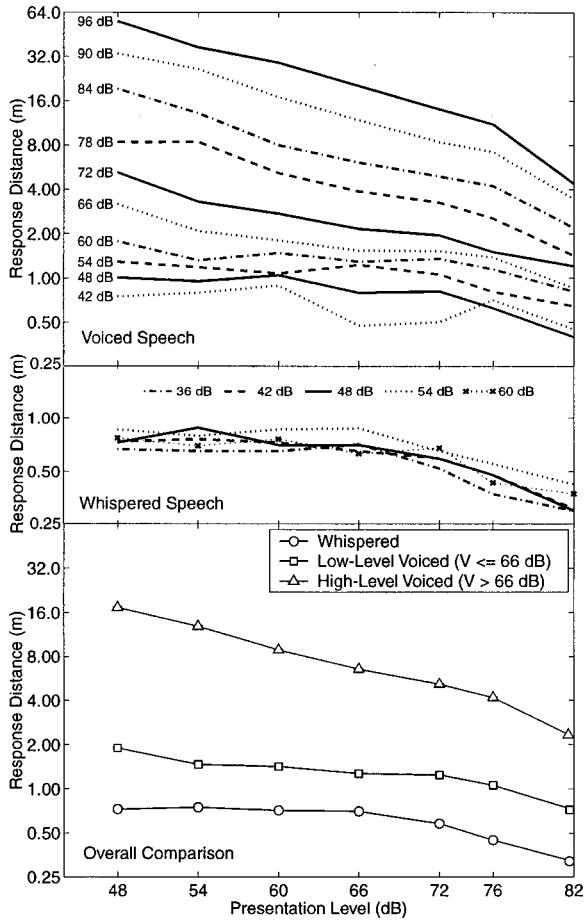


FIG. 7. Effects of presentation level on the perceived distance of speech in experiment 2. The top panel shows the mean distance judgments of voiced speech as a function of presentation level for each of the ten production levels of the voiced speech in the experiment (shown to the left of each curve). The middle panel shows the mean distance judgments of whispered speech as a function of presentation level at each of the five production levels used with whispered speech (shown in legend). The bottom panel shows the average of the distance judgments across the production levels for three categories of speech: whispered speech, voiced speech produced at or below 66 dB SPL at a distance of 1 m, and voiced speech produced above 66 dB SPL at 1 m.

changes in presentation level are required to produce a factor of 2 change in the perceived distance of speech signals than the 6-dB change in P that occurs with each doubling of talker distance in the free field.

Although presentation level had much less impact on the distance judgments than production level, with an ω^2 value of 0.085 it was still far more influential than any other main effect or interaction in the experiment (Table II). The ANOVA results also show that the interaction between presentation level and production level was the third most influential factor in this experiment ($\omega^2 = 0.019$). In part, this interaction can be explained by the ceiling effect that occurred at low presentation levels with whispered and low-level voiced speech. Another factor that contributed to this interaction can be seen in the top panel of Fig. 6: the responses were generally influenced less by production level at high presentation levels (especially at $P = 82$ dB).

3. Other factors

As in the first experiment, there was little difference in the distance judgments across the talkers and utterances used in the experiment. Mean responses varied only 30% across the six talkers used in the study, and only by 20% across the three utterances. Neither talker nor utterance could account for more than 0.5% of the total variance in the responses (Table II). The responses were also quite similar across the different listeners used in the experiment. The effects of production level and presentation level on the perception of speech seem to be robust across a wide range of talkers, utterances, and listeners.

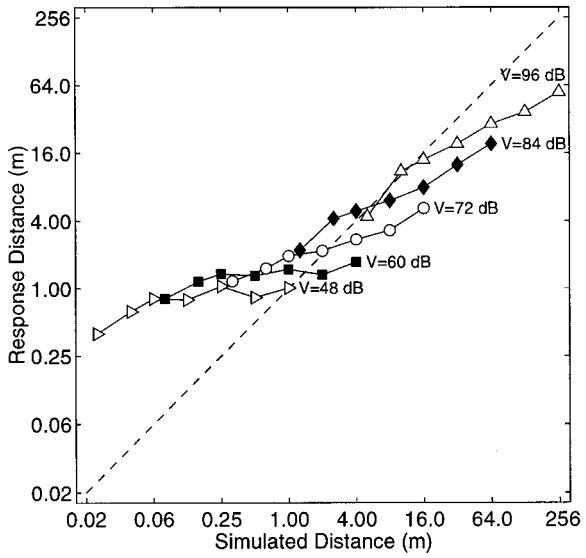


FIG. 8. Response distance for voiced speech as a function of simulated talker distance at five different production levels from experiment 2. Each curve plots response distance as a function of the simulated distance of the talker (the distance of the talker that would produce the same combination of production level and presentation level at the location of the listener in a free field) for all the trials in experiment 2 conducted at the production level shown to the right of the curve. Note that the seven symbols in each curve represent trials at different presentation levels. The dashed line represents ''correct'' responses where the distance judgments would match the simulated distance of the stimulus.

C. Discussion

1. The influence of production and presentation levels on the perception of talker distance in the free field

The results of experiment 2 provide some insights into the types of errors listeners are likely to make when judging the distances of live talkers in free-field environments. Figures 6 and 7 show how the apparent distance of a speech signal varies with the production and presentation levels of the utterance, but they do not provide much information about the types of distance errors the listeners are likely to make when judging the distances of live talkers. Figure 8 replots the response data for voiced speech as a function of the simulated talker distance in each trial. The five curves in the figure represent five different production levels (as indicated by the numbers to the right of each curve), and the seven symbols within each curve represent seven different presentation levels. The abscissa of each point represents the simulated distance of the talker for that particular combination of P and V. Recall from experiment 1 that simulated talker distance is the distance of a talker that would produce

the same combination of production and presentation levels at the location of the listener in a free field. Thus, to the extent that the virtual stimuli accurately replicate the speech signals that the listener would hear from a live free-field talker, the curves shown in Fig. 8 reflect the distance judgments that listeners are likely to make in a free-field listening environment.

These curves indicate that the apparent distance of a speech signal in the free field is highly dependent on the production level of the talker's voice. Most experiments concerning distance perception have shown a pattern of responses similar to the one found in experiment 1, where the listeners tended to overestimate the distances of nearby sounds and underestimate the distances of far-away sounds (Zahorik, 1998). The results in Fig. 8 confirm that, when judging the distance of a voice at a fixed production level, listeners will tend to overestimate the distance when the talker is close and underestimate the distance when the talker is far. For example, listeners will overestimate the distance of a talker at $V = 72$ dB when the simulated distance of that talker is closer than 2.0 m, and underestimate the distance when the simulated distance of the talker is farther than 2.0 m. However, the results also make it clear that, at any given talker distance, the listener's responses increased systematically with the production level of the stimulus. Thus, at each simulated distance, a listener's tendency to underestimate, overestimate, or correctly estimate the distance of the talker depended directly on the production level of the speech. When the simulated distance was 2.0 m, for example, listeners tended to overestimate the distance of the talker when V was greater than 72 dB, underestimate the distance when V was less than 72 dB, and correctly estimate the distance of the talker when V was equal to 72 dB. Gardner (1969) reported a similar relationship between production level and the apparent distance of speech: his listeners systematically increased their judgments about the distance of a live talker at a fixed location when that talker's vocal effort level increased from conversational speech to shouted speech. Thus, it is apparent that a listener's ability to make accurate judgments about the distance of a talker in the free field depends on the production level of the speech. This makes it very difficult to make general statements about the accuracy of distance perception for live speech stimuli in the free field. Indeed, it is possible that variations in the production levels of the talkers can account for at least some of the discrepancies between the highly accurate distance judgments presented in the live-speech studies of Gardner (1969), Von Bekesy (1949), and Cochran *et al.* (1968) and the somewhat less accurate distance judgments that occurred in experiment 1.

At each simulated distance tested, unbiased distance judgments generally occurred only at a single production level value. The production level required for unbiased distance estimation increased systematically with the distance of the talker. For example, listeners made unbiased distance judgments for a talker at 1 m only when V was approximately 50 dB, for a talker at 2.0 m only when V was approximately 70 dB, and for a talker at 8.0 m only when V was approximately 90 dB. One might expect the production levels that produce unbiased distance judgments to correspond to the natural adjustments in voice level that talkers make when communicating with distant listeners (Warren, 1968). However, the results of this experiment suggest that talkers would have to increase their production level much faster than they normally do in response to increasing conversational distance in order to produce consistently unbiased distance judgments on the part of the listener. At a talker distance of 1.0 m, unbiased distance perception occurred only when the production level was 50 dB. As distance doubled from 1.0 to 2.0 m, a 21-dB increase in V was required to maintain unbiased distance perception. At distances beyond 2.0 m, V had to be increased by about 9.3 dB for each doubling in distance in order to maintain unbiased distance perception. These increases are substantially larger than those that naturally occur in conversational speech: Lienard and Di Benedetto (1999) found that talkers increase their output level 9 dB when the distance to the listener increases from 0.4 to 6 m, which is only about 2.3 dB for each doubling of distance; Traunmüller and Eriksson (2000) found that talkers increase their output level only 1.3 dB when the distance to the listener increases from 0.3 to 1.5 m, and only by 4.6 dB for each doubling in distance beyond 1.5 m. These results suggest that listeners would not be able to use changes in the production and presentation levels of speech to accurately judge the distances of live talkers in real-world conversations: they would probably overestimate the distances of nearby talkers and severely underestimate the distances of far-away talkers. This pattern of response biases was also found in experiment 1 and in the listening experiments conducted by Traunmüller and Eriksson (2000).

The results of experiments 1 and 2 also suggest that listeners are unable to accurately judge the distances of whispering talkers. The data in Figs. 4 and 6 show that the production level of the talker has essentially no influence on the perceived distance of whispered speech. This implies that listeners cannot determine the voice level of a whispering talker, and that they base their distance estimates for whispered speech solely on presentation level. Because listeners are unable to determine whether a change in the level of a whisper was due to a change in the production level of the talker or in the distance of the talker, they should not be capable of accurately judging the distance of a whispering talker in the free field.

2. The role of categorization in the distance perception of speech

In our initial investigations of distance perception with speech stimuli, we hypothesized that listeners would derive their distance judgments from a direct comparison between the presentation level of the speech signal and the apparent production level of the talker. This hypothesis implicitly assumes that production and presentation levels play roughly equal roles in determining the perceived distance of the talker. However, the results of experiment 2 demonstrate that these two variables are interpreted differently for high-level voiced speech, low-level voiced speech, and whispered speech. Thus, it is clear that the processes used to classify speech signals into one of these three categories are at least

as important as production and presentation levels in determining the apparent distance of a speech stimulus.

At this point, it is difficult to determine how or why speech signals are categorized in this way. One possible explanation is that the acoustic cues used to determine the production level of the speech waveform are most salient for high-level voiced speech, somewhat less salient for low-level voiced speech, and least salient for whispered speech. Fundamental frequency ($F0$) is one example of an acoustic feature that is known to increase systematically with V for voiced speech (Lienard and Benedetto, 1999) but is effectively nonexistent in whispered speech. If listeners use $F0$ to determine the production levels of the stimuli, it would explain why their distance judgments are essentially independent of V for unvoiced speech. Other acoustic features might explain why production level has a much greater influence on apparent distance for high-level voiced speech than it does for low-level voiced speech.

Differences in acoustic features may explain why production level is interpreted differently for different categories of speech, but they cannot explain why presentation level is interpreted differently. A 3-dB change in intensity should be just as detectable in whispered speech as it is in high-level voiced speech. Some other nonacoustic factor must be responsible for the much smaller influence that presentation level had on the perceived distance of whispered speech. One possibility that is consistent with the results of this experiment is that the listeners were influenced by their past experiences with the ''natural'' voice levels that talkers use when they are engaged in conversations at different distances. This kind of influence could explain the ceiling effects that occurred in the distance judgments for unvoiced speech. The curves in the center panel of Fig. 7 show that the distance judgments for whispered speech were limited to a distance of approximately 0.8 m even at the lowest presentation levels tested in the experiment, and that there was essentially no variation in perceived distance of whispered speech when its presentation level was varied from 48 to 66 dB. This suggests that the listeners inherently associated whispered speech with nearby talkers, and that they restricted their responses to distances less than 1 m even when the presentation level of the whispered speech was low enough to indicate a more distant source. If their past experiences indicated that talkers only whisper when they are close, the listeners were probably inclined to restrict their responses to relatively short distances for whispered stimuli.

A related effect based on past experience might also explain why production level had less influence on the listeners' responses with low-level voiced speech than it did with high-level voiced speech. A recent experiment by Traunmüller and Eriksson (2000) examined the natural tendencies of talkers to adjust their voices to an output level appropriate for the distances of their intended listeners, and found that the talkers did not systematically increase their production levels until they were addressing a listener who was farther than 1.5 m away. At distances less than 1.5 m, the talkers did not systematically vary their production levels to adjust for changes in the distances of the listeners. This result indicates that production level does not vary systematically with distance in conversations over short distances, perhaps because intonation and emotion have a greater impact on voice level than listener proximity in these conversations. If listeners have learned through experience that the relationship between production level and talker distance is less reliable in conversations with nearby talkers producing low-level voiced speech than it is for distant talkers producing high-level voiced speech, it is not surprising that their distance judgments are influenced less by changes in production level for low-level speech.

Of course, other factors may also play a role in this speech categorization process. Further investigation is needed to fully explore the acoustic and cognitive factors that cause dramatically different distance perception for whispered, low-level voiced, and high-level voiced speech.

V. EXPERIMENT 3: EFFECTS OF TIME REVERSAL ON THE DISTANCE PERCEPTION OF SPEECH

The results of the first two experiments provide substantial evidence that listeners use changes in the production and presentation levels to judge the distances of speech signals. These results do not, however, provide much information about the specific acoustic properties of the speech signal that allow listeners to make these distance judgments. It is not clear whether the signal has to be normal comprehensible speech or if a speechlike signal with similar spectral characteristics is sufficient. In order to further investigate this issue, a third and final experiment was conducted to determine how listeners would judge the distances of time-reversed speech, which is spectrally and temporally similar to normal speech but lacks the semantic information that is present in ordinary speech.

A. Experiment design

The same six listeners used in experiment 2 were also used in experiment 3, and the basic design of the third experiment was almost identical to that used in the second experiment. The major difference was that the speech samples in the third experiment were time-reversed prior to being processed with the HRTFs and recorded onto the DAT tape. In addition, a seventh ''talker'' condition was added in which the speech signal was replaced with a 400-ms burst of white noise. The range of production and presentation levels was also reduced.

A total of 630 speech samples were recorded on the DAT tape for each listener in the experiment. The trials represented a full-factorial design with four independent variables: seven ''talkers'' (three male, three female, and white noise); three utterances (''warning,'' ''threat,'' and ''over here'');[2] six levels of production level (36, 48, 60, 72, 84, and 96 dB); and five presentation levels (54, 60, 66, 72, and 76 dB SPL). The trials were recorded on the tape in blocks of 100 trials with a beep and a short pause between each block. The entire experiment took approximately 45 min for each listener and was completed in a single experimental session.

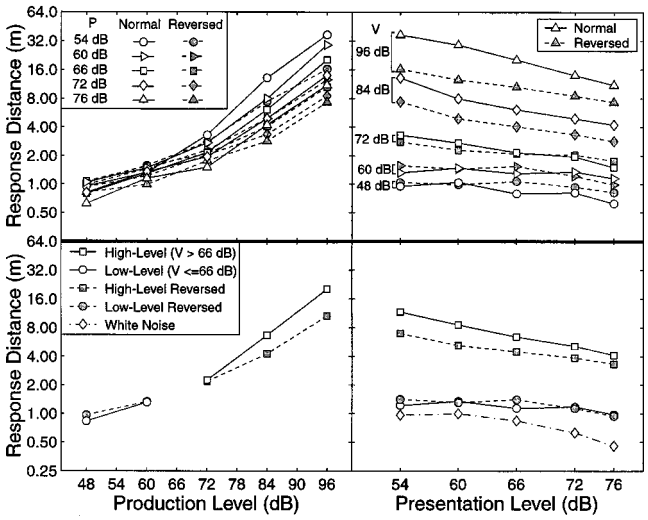


FIG. 9. Comparison of mean response distances as a function of production level (left panels) and presentation level (right panels) for normal voiced speech in experiment 2 (white symbols) and for time-reversed voiced speech in experiment 3 (shaded symbols). The data in the top panels show the data for each combination of production level and presentation level averaged across the different subjects, utterances, and talkers used in the experiment. In the upper left panel, the production level V is indicated by the abscissa of each point, and the presentation level P is indicated by the symbols shown in the legend. In the upper right panel, P is indicated by the abscissa of each point, and V is indicated by the numbers shown to the left of each curve. The mean values in the bottom panels have been averaged separately for the ranges of high-level and low-level voiced speech plotted in the bottom panels of Figs. 6 and 7. Note that the data shown for experiment 2 include only those data points that were collected at exactly the same production levels and presentation levels used in experiment 3. The bottom right panel also shows data from the white-noise stimulus tested in experiment 3 (diamonds).

B. Results and discussion

Figure 9 compares the average distance judgments for the reversed voiced speech stimuli from experiment 3 (shown by shaded symbols in the figure) to the average distance judgments for the nonreversed voiced speech from experiment 2 (shown by the open symbols). The left panels show the data in the same format used to show performance as a function of production level in Fig. 6, with the data for each combination of P and V in the top panel and the mean value across P for high- and low-level voiced speech in the bottom panel. The right panels show the data in the same format used to show performance as a function of presentation level in Fig. 7, with the data for each combination of P and V in the top panel and the mean value across V for high- and low-level voiced speech in the bottom panel. These results show that time reversal had little or no influence on the apparent distance of low-level voiced speech, but that it substantially reduced the apparent distance of high-level voiced speech. At the highest production level tested ($V=96$ dB), the distance judgments for the time-reversed speech were only about half as far as the distance judgments for the normal speech. The results of a three-factor, repeated-measures ANOVA (on the factors of playback direction, production level, and presentation level) indicate that the interaction between playback direction and production level shown in the bottom left panel of Fig. 9 [$F(4,20)=7.76$, $p=0.0006$] was the only interaction involving playback direction that was statistically significant at the $p<0.05$ level.

These results suggest that time-reversed speech contains all of the important acoustic cues that listeners use to determine the apparent distances of low-level conversational speech, but that it does not contain the acoustic cues that listeners use to determine the apparent distance of high-level shouted speech. The spectral content of time-reversed speech is identical to that of normal speech, so it does not appear that the distance perception of talkers at high production levels can be fully explained by the increases that occur in the fundamental and formant frequencies and the ratio of high- to low-frequency energy in shouted speech. One possible explanation for the different distance judgments for forward and reversed speech at high production levels is that listeners require both acoustic and phonetic information in order to accurately determine the production level of shouted speech. The acoustic effects of increasing production level are probably different for different phonetic elements, and listeners may need phonetic information in order to interpret the acoustic properties of shouted speech. In time-reversed speech, the phonetic information is scrambled and the listeners may be unable to properly interpret the acoustic information in the speech waveform. It may also be the case that the listeners simply failed to identify the reversed high-level speech as a shouted speech signal and that they interpreted the apparent production and presentation levels of the reversed shouted speech as if it were associated with a low-level voiced speech signal. This might explain why the apparent distances of the high-level voiced speech signals in the left panel of Fig. 9 appear to be an extension of the nearly linear relationship between log response distance and production level that exists for low-level forward or reversed voiced speech.

The nonspeech white-noise signal did not vary with production level, so it can be analyzed only in terms of the influence of presentation level on its apparent distance (diamonds in the lower right panel of Fig. 9). The apparent distance of the noise signal was roughly twice as sensitive to changes in overall level as the low-level speech signal, dropping by approximately 50% as the level was increased from 54 to 76 dB (versus a 22% drop for the low-level speech in experiment 2), but was influenced substantially less by changes in P than the high-level voiced speech in experiment 2 (which fell 67% with the same increase in P). Note that the overall influence of presentation level on perceived distance was similar for the noise and for the conversational speech: in both cases, presentation level had a substantially larger influence on distance perception when P was greater than 66 dB.

VI. CONCLUSIONS

The results of these experiments provide insights into the strategies that listeners use to integrate the apparent presentation and production levels of a speech signal into a single estimate of the distance of the talker. From these results, it is clear that there are fundamental differences in distance perception for three distinct classes of speech:

Whispered speech: Neither production level nor presentation level has much influence on the apparent distance of whispered speech, which is nearly always perceived at a distance approximately 0.8 m from the listener.

Low-level voiced speech: The perceived distance of low-level voiced speech (produced at or below 66 dB SPL at a distance of 1 m) doubles with each 15-dB increase in the production level of the talker. However, presentation level has relatively little impact on the apparent distance of low-level voiced speech until the presentation level of the stimulus exceeds 72 dB SPL. The perceived distance of conversational speech may rely primarily on the spectral properties of the speech: time reversal has little or no impact on the distance perception of low-level voiced speech.

High-level voiced speech: The perceived distances of high-level voiced speech signals (produced at levels exceeding 66 dB SPL at 1 m) are influenced substantially more by changes in production level and presentation level than the perceived distances of other types of speech signals. The apparent distance of high-level voiced speech doubles with each 8-dB increase in production level and with each 12-dB decrease in presentation level. The perceived distance of shouted speech cannot be explained by its spectral properties: time reversal results in a substantial decrease in the apparent distance of high-level voiced speech.

These results have important practical applications in the design of advanced virtual audio displays. Two important characteristics of the distance perception of speech make voice-based distance cues an appealing way to add robust, intuitive distance information to virtually synthesized speech sounds. The first is that production level and presentation level have substantially larger influences on the apparent distance of a speech signal than the vocal characteristics of the talker, the particular utterances spoken, or the direction of the stimulus. The second is that listeners are able to judge the distances of speech signals without *a priori* training on the vocal characteristics of the talker. These results suggest that variations in the production and presentation levels of virtually presented speech signals could be used to manipulate the apparent locations of those speech signals over an exceptionally wide range of distances, from less than 0.5 m to nearly 64 m, and that these manipulations could generate relatively robust distance simulations across a variety of different talkers, listeners, and utterances. In order to fully exploit the capabilities of a voice-based distance display, it will be necessary to determine the acoustic properties of speech that influence perceived production level and find ways to modify the apparent distances of arbitrary speech signals. Until this capability is achieved, prerecorded utterances like the ones used to generate the stimuli in these experiments can be used to simulate different talker distances in virtual audio display systems.

ACKNOWLEDGMENTS

The authors would like to thank Syed A. Abbas for his assistance in conducting this research, Chuck Goodyear for his assistance in the statistical analysis, and Don Mershon for his helpful comments on revising the manuscript. This research was supported by AFOSR Grant Nos. F49620-98-1-0108 and 01HE02COR.

APPENDIX: VOICING DIFFERENCES ACROSS TALKERS

Although the overall pattern of responses was similar across the six talkers (Fig. A1, top panel), there was a significant interaction between talker and distance that accounted for more than 1% of the overall variance in the responses (Table I). The nature of this interaction is illustrated by the data from talker 3, highlighted in Fig. A1 by the large asterisks. The speech samples of talker 3 were generally perceived closer than those of the other talkers at distances of 4.0 m or less, and farther away than the other talkers at distances more than 4.0 m. This effect is easily explained by the ability of talker 3 to generate higher-level whispered speech and lower-level conversational speech than the other talkers. Because the distance simulation algorithm simply selected the speech sample closest to the desired production level without regard for the voicing of the speech, this resulted in a much larger percentage of whispered speech samples at 2.0 and 4.0 m for talker 3 than for the other talkers (Fig. A1, bottom panel). Consequently, the mean response distances at those locations for talker 3 were

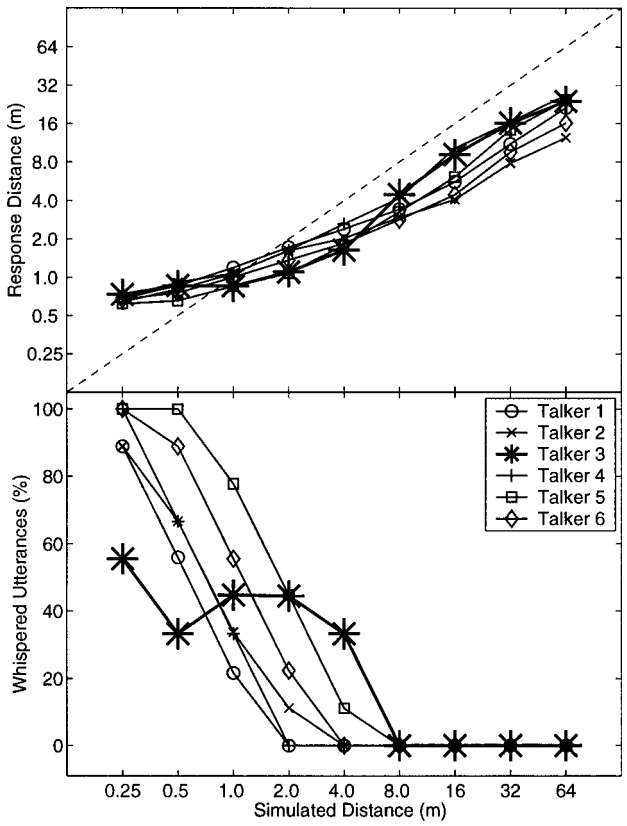


FIG. A1. Effects of talker and voicing on the distance perception of speech. The top panel shows the stimulus and response distances for the speech samples of each of the six talkers used in the experiment. The bottom panel shows the percentage of whispered speech samples at each distance for each of the six talkers. Note that the significant interactions between talker, level, utterance, and distance can largely be explained by talker 3's (*) tendency to produce more intense whispers than the other talkers, resulting in relatively short mean distance judgments for stimuli at 2.0–4.0 m.

substantially reduced by the tendency of listeners to associate all whispered speech with distances less than 1.0 m (Fig. 4). Thus, the interaction between talker and distance can be explained by the number of whispered utterances produced by each talker at each distance. This same effect can also account for all of the other significant interactions between talker, utterance, level, and distance (Table I), since all of these factors played a role in the selection of the speech samples.

[1]Note that both Zahorik and Begault (Begault, 1987) found that listeners placed more weight on intensity and less weight on the direct-to-reverberant energy ratio with speech than with other types of stimuli. This may occur because listeners have a better reference for the intensity of speech than for arbitrary noise and therefore are able to derive more absolute distance information from the intensity of speech.

[2]In the noise condition, three replications of the noise burst were substituted for the three different utterances.

Begault, D. R. (**1987**). ''Control of Auditory Distance,'' Ph.D. thesis, University of California, San Diego.

Cochran, P., Throop, J., and Simpson, W. (**1968**). ''Estimation of distance of a source of sound,'' Am. J. Psychol. **36**, 1171–1184.

Cohen, J. (**1977**). *Statistical Power Analysis for the Behavioral Sciences* (Academic, New York).

Coleman, P. (**1963**). ''An analysis of cues to auditory depth perception in free space,'' Psychol. Bull. **60**, 302–315.

Foster, S. (**1986**). ''Impulse response measurements using Golay codes,'' *IEEE Conference on Acoustics, Speech and Signal Processing*, Vol. 2, pp. 929–932.

Gardner, M. B. (**1969**). ''Distance estimation of 0 degree or apparent 0 degree-oriented speech signals in anechoic space,'' J. Acoust. Soc. Am. **45**, 47–53.

Grandstrom, B., and Nord, L. (**1992**). ''Neglected dimensions in speech synthesis,'' Speech Commun. **11**, 459–462.

Junqua, J. (**1993**). ''The Lombard reflex and its role on human listeners and automatic speech recognizers,'' J. Acoust. Soc. Am. **93**, 510–524.

Lienard, J.-S., and Benedetto, M.-G. D. (**1999**). ''Effect of vocal effort on spectral properties of vowels,'' J. Acoust. Soc. Am. **106**, 411–422.

Mershon, D., and Bowers, J. (**1979**). ''Absolute and relative cues for the auditory perception of egocentric distance,'' Perception **8**, 311–322.

Mershon, D., and Philbeck, J. (**1991**). ''Auditory perceived distance of familiar speech sounds,'' *Proceedings of the Psychonomic Society 32nd Annual Meeting*.

Moller, H., Sorensen, M., Hammershoi, D., and Jensen, C. B. (**1995**). ''Head-related transfer functions of human subjects,'' J. Audio Eng. Soc. **43**, 300–320.

Nielson, S. (**1991**). ''Distance Perception in Hearing,'' Technical report, Aalborg University, Denmark.

Pickett, J. (**1956**). ''Effect of vocal force on the intelligibility of speech sounds,'' J. Acoust. Soc. Am. **28**, 902–905.

Rostolland, D. (**1982**). ''Acoustic features of shouted voice,'' Acustica **50**, 118–125.

Shulman, R. (**1989**). ''Articulatory dynamics of loud and normal speech,'' J. Acoust. Soc. Am. **85**, 295–312.

Sluijter, A., and Heuven, V. V. (**1996**). ''Spectral balance as an acoustic correlate of the linguistic stress,'' J. Acoust. Soc. Am. **100**, 2471–2485.

Traunmüller, H., and Eriksson, A. (**2000**). ''Acoustic effects of variation in vocal effort by men, women, and children,'' J. Acoust. Soc. Am. **107**, 3438–3451.

Von Bekesy, G. (**1949**). ''The moon illusion and other similar auditory phenomena,'' Am. J. Psychol. **62**, 540–552.

Warren, R. (**1968**). ''Vocal compensation for change in distance,'' *in Proceedings of 6th International Congress of Acoustics*, Tokyo, pp. 61–64.

Zahorik, P. (**1998**). ''Experiments in Auditory Distance Perception,'' Ph.D. thesis, University of Wisconsin—Madison.

Dynamic articulatory model based on multidimensional invariant-feature task representation

Tokihiko Kaburagi
Department of Acoustic Design, Kyushu Institute of Design, 4-9-1, Shiobaru, Minami-ku, Fukuoka, 815-8540 Japan and CREST, Japan Science and Technology Corporation, Japan

Masaaki Honda
Human and Information Science Laboratory, NTT Communication Science Laboratories, 3-1, Morinosato-Wakamiya, Atsugi, Kanagawa, 243-0198 Japan and CREST, Japan Science and Technology Corporation, Japan

(Received 7 August 2000; accepted for publication 30 March 2001)

A dynamic model of articulatory movements is introduced. The research presented herein focuses on the method of representing the phonemic tasks, i.e., phoneme-specific articulatory targets. Phonemic tasks in our model are formally defined using invariant features of articulatory posture. The invariant features used in the model are characterized by the linear transformation of articulatory variables and found using a statistical analysis of measured articulatory movements, in which the articulatory features with minimum variability are taken to be the invariant features. Articulatory movements making vocal-tract constrictions or relative movements among articulators reflecting task-sharing structures are typical examples of the features found to have low variability. In the trajectory formation of articulatory movements, the dimension number of the phonemic task is set at a smaller value than that of articulatory variables. Consequently, the kinematic states of the articulators are partly constrained at given time instants by a sequence of phonemic tasks, and there remain unconstrained degrees of freedom of articulatory variables. Articulatory movements are determined so that they simultaneously satisfy given phonemic tasks and dynamic smoothness constraints. The dynamic smoothness constraints coupled with the underspecified phonemic targets allow our model to explain contextual articulatory variability using context-independent phonemic tasks. Finally, the capability of the model for predicting actual articulatory movements is quantitatively investigated using empirical articulatory data. [DOI: 10.1121/1.1373707]

PACS numbers: 43.70.Aj, 43.70.Bk, 43.70.Jt [AL]

I. INTRODUCTION

A. Dynamic modeling of articulatory behaviors

Contextual coarticulatory movements, such as carry-over and anticipatory effects, are characteristic features of continuous speech utterances, which form one of the central problems in the motor process of speech production (Sharf and Ohde, 1981). Coarticulation causes temporal spreading of phonemic features and can therefore be interpreted as sources of variability in phoneme articulation. Investigations aimed at revealing the influences of phoneme contexts on the attainment of phonemic articulatory targets have been carried out in physiological (MacNeilage and DeClerk, 1969), kinematic (Gay, 1977; Farnetani, 1989; Honda and Kaburagi, 1996), and acoustic (Stevens and House, 1963; Öhman, 1966; Blumstein, 1986) domains.

The representation of contextual coarticulatory effects is a major issue in the task-oriented trajectory formation of articulatory movements (Saltzman and Munhall, 1989; Bailly *et al.*, 1991; Kaburagi and Honda, 1996; Okadome *et al.*, 1998, 1999). Trajectory formation is a computational approach for generating spatial-temporal patterns of articulatory movements based on the specification of phonemic tasks and dynamic constraints. Phonemic tasks represent phoneme-specific articulatory goals and are used to constrain the kinematic states of the organs to achieve those goals. The dynamic constraints restrict the behavior of articulators to make a smoothly changing trajectory. They are frequently represented by a smoothness criterion, i.e., power consumption during movements defined locally (Saltzman and Munhall, 1989) or globally (Kaburagi and Honda, 1996) in the time domain. Movements of the articulatory organs are then determined so that they satisfy specified phonemic tasks and dynamic constraints simultaneously.

The mechanism for generating contextual articulatory variability is explained in the trajectory formation approach by means of the dimensionality of the task space and dynamic constraints regarding the smoothness of movements. The dimension number of the task space generally becomes smaller than that of the articulator variables because of the coordinated structure of the articulators (Saltzman, 1979) and the existence of priority among them (Coker, 1976). Because of the coordinated structure, a task can be shared by different articulators. This results in the so-called excess degrees-of-freedom problem; the task cannot determine uniquely the kinematic states of task-sharing articulators. In addition, articulators are constrained by the task only when they are essential to produce the phoneme, and articulators with low priority are free from it. Consequently, the kinematic states of the articulators are only partly constrained by

specified phonemic tasks and there remain unconstrained degrees of freedom of the articulator variables. These redundancies are resolved in trajectory formation by minimizing the smoothness criterion, and the values of unconstrained components are determined so that the tasks of adjacent phonemes are smoothly interpolated. Thus the redundant components are influenced by phonemic contexts and the model can automatically produce contextual coarticulatory movements using context-independent phonemic tasks.

Dynamic articulatory models based on the trajectory formation approach are thus capable of explaining contextual coarticulatory effects and the variability of nominal phonemic targets. Therefore, they provide the basis for behavior-oriented modeling of phonetic structure (Browman and Goldstein, 1985) and the recovery of articulatory gestures from speech input (McGowan, 1995). In addition, the models can be applied for dynamical control of articulatory models (Mermelstein, 1972; Engwall, 1999) or articulatory-based speech synthesis which mimics the human speech production process.

B. Types of phoneme-specific articulatory tasks

Phoneme-specific tasks in the trajectory formation model are represented as features of the articulatory gestures or speech acoustics that should be achieved by articulatory movements. They can formally be interpreted as static constraints from which the kinematic states of the articulatory organs are determined. In the generation of articulatory movements, tasks are organized according to the input sequence of symbols whereby the linguistic information of the intended utterance is expressed. Tasks are assigned for each of the input phonemes and their temporal structure (articulatory timing) is also specified.

Gesture-based phonemic tasks proposed thus far can be classified into context dependent and independent types. Context-dependent task models are referred to as syllable or allophone based models by Sharf and Ohde (1981). This model takes an allophone or syllable as the production unit and does not require a mechanism for producing contextual variability. Okadome *et al.* (1998, 1999) studied a representation method of the phonemic task following this context-dependent approach using diphone and triphone models. Their model is constructed as a stiff mapping from a string of input phoneme symbols to a temporal sequence of the context-sensitive target positions of every articulator. Trajectory formation based on context-dependent tasks provides a basis for accurate representation of articulatory movements in continuous speech production, but it inherently requires a lot of task patterns in order to cover every phonemic combination.

On the contrary, gesture-based context-independent phonemic tasks have been proposed in task-dynamic approaches (Saltzman and Munhall, 1989; Kaburagi and Honda, 1996). These task models, called the target based model by Sharf and Ohde (1981), use a phoneme-sized production unit internally specified by invariant articulatory targets. Motor control for each phoneme is programmed contingent upon the preceding and succeeding phonemes, and then contextual organization of articulatory movements is explained. Saltzman and Munhall (1989) proposed the blending of target vocal-tract gestures as such a mechanism. In the model studied by Kaburagi and Honda (1996), a smoothness criterion is used to interpolate tasks of adjacent phonemes. These trajectory formation models are then capable of automatically producing contextual articulatory variability from a small number of phonemic tasks. The potential task representation (Vatikiotis-Bateson *et al.*, 1992; Honda and Kaburagi, 1994) offers another type of context-independent model by introducing a metric criteria of the difference between articulatory variables and their target positions. This approach also has the advantage that it can explicitly explain reduced articulatory movements as well as contextual effects.

C. Invariant-feature task representation

Addressing the problem of context-dependent articulatory variability, this paper presents a novel method for representing phonemic tasks in the trajectory formation of articulatory movements and describes a dynamic model that can explain contextual coarticulatory movements.

The task in our approach is based on invariant articulatory features. These invariant features are calculated by finding linear transformations of measured articulator positions that minimize the ratio of within-class variance to total variance. According to this formal definition, the invariant feature can represent several types of consistent articulatory features: the posture of the articulator making a vocal-tract constriction, relative movements among articulators reflecting task-sharing or coordinated structures, and the relationship between different parts of an articulator, such as the tip and body of the tongue. Finally, the phonemic task is defined using invariant features to constrain the values of articulatory variables projected on the feature subspaces.

This task representation has several advantages when compared with the vocal-tract tasks used thus far. First, articulatory movements making vocal-tract constrictions are flexibly represented by the invariant features as a result of small articulatory variance at the constrictions. In the vocal-tract task, on the other hand, the place and degree of the vocal-tract constriction should be explicitly specified as phonemic targets. Second, by combining articulatory variables as features, relative movements among articulators in the spatial domain are automatically represented by the invariant features. This contrasts with the vocal-tract task, which requires an explicit model representing the coordinated structure among the articulators. Third, the dimension of the task space can be changed continuously to flexibly control constrained and unconstrained degrees of freedom of articulatory variables. The vocal-tract task has less freedom of task-dimensionality control because it is represented by a combination of a limited number of tract variables.

Our task model aims at deriving a compact set of context-independent phonemic tasks, and providing a means of constructing a dynamic articulatory model that can accurately predict actual articulatory movements in the production of continuous speech utterances. We therefore compare several types of task representation methods, i.e., context-dependent, vocal-tract, and invariant-feature tasks, and evaluate them with respect to the number of task patterns and

prediction accuracy of articulatory movements. In addition, an electromagnetic articulograph (EMA) system is used to obtain a set of articulatory data. This data set is used for the analysis of phoneme invariant features and a quantitative evaluation of the model's prediction accuracy.

This paper is organized as follows. Section II mathematically explicates our method of representing the phonemic invariant features and articulatory tasks. Section III describes our trajectory formation model of articulatory movements based on invariant-feature task representation. Section IV presents the results of an invariant feature analysis using an articulatory data set and simulation results of articulator movements. In addition, the accuracy of the movement model in predicting actual articulatory movements is quantitatively evaluated and compared with respect to the type of task representation method. Section V summarizes this work and gives our conclusions.

II. INVARIANT-FEATURE TASK REPRESENTATION

Trajectory formation of articulatory movements, which will be discussed in the next section, is a constructive approach to explain context-dependent articulatory movements based on static target patterns in the phoneme articulation. These phoneme-specific articulatory targets are referred to as phonemic tasks. The articulatory movements are represented in the trajectory formation model as a time sequence of articulatory variables representing kinematic states of the articulators. The spatial-temporal pattern of these articulatory variables is determined so that the articulatory targets are achieved by articulatory movements. In other words, phonemic tasks are used to constrain unknown values of the articulatory variables at given points in time.

In this section, the representation of the phonemic tasks is mathematically described. The task in our model is defined based on invariant features of the articulatory gestures in producing each phoneme. An articulatory posture that constricts the vocal-tract is a typical example of such a gesture. The invariant features are determined using covariance matrices of the articulatory variables by minimizing a normalized criterion. Based on this statistical approach, the phoneme-specific consistent posture of the articulators is automatically extracted as invariant features and these features are used to specify phonemic tasks. The articulatory variables in our model are described in Sec. III and examples of analyzed invariant features will be presented in Sec. IV.

A. Statistical determination of the invariant features

The invariant feature of a phoneme p is calculated by finding a linear transformation $\mathbf{f}_p$ of articulatory variables $\mathbf{x}$ that minimizes the following normalized articulatory variance (Honda and Kaburagi, 1996):

$$J(\mathbf{f}_p)=\frac{\mathbf{f}_p^t\Sigma_p\mathbf{f}_p}{\mathbf{f}_p^t\Sigma_T\mathbf{f}_p}, \tag{1}$$

where Σ_p and Σ_T, respectively, represent within-class and total covariance matrices of the articulatory variables, and t denotes the vector transposition. The optimal transformation is determined by solving the following generalized eigenvalue problem (Strang, 1976):

$$\Sigma_p F_p=\Sigma_T F_p\Lambda_p, \tag{2}$$

where $F_p=(\mathbf{f}_{p1},\mathbf{f}_{p2},\ldots,\mathbf{f}_{pL})$ and $\Lambda_p=\mathrm{diag}(\lambda_{p1},\lambda_{p2},...,\lambda_{pL})$ represent the eigenvector and eigenvalue matrices ($\lambda_{p1}<\lambda_{p2}<\ldots<\lambda_{pL}$), respectively. L is the dimension number of the articulator variables.

The criterion defined in Eq. (1) differs from one ordinarily used in the principal component analysis (PCA) in that it is normalized by the denominator associated with the total variance. In addition, the criterion is minimized in our method to determine the optimal transformation, while it is maximized in the PCA. By assigning every phoneme to one of the classes, phoneme-specific eigenvector and eigenvalue are derived from within-class and total covariance matrices.

Equation (1) indicates that the normalized articulatory variance is minimum in the subspace spanned by the optimal transformation. In other words, the transformation represents a characteristic feature of the articulatory gesture in the sense of minimum articulatory variance. For example, it can represent an articulatory gesture making a vocal-tract constriction by transforming the horizontal and vertical variables of an articulator into an axis that is perpendicular to the shape of the hard palate. The articulatory feature of bilabial consonants could also be represented by transforming the variables of the lips into their closing positions. In addition, relative movements among articulators can be represented by transforming the variables of different articulators. It should be noted that the eigenvalue λ_{pl} in Eq. (2) is equivalent to the value of the normalized variance $J(\mathbf{f}_{pl})$ for the corresponding transformation $\mathbf{f}_{pl}$. Therefore, the eigenvalue can be interpreted as a degree of articulatory consistency for each transformation.

In the solution of the eigenvalue problem in Eq. (2), covariance matrices are first calculated using data related to the articulatory variables as

$$\Sigma_p=\frac{1}{N_p}\sum_{i=1}^{N_p}(\mathbf{x}_{pi}-\overline{\mathbf{x}}_p)^t(\mathbf{x}_{pi}-\overline{\mathbf{x}}_p),\quad \overline{\mathbf{x}}_p=\frac{1}{N_p}\sum_{i=1}^{N_p}\mathbf{x}_{pi}$$

and

$$\Sigma_T=\frac{1}{\Sigma N_p}\sum_{p=1}^{P}\sum_{i=1}^{N_p}(\mathbf{x}_{pi}-\overline{\mathbf{x}}_T)^t(\mathbf{x}_{pi}-\overline{\mathbf{x}}_T),$$

$$\overline{\mathbf{x}}_T=\frac{1}{\Sigma N_p}\sum_{p=1}^{P}\sum_{i=1}^{N_p}\mathbf{x}_{pi},$$

where $\overline{\mathbf{x}}_p$ is the mean vector of a specific phoneme, $\overline{\mathbf{x}}_T$ is the total mean, N_p is the number of data samples, and P represents the number of phonemes. Next, Σ_T is factorized as $\Sigma_T=Q\Delta^2Q^t$ and a matrix A is set as $A=Q\Delta^{-1}$. $A^t\Sigma_pA$ is then factorized as $A^t\Sigma_pA=R\Lambda R^t$ and finally the eigenvector matrix is obtained as $F_p=AR=Q\Delta^{-1}R$. Based on the ei-

genvector matrix F_p, the numerator and denominator of Eq. (1) can be simultaneously diagonalized as $F_p^t \Sigma_T F_p = I$ and $F_p^t \Sigma_p F_p = \Lambda$, respectively. Note that matrices Q and R are orthogonal and Δ is diagonal.

B. Definition of the invariant-feature phonemic task

The linear transformation $\mathbf{f}_{p1}$ in Eq. (2) with the smallest eigenvalue λ_{p1} represents an articulatory feature that is less variant across phoneme contexts and other utterance conditions and hence would be most characteristic in producing the phoneme. In general, a phonemic task in the trajectory formation model is specified by setting the projection of the articulatory variables into an invariant feature space to zero:

$$0 = \mathbf{f}_{pl}^t (\mathbf{x} - \overline{\mathbf{x}}_p), \tag{3}$$

where vectors $\mathbf{x}$ and $\overline{\mathbf{x}}_p$, respectively, represent the articulatory variables and mean of the phoneme.

It is appropriate to use the first component of features ($l=1$) to constrain the values of the articulatory variables in Eq. (3), because it represents a characteristic articulatory feature of the phoneme. Then, the dimension of the task becomes one and the remaining $L-1$ kinematic degrees of freedom of the articulatory variables are not constrained. They become redundant components in the trajectory formation problem each of which can take an arbitrary value.

It is also possible to use two, three, or even more components together with the first one as the phonemic task. This multidimensional task representation results in simultaneous linear equations in which each equation is expressed in the form of Eq. (3). The task dimension number L_p, which represents the number of features used and hence is equivalent to the order of the simultaneous linear equations, can be set according to the eigenvalue of each transformation to control the unconstrained kinematic degrees of freedom of the articulator variables. When the task dimension number L_p is set to a small value, there are many degrees of freedom that can be used to explain the contextual variability. To the contrary, when L_p is the same as that of the articulator variables L, the position of every articulator is completely constrained and fixed to the mean vector $\overline{\mathbf{x}}_p$.

C. Allophonic targets

Although phonemes tend to exhibit consistent articulatory behaviors when they are articulated by making a vocal-tract constriction between the two lips, or the hard palate and the tip of the tongue, it is rather difficult to find such consistencies for back vowels and velar consonants even at the vocal-tract constriction due to contextually variable movements of the tongue body. It is therefore difficult to determine the less-variable articulatory feature and a unique phonemic target for each of these phonemes. The use of allophonic targets is investigated as a way to overcome this problem and achieve an accurate representation of contextual articulatory movements. By allowing a small number of context-sensitive variations of articulatory targets, the articulatory targets are switched according to the phoneme context to specify phonemic tasks of the trajectory formation model.

Allophonic targets are automatically constructed for each phoneme by clustering the data of the articulatory variables in which a label indicating preceding and succeeding phoneme symbols is assigned to each data sample. The LBG algorithm (Linde *et al.*, 1980) is used for clustering of this triphonic articulatory data space to determine cluster centroids of a given number. Meanwhile, articulatory data are also classified based on the phoneme context and the mean data are calculated for each context as the triphonic mean vector. Next, to assign every triphonic context to one of the LBG-based clusters, the Euclidean distance is calculated for every combination of cluster centroids and triphonic mean vectors. As such, each triphonic context is assigned to an LBG cluster for which the distance becomes minimum. Finally, invariant features of each allophonic target are calculated for each LBG cluster using the articulatory data with every triphonic context assigned to that cluster. In the generation of articulatory movements, these allophonic targets are switched based on the match between the input phoneme context and the triphonic contexts assigned to each cluster.

III. DYNAMIC MODEL OF ARTICULATORY MOVEMENTS

This section presents the dynamic articulatory model that predicts movements of the articulatory organs from an input string of phoneme symbols expressing the intended utterance. The basic framework of the model is based on Kaburagi and Honda (1996), where features of the vocal-tract shape were used to represent phonemic articulatory targets. However, in the present model, phoneme-specific articulatory tasks are represented on the basis of multidimensional invariant features. This task representation has the advantage, as mentioned in Sec. II, that it can flexibly reflect features of consistent articulatory gestures. In addition, an explicit articulatory model for representing coordinated structure among articulators is no longer required. Invariant-feature tasks can also represent the interdependency of different parts of the same articulator, such as the tip and body of the tongue, that reflects the physiological structure of the articulator. As discussed later, our task has extensive freedom in controlling the dimensionality of the task space. This feature is very important in explaining context dependent articulatory movements. The vocal-tract task has less freedom of task-dimensionality control because it is represented by a combination of a limited number of tract variables.

Figure 1 outlines the dynamic articulatory model which determines the trajectory of articulatory movements based on specifications of the invariant-feature phonemic tasks. A phonemic task and its articulatory timing are first specified for each input phoneme. Phonemic invariant features are used to construct a task, as defined in Eq. (3), as a simultaneous linear equation with respect to unknown values of the articulatory variables of the given time instant. Task dimension number L_p, a phoneme-dependent model parameter, should also be specified. L_p is set at a smaller value than the dimension number of the articulatory variables L. Then each phonemic task is represented as a constraint region in the articulatory variable space. This means that each task con-

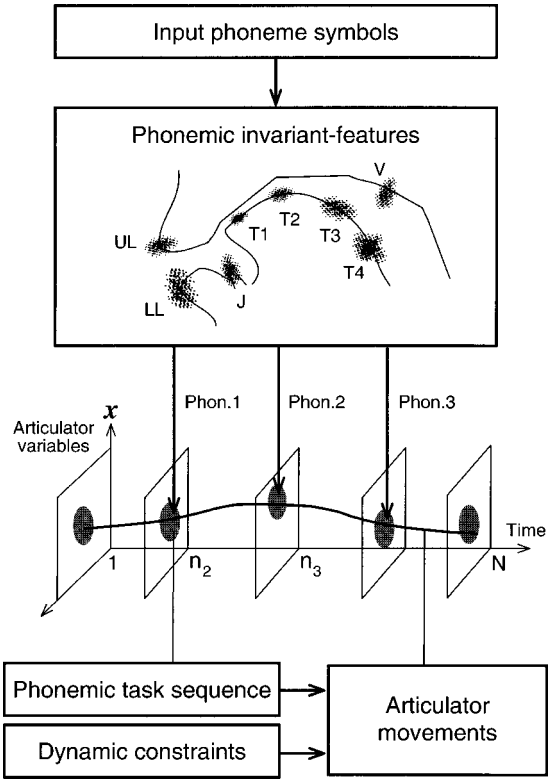


FIG. 1. Outline of the dynamic articulatory model based on invariant-feature task representation. Articulatory targets are represented by means of a sequence of phonemic tasks specified for input phoneme symbols. Each phonemic task is constructed by using phoneme-specific invariant features, i.e., a consistent articulatory posture in producing the phoneme, and it partially constrains the values of articulatory variables at given points in time. The trajectory of articulatory movements is determined so that it simultaneously satisfies given phonemic tasks and two types of dynamic constraint representing the smoothly moving behavior of the articulators. Redundant components of articulatory variables that are not constrained by the task are determined by smoothly interpolating the tasks of the adjacent phonemes. Context-dependent articulatory movements are then explained by the dynamic model on the basis of context-independent phonemic tasks represented using the phoneme invariant features.

strains the values of the articulatory variables only partly and consequently there remain unconstrained components of the articulatory variables. In addition, none of the articulatory variables are constrained during the time interval between the adjacent phonemic tasks. To resolve these spatial and temporal redundancies and to determine the articulatory trajectory uniquely, two types of dynamic constraints are introduced. Finally, the trajectory of articulatory movements is determined so that it satisfies specified phonemic tasks and the dynamic constraints simultaneously.

In the following, articulatory variables and each part of the dynamic articulatory model are explained more concretely. The method for determining articulatory movements is then discussed.

A. Articulatory variables

In our trajectory formation model, articulatory variables express the absolute position of points fixed on the surface of the articulatory organs as shown in Fig. 2. One of major reasons for selecting this type of articulatory variable is its direct link to the electromagnetic observation method of articulatory movements (Schönle *et al.*, 1987; Perkell *et al.*, 1992; Kaburagi and Honda, 1994, 1997). A set of fixed-point articulatory data can serve to calculate covariance matrices of the articulatory variables. Phonemic invariant features are then determined by solving an eigenvalue problem constructed from them. In addition, these articulatory variables are beneficial in comparing predicted articulatory movements, i.e., the temporal pattern of every articulatory variable, with measured movements. Then, the prediction accuracy of the trajectory formation model can be quantitatively evaluated.

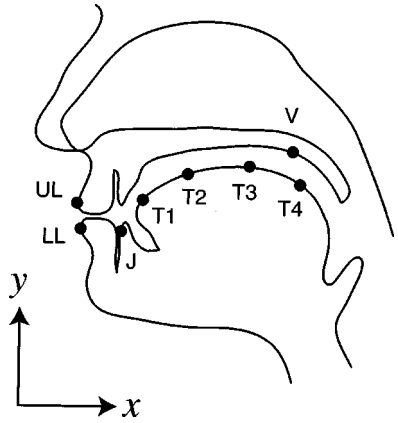


FIG. 2. Articulatory variables used in the dynamic articulatory model are represented as the horizontal and vertical positions of fixed points on the jaw (J), upper lip (UL), lower lip (LL), tongue (T1, T2, T3, and T4), and velum (V). The coordinate system of these articulatory variables is fixed to the head position; hence the absolute coordinate system represents the position of every articulatory organ as the articulatory variables.

B. Phonemic task sequence

In the trajectory formation model, the phonemic tasks are first sequenced according to the input string of phoneme symbols. Each task in the sequence is specified based on phoneme-specific invariant features $\mathbf{f}_{pl}$, task dimension number L_p, and the time instant at which the phoneme is articulated. The task representation in Eq. (3) constructs a simultaneous linear equation on the order of L_p with respect to the unknown values of the articulatory variables $\mathbf{x}$. Therefore, the sequence of phonemic tasks can be generally written as

$$\mathbf{z}_k = E_k \mathbf{x}(n_k) \quad (1 \leqslant k \leqslant K), \tag{4}$$

where $\mathbf{z}_k$ represents the target vector, E_k is the transformation matrix, n_k represents the articulatory timing, and k is an index of the task sequence. $n_1(=1)$ and $n_K(=N)$, respectively, correspond to the initial and terminal ends of articulatory movements. The articulatory timing n_k must be specified explicitly.

In specifying each phonemic task, context-independent invariant features are used unless the input phoneme is a back vowel or velar consonant. When a back vowel or velar consonant is input, context-sensitive allophonic tasks are used while taking the preceding and succeeding phoneme contexts into account. As described in the preceding section, the selection of allophonic tasks is based on the match between the input phoneme context and the triphonic context assigned to each allophonic task.

C. Dynamic constraints

Two types of dynamic constraints are introduced in the trajectory formation model. First, the trajectory of each articulatory variable is represented as the output of a linear second order dynamic system as

$$x_l(n)-2\tau_l x_l(n-1)+\tau_l^2 x_l(n-2) = (1-\tau_l)^2 y_l(n) \quad (1\leqslant l\leqslant L), \tag{5}$$

where y_l represents an input force to the system. Equation (5) represents a critically damped dynamical system and the time-invariant model parameter τ_l is specified depending on each articulator, because it determines the time constant of the dynamical system.

Second, an objective function representing the smoothness of articulatory movements is defined as (Okadome and Honda, 1992)

$$C=\sum_{n=1}^{N-1}(C_x(n)+C_y(n))+C_x(N), \tag{6}$$

where

$$C_x(n)=(\mathbf{x}(n)-\mathbf{x}(n-1))^t G W_x G^t(\mathbf{x}(n)-\mathbf{x}(n-1)) \tag{7}$$

represents the energy criterion with respect to the velocity of movements and

$$C_y(n)=(\mathbf{y}(n)-\mathbf{y}(n-1))^t G W_y G^t(\mathbf{y}(n)-\mathbf{y}(n-1)) \tag{8}$$

represents it with respect to the change of system inputs. The total energy C is calculated by summing the instantaneous power related to the change of system inputs and outputs over the entire movement.

Weighting matrices of the objective function are set as $W_x=\mathrm{diag}\{w_1,w_2,...,w_L\}$ and $W_y=\mathrm{diag}\{d_1w_1,d_2w_2,...,d_Lw_L\}$. Parameter w_l determines the cost weight between each articulatory variable and can be used to control the relative amplitude of the articulatory variables. Parameter d_l determines the relative cost between C_x and C_y and is used to adjust the overall shape of the trajectory together with parameter τ_l. G is an orthogonal matrix constructed from the eigenvectors of the total covariance matrix Σ_T. The total cost is weighted to each principal component of the articulator variables according to their variation by means of this orthogonal transformation. The influences of model parameters τ_l, w_l, and d_l on the determination of articulatory movements were previously described in Kaburagi and Honda (1996).

D. Determination of articulatory movements

The trajectory of articulatory movements is determined so that it satisfies the phonemic task sequence [Eq. (4)] and minimizes the objective function [Eq. (6)] under the constraint of the dynamic system representation [Eq. (5)]. From the linear system representation, it is clear that phonemic tasks constrain the values of outputs from the linear systems at given time instants. In addition, the objective function is related to the total energy consumption of system inputs and outputs. Therefore, the objective of the trajectory formation planning is to determine the optimal inputs to the systems so that these static and dynamic constraints are simultaneously satisfied. This framework produces an optimal control problem with linear dynamics and quadratic criteria. This problem can be solved explicitly by dynamic programming (Dreyfus and Law, 1977) when the values of the model parameters are fixed, and a specific algorithm for the problem is presented in Kaburagi and Honda (1996). Finally, the trajectory of the articulatory variables is calculated as the outputs of the linear systems driven by the optimal input forces.

The phonemic task representation in Eq. (3) indicates that only linear components of articulatory gestures are used as invariant articulatory features. In addition, the inter-articulator relationship is also restricted to a class which can be represented in a linear form. Although the representation of tasks and coordinated structures has such limitations in our model, these simplifications allow us to solve the optimal control problem explicitly. On the other hand, the relationship between the articulatory parameters and features of the vocal-tract shape generally becomes nonlinear in a geometrical articulatory model (Mermelstein, 1972). Therefore, the kinematic inverse problem for determining the values of articulatory parameters from specified vocal-tract tasks requires very complicated nonlinear optimization techniques. It is also difficult to derive an optimal solution for the nonlinear optimal control problem that satisfies given constraints and simultaneously minimizes the cost function.

IV. EXPERIMENT

This section presents the experimental results in analyzing phonemic invariant features and predicting articulatory movements. Phonemic invariant features were calculated using electromagnetically obtained articulatory data. Phonemic labels were first assigned to the data set for the time instant at which each phoneme was articulated. The phoneme-related positions of the articulatory organs were then used to construct the covariance matrices and the invariant features were determined by solving the generalized eigenvalue problem as described in Sec. II. Next, articulatory movements were simulated using the analyzed invariant features as the phonemic tasks. Movements for VCV utterances were first predicted to demonstrate the ability of our dynamic articulatory model to explain context-dependent variability. Then, simulated articulatory movements were compared with measured movements to evaluate the prediction accuracy of the proposed model.

A. Analysis of phonemic invariant features

The articulatory positions of the jaw, lips, tongue, and velum when each of two vowels and four consonants was produced are shown as scatter plots in the left part of Fig. 3. A set of articulatory data was obtained for a Japanese male subject using an electromagnetic articulograph system (Schönle *et al.*, 1987; Kaburagi and Honda, 1994, 1997) while the subject uttered 354 sentences. The articulatory timing of each phoneme was determined manually as the time instant at which the primary articulator exhibited its maximum displacement. The primary articulator of a consonant was selected as one that made the vocal-tract constricted. The articulatory timing of vowels was assigned, on the other

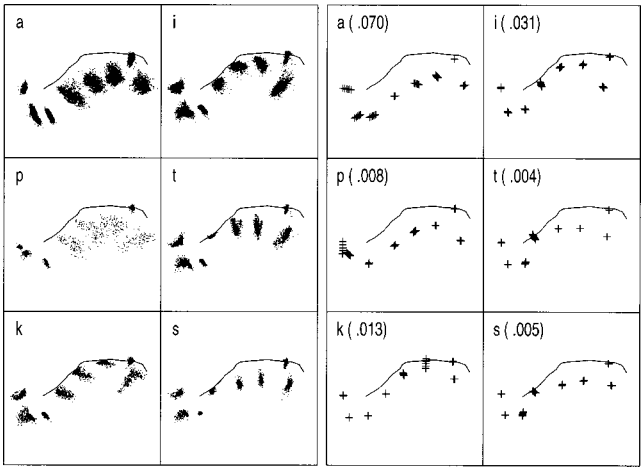


FIG. 3. Scatter plots of the articulator position in producing phonemes /a/, /i/, /p/, /t/, /k/, and /s/, and their first component of the phonemic invariant features. Scatter plots were obtained by measuring fixed points on the surface of the articulators using an electromagnetic device and collecting position data samples of time instants at which the primary articulator exhibited the maximum displacement. In the right part, plus signs indicate the direction of the linear transformation with the smallest eigenvalue obtained by solving the generalized eigenvalue problem [Eq. (2)], in which the covariance matrices were constructed using the articulatory data shown as the scatter plots. Phoneme-specific consistent articulatory postures, such as those constricting the vocal-tract, are emphasized by the transformation, i.e., the phonemic invariant feature. These plots were calculated by multiplying several gain factors to the linear transformation and adding it to the mean position for each phoneme.

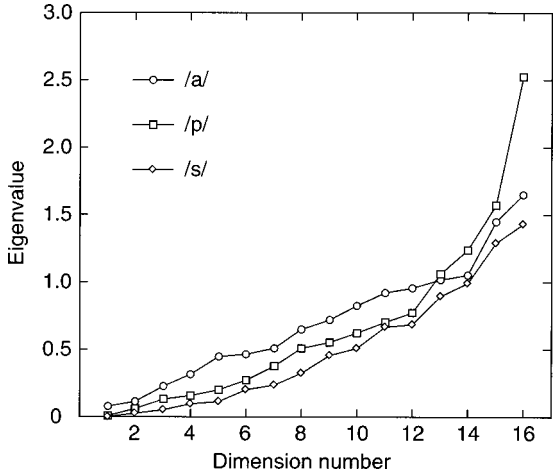


FIG. 4. Eigenvalue of the phonemic invariant-feature analysis as a function of the dimension number. The eigenvalue at each dimension is equivalent to the invariance measure [Eq. (1)] of the corresponding linear transformation. A dimension with a small eigenvalue therefore represents a consistent articulatory feature in producing the phoneme.

hand, based on specific movements of organs, such as jaw lowering, lip protrusion, or the raising or lowering of the tongue. The scatter plot for each phoneme was drawn by collecting data samples for which timing markers were assigned.

The right part of the figure shows the results of the phonemic invariant feature analysis. Plus signs indicate the direction of the linear transformation with the smallest eigenvalue, which were calculated by multiplying several gain factors to the eigenvector $\mathbf{f}_{p1}$ and adding it to the mean vector $\overline{\mathbf{x}}_p$ [see Eq. (1)]. It should be noted here that the markers staggered when the articulatory variation of the corresponding component was small, because the linear transformation was associated with a consistent articulatory feature. The numbers show the eigenvalues λ_{p1}, which are equivalent to the value of the normalized variance J for the corresponding transformation $\mathbf{f}_{p1}$. It therefore indicates the degree of articulatory consistency. It is clear from the figure that the transformations emphasize articulatory behaviors making vocal-tract constrictions, especially for the vowel /i/ and plosive consonants. The eigenvalue was smaller for consonants than for vowels, indicating that the position of the primary articulator was less variant in consonants. In addition, the eigenvalue tended to be larger when the phoneme was articulated using the body of the tongue. For instance, the eigenvalue was larger for the velar consonant than for the dental or labial consonant, and it was also larger for the back vowel than for the front vowel. These data can be interpreted as resulting from the relatively large context-dependent variability of the tongue body.

Figure 4 compares the eigenvalue of three phonemes as a function of the dimension number of the feature space. The eigenvalue increased as the dimension number increased because it was sorted in the ascending order [see Eq. (2)]. When the dimension number was smaller than 12, the eigenvalue was smaller for consonants than the vowel, reflecting the smaller variability of the primary articulators. On the other hand, the eigenvalues of the bilabial consonant were larger than the vowel eigenvalues when the dimension increased to more than 13. This reflects a context-dependent variability of the tongue, which has only low priority for producing the consonant.

B. Representation of contextual variability

To demonstrate the capability of the dynamic articulatory model to represent the contextual variability in continuous speech utterances, articulatory movements were simulated for several VCV sequences using the horizontal and vertical positions of eight fixed points as articulatory variables (Fig. 2). The task of each phoneme was specified based on invariant features, as shown in Fig. 3, obtained from the articulatory data set. The time interval between the initial V and the C was set equal to the interval between the C and the final V:50 ms. Simulation results are shown as the trajectory of each fixed point in Fig. 5. The shape of the tongue when the medial consonant is articulated is shown as polygonal lines by concatenating the positions of four fixed points along the tongue surface. Model parameters L_p, τ_l, w_l, and d_l were selected as their typical values at which simulation results agreed well with measured movements.

The results indicate that, although articulatory movements were simulated using context-independent phonemic tasks, tongue movements in articulating each consonant were influenced by the vowel context. It is also clear from the figure that, regardless of the vowel context, tongue movements for vocal-tract closure were maintained. In simulations containing the consonant /p/, the tongue was located in the front portion of the vocal tract for the /epi/ utterance and in the back for /epa/. This context-dependent difference of the

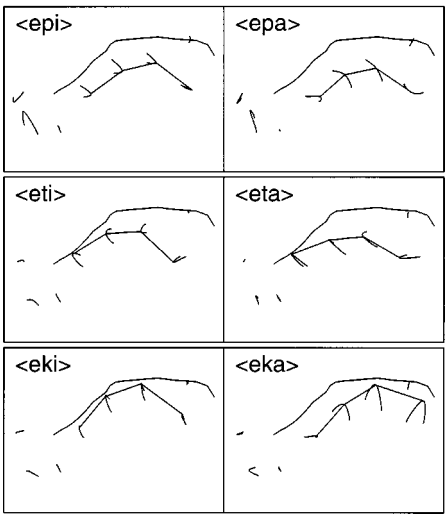


FIG. 5. Simulated articulatory movements for VCV sequences based on invariant-feature task representation. The trajectory of articulatory movements is shown for eight fixed points on the jaw, lips, tongue, and velum. The shape of the tongue when the intermediate consonant is articulated is indicated by connecting four points on the tongue. The shape of the hard palate is obtained by tracing an x-ray of the same subject.

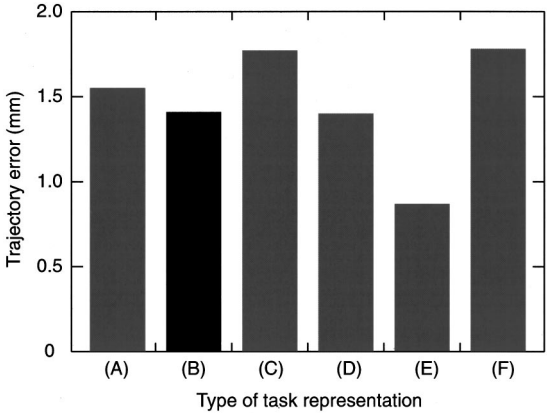


FIG. 6. Accuracy in predicting articulatory movements of continuous speech utterances using different types of the phonemic task model. Prediction accuracy is shown for (A) the invariant-feature task representation without the allophonic model, (B) the invariant-feature task representation with the allophonic model, (C) the uniphone task model, (D) the diphone task model, (E) the triphone task model, and (F) the vocal-tract task model. The prediction error was calculated as the mean error between simulated and actual movements for utterances of 25 sentences. In the uniphone, diphone, and triphone models, phonemic tasks constrained the value of every articulatory variable at their context-dependent positions (Okadome *et al.*, 1998, 1999). In the vocal-tract task model, tract variables representing the tongue position constricting the vocal-tract and the height and protrusion of the lips were used to specify the phonemic tasks (Kaburagi and Honda, 1996).

tongue posture can be interpreted as the result of the model-generated anticipatory effect reflecting the influence of the articulatory target of the succeeding vowels. For the consonant /t/ in /eti/ and /eta/ productions, it can be seen that the vocal-tract closure was attained because the behavior of the tongue tip was strictly specified in the phonemic task. However, the position of the tongue back was affected by the succeeding vowels. The tongue articulation was also affected by the vowel context for /k/.

These simulations rather qualitatively demonstrate the ability of the model to capture contextual variability of articulatory movements. It should be noted that articulatory behaviors making the vocal-tract constrictions are strongly constrained to produce these consonants and thus they were represented as phonemic invariant features. These features were used to specify phonemic tasks of the dynamic model and vocal-tract constrictions were therefore attained regardless of the phonemic context. On the other hand, variability of the tongue in the bilabial consonant and that of the tongue back in the dental consonant were represented by using redundant components of the articulatory variables that were not specified by the tasks. The values of these components were determined by smoothly interpolating the tasks of the preceding and succeeding vowels as a result of applying the dynamic constraints. The dynamic articulatory model could then exhibit contextual coarticulatory movements in VCV utterances using context-independent phonemic tasks.

C. Prediction accuracy

Finally, the prediction accuracy of the dynamic articulatory model was quantitatively investigated by comparing simulated movements with measured articulatory data for the same utterance. In addition, to determine the influence of the phonemic task representation on simulation accuracy, the invariant-feature task representation was compared with different task types such as vocal-tract and context-dependent tasks. In the vocal-tract task, features representing the shape of the vocal tract, such as the aperture and height of the lips, the position of the tongue making the vocal-tract constricted, and the position of the velum, were used to specify the tasks (Kaburagi and Honda, 1996). On the other hand, the positions of every articulatory variable in uniphone, diphone, and triphone contexts were used as the context-dependent phonemic task (Okadome *et al.*, 1998, 1999).

The target values of each task type were determined in this experiment using an articulatory data set of 100 sentences to which phonemic labels were assigned. Articulatory movements were simulated for 25 sentences in the same data set, and the error between simulated and measured movements, i.e., the rms (root mean square) error averaged over articulatory variables and time, was used to calculate the mean prediction error of each task type. The number of tasks in each utterance and their articulatory timing were identically used among the task types. Model parameters, i.e., the phoneme-dependent task dimension L_p, weighting parameters w_l and d_l of the smoothness criterion, and the parameter τ_l of the second order system, were determined using a search method so that the mean trajectory error was minimum for the same 25 sentences. It should be noted that the values of these model parameters were identical among test utterances.

Figure 6 compares the mean prediction error between model-generated and actual articulatory movements for six task types. As the task type, invariant-feature task representation was used in (A) and (B). In (B), allophonic tasks were used for the vowels /a/, /o/, and /u/ and velar consonants /k/ and /g/. The number of allophonic targets was set at four for each of these phonemes and the total number of the phonemic tasks was 35 in this study. Results from (C) to (E) cor-

TABLE I. Prediction error of articulatory movements simulated by using the invariant-feature task representation. The error was calculated as the difference of simulated and measured articulator positions at the time instant where a phonemic task was imposed.

Phoneme	Sample number	Mean error (mm)	Maximum error (mm)
a	122	1.596	3.223
i	94	1.175	2.274
u	48	1.245	2.098
e	61	1.455	2.532
o	86	1.624	3.197
pbm	46	1.619	3.053
tdn	112	1.267	2.009
kg	66	1.290	2.010
sz	66	0.817	1.566
r	26	1.315	1.906

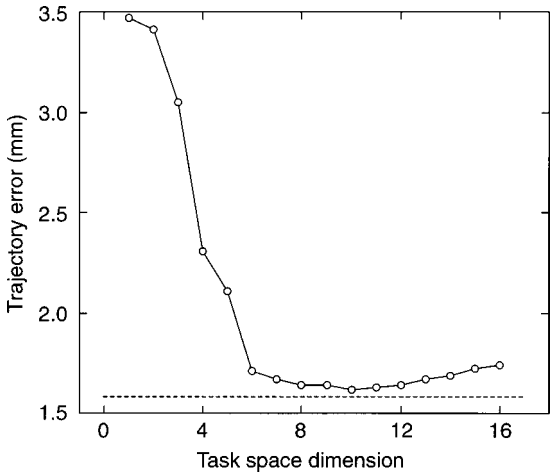


FIG. 7. Trajectory error as a function of the task dimension in invariant-feature task representation. The task dimension was set at a common value for every phoneme and the trajectory error was calculated as the mean difference between simulated and actual articulatory movements. The dotted horizontal line indicates the error when the task dimension was set independently for each phoneme at an optimal value.

respond to the context-dependent task, and (F) to the vocal-tract task.

Phonemic tasks in (A), (C), and (F) were context-independent and they had the same number of task patterns. By comparing the prediction errors among these three task types, it is shown that the invariant-feature task representation predicts articulatory movements better than the uniphone model in (C) and the vocal-tract task model in (F). This experimental result supports the usefulness of the proposed task representation method in explaining contextual articulatory movements. The invariant-feature tasks have flexible properties in representing consistent articulatory gestures to be specified as phoneme features and in controlling the task dimension which determines the constrained and unconstrained kinematic degrees of freedom of the articulatory variables. The vocal-tract task has less freedom with respect to these task representation properties because it is represented by a combination of a limited number of tract variables. In the uniphone model, tasks are represented as phoneme-specific mean vector $\overline{\mathbf{x}}_p$. Therefore, the value of every articulatory variable is constrained and there is no redundant component that can be used to represent contextual variability.

Experimental results for (B) and (D) indicate that the invariant-feature task representation achieves the same prediction accuracy as the diphone model when a small number of allophonic targets are used. The diphone model stores context-dependent articulatory targets, by taking the combination of successive two phonemes into account, and 248 task patterns were required to cover these phoneme contexts in the experiment. The value of every articulatory variable is constrained by the task, and the contextual variability of articulatory movements is fully explained by the contextual sensitivity of the task patterns. On the other hand, the number of task patterns was 35 in the invariant-feature task representation including the allophonic targets, indicating that our task model can achieve a fine prediction accuracy from a quite small set of articulatory targets.

Table I shows the phoneme-dependent simulation error in the invariant-feature task representation. The simulation errors were calculated at time instants for which the tasks were assigned, and their mean and maximum values were shown together with the number of samples appearing in 25 test utterances. The results indicated that the error was relatively large for back vowels /a/ and /o/ although the allophonic targets were used. On the other hand, the error tended to be small when the phoneme was articulated by the front portion of the tongue.

The effect of the task dimension L_p on the prediction accuracy was examined for the invariant-feature task representation and the results are shown in Fig. 7. When the task dimension was set identically among every phoneme and it was increased from 1 to 16, the error first decreased and then increased. The minimum prediction error was achieved at the dimension number of 10. When the dimension number is smaller than this optimal number, the increase in error occurs possibly due to the existence of too many redundant components of the articulatory variables that are not constrained by the tasks. On the other hand, when the dimension number is larger than ten, there are fewer redundant components usable to explain contextual effects resulting in an increase of the prediction error. The dotted horizontal line in the figure shows the prediction error when the dimension number was set at an optimal value for each phoneme. The optimal numbers of task dimensions for vowels /a/, /i/, and /u/ were 11, 5, and 13, respectively. Those for labial, alveolar, and velar consonants were 11, 9, and 13.

An example of simulated movements is shown in Fig. 8, where the horizontal axis represents time and the traces show the speech waveform and displacements of articulatory movements. Vertical lines above the speech waveform indicate articulatory timings determined from measured movements as time instants at which the primary articulator of each phoneme exhibited the maximum displacement. Thin and thick lines correspond to measured and simulated movements, respectively. The lowest trace shows the rms prediction error between measured and simulated movements as a function of time, in which the amplitude is magnified by a scaling factor of 8.

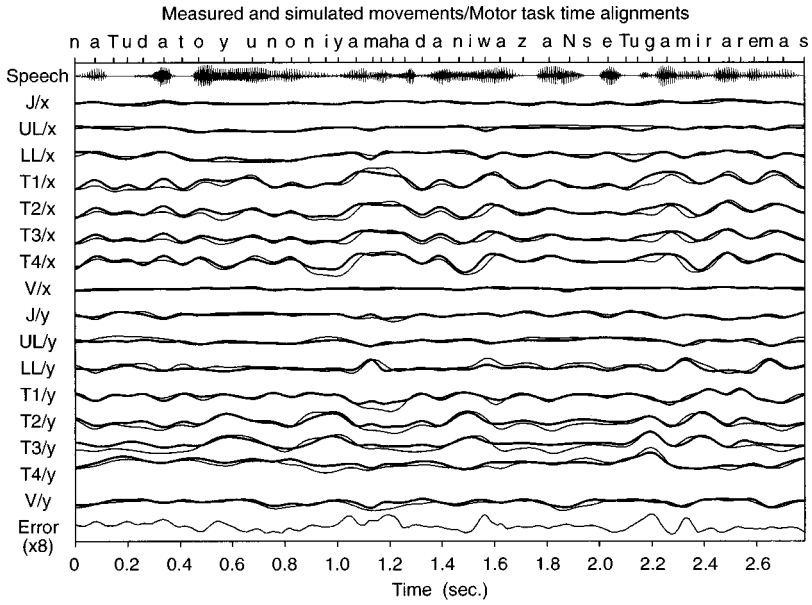


FIG. 8. Comparison of simulated and actual articulatory movements. Traces show the speech waveform, measured (thin lines) and simulated (thick lines) movements of the jaw (J), the upper lip (UL), the lower lip (LL), tongue points (T1, T2, T3, and T4), and velum (V) in the horizontal (*x*) and vertical (*y*) directions. The lowest trace shows the instantaneous difference between simulated and measured movements. Phonemic labels and corresponding time instants, at which phonemic tasks were temporally assigned, are shown above the speech waveform.

V. SUMMARY AND CONCLUSIONS

The representation of contextual effects is a common problem in speech information processing such as speech synthesis and recognition technologies. Instead of concatenating context-dependent spectrum patterns, trajectory formation offers a constructive way to automatically generate contextual variability from context-independent phonemic patterns.

This paper presented a novel approach for representing phonemic tasks of a dynamic articulatory model. The phonemic task in our model is represented using gesture-based invariant features in articulating each phoneme. The invariant feature is a linear transformation of articulatory variables determined statistically so that it minimizes the normalized variance of measured articulator positions. Given an input sequence of phonemes and their articulatory timings, the model produces articulator motions that reach phoneme targets at the time specified for each phoneme while also meeting two dynamic smoothness constraints. Contextual articulatory variability was explained in this dynamic articulatory model by means of the dimensionality of the task space and smoothness of articulator movements represented as the dynamic constraints.

In addition, use of allophonic targets was examined for back vowels and velar consonants to represent context-sensitive variability of the tongue movement more accurately. Invariant features of allophonic targets were calculated from articulatory data samples categorized by taking phoneme context of each data sample into account. Generated allophonic targets were switched in the trajectory formation based on contexts of the preceding and following phonemes in the input phoneme symbols.

The main experimental results and related considerations are as follows:

(1) Statistical analysis of articulatory movements based on the invariance criterion showed that consistent articulatory gestures could be extracted as invariant features. Articulatory gestures emphasized by the invariant features were interpreted graphically as the displacement of articulatory variables transformed by the eigenvectors. Furthermore, experiments showed that the value of the invariance measure, representing the articulatory variability, was especially small for labial and alveolar consonants, indicating that their articulatory configurations are less variable and consistent across phonemic contexts. On the other hand, the value was relatively large when the phoneme was articulated by the tongue in back vowels or velar consonants.

(2) Invariant feature phonemic tasks with respect to the articulatory variables were determined from electromagnetic fixed-point articulatory data, and articulatory movements in VCV contexts were simulated by the dynamic articulatory model. Experimental results showed that the model can predict context-dependent articulatory gesture of the intermediate consonants although each task was specified in a context-independent manner. Namely, while the primary articulator of each consonant achieved the vocal-tract constriction required to produce it, articulators with lower priority exhibited a variability due to the vowel context.

(3) Articulatory movements were simulated for utterances of short sentences and compared with actual movements obtained by an electromagnetic device. In addition, to examine influences of the task type on simulation accuracy, context-dependent and vocal-tract tasks were used in the simulation as well as the invariant feature task. As a result, it was revealed that the prediction accuracy of the proposed task model (invariant feature task) is better than that of the vocal-tract task. In addition, its accuracy is almost the same as that of the diphone task model when a small number of allophonic targets are used for back vowels and velar consonants. The number of task patterns is much smaller in the invariant feature task model than in the diphone model with the same prediction accuracy maintained. These differences among task types result from the property of the invariant feature task in which articulatory variability is automatically generated by the dynamic model. In the diphone model, such

variability is represented as the context-dependent target pattern itself.

(4) Simulation experiments indicated that the setting of the task dimension is quite important in controlling specified and the unspecified kinematic degrees of freedom of the articulatory variables to predict articulatory movements accurately. It was shown that the optimal task dimension number is 10 when it is identically set among phonemes (the number of dimensions of the fixed-point articulatory variables is 16). Experiments also showed that the prediction error slightly decreases when the number of task dimensions is set for each phoneme independently.

In conclusion, our simulation of articulatory movements showed that proposed task representation method has two major advantages when used in a dynamic articulatory model. One is that the phonemic invariant feature is defined statistically. Therefore, phonemic task patterns can be obtained automatically given enough articulatory data. In addition, although they are defined formally, the obtained phonemic patterns represent characteristic features of articulatory gestures in producing each phoneme. The other advantage is that the prediction accuracy of the dynamic articulatory model can be improved by using the invariant feature task representation. Experimental results showed that the prediction accuracy of the proposed task model is better than that of the vocal-tract task model and almost the same as that of the context-sensitive model (diphone task model).

It should be noted that the model can account for long term coarticulatory phenomena as long as the state of certain articulator is not constrained by the tasks, because the total energy consumption during an utterance is minimized by the dynamic criterion. In addition, invariant features are expected to represent variables that could be effectively controlled by the central nervous system, because the features can capture characteristic organization of articulatory behaviors. However, there are some limitations of the invariant features in interpreting the strategy of the speech motor control. First, the invariant features by their definition only capture linear relations of the articulatory behaviors measured in the Cartesian coordinates. Second, actively and passively controlled components of articulatory behaviors are hardly distinguished. For the interpretation of the behavioral features in relation to the motor control process, the magnetic resonance imaging of articulators would be useful in which the length of muscles and the posture of the articulators are simultaneously measured (Niitsu *et al.*, 1992; Takano and Honda, 2001).

Further investigations should address the construction of an articulatory-based speech synthesis model by combining the proposed articulatory model and a sample-based articulatory-to-acoustic mapping (Kaburagi and Honda, 1998). Spectrum parameters are controlled dynamically in this approach by articulatory movements and hence acoustic variability in continuous speech is explained in the articulatory dimension. This speech synthesis model is also expected to serve as a basic framework for acoustic evaluation of the dynamic articulatory model, articulatory-based speech inversion, and the audio-visual speech synthesis which can provide visible information of articulatory movements and audible speech signals simultaneously.

ACKNOWLEDGMENTS

This research was performed while the first author was in NTT Laboratories. The authors would like to thank Dr. John Hogden, Professor Pascal Perrier, and associate editor, Professor Anders Löfqvist, for their constructive comments on the paper.

Bailly, G., Laboissiere, R., and Schwartz, J. L. (**1991**). ''Formant trajectories as audible gestures: an alternative for speech synthesis,'' J. Phonetics **19**, 9–23.

Blumstein, S. E. (**1986**). ''On acoustic invariance in speech,'' in *Invariance and Variability in Speech Processes*, edited by J. S. Perkell and D. H. Klatt (Lawrence Erlbaum, NJ).

Browman, C. P., and Goldstein, L. (**1985**). ''Dynamic modeling of phonetic structure,'' in *Phonetic Linguistics*, edited by V. Fromkin (Academic, New York).

Coker, C. H. (**1976**). ''A model of articulatory dynamics and control,'' Proc. IEEE **64**, 452–460.

Dreyfus, S. E., and Law, A. M. (**1977**). *The Art and Theory of Dynamic Programming* (Academic, New York).

Engwall, O. (**1999**) ''Modeling of the vocal tract in three dimensions,'' Proc. Eurospeech 1999, pp. 113–116.

Farnetani, E. (**1989**). ''V-C-V lingual coarticulation and its spatiotemporal domain,'' in *Speech Production and Modeling*, edited by W. J. Hardcastle and A. Marchal (Kluwer, Dordrecht).

Gay, T. (**1977**). ''Articulatory movements in VCV sequences,'' J. Acoust. Soc. Am. **62**, 183–193.

Honda, M., and Kaburagi, T. (**1994**). ''A dynamical articulatory model using potential task representation,'' Proc. Int. Conf. Spoken Language Processing (ICSPL94), pp. 179–182.

Honda, M., and Kaburagi, T. (**1996**). ''Statistical analysis of a phonemic target in articulatory movements,'' J. Acoust. Soc. Am. **100**, 2598.

Kaburagi, T., and Honda, M. (**1994**). ''Determination of sagittal tongue shape from the positions of points on the tongue surface,'' J. Acoust. Soc. Am. **96**, 1356–1366.

Kaburagi, T., and Honda, M. (**1996**). ''A model of articulator trajectory formation based on the motor tasks of vocal-tract shapes,'' J. Acoust. Soc. Am. **99**, 3154–3170.

Kaburagi, T., and Honda, M. (**1997**). ''Calibration methods of voltage-to-distance function for an electro-magnetic articulometer (EMA) system,'' J. Acoust. Soc. Am. **101**, 2391–2394.

Kaburagi, T., and Honda, M. (**1998**).''Determination of the vocal tract spectrum from the articulatory movements based on the search of an articulatory-acoustic database,'' Proc. Int. Conf. Spoken Language Processing (ICSLP98), pp. 433–436.

Linde, Y., Buzo, A., and Gray, R. M. (**1980**). ''An algorithm for vector quantization,'' IEEE Trans. Commun. **COM-28**, 84–95.

MacNeilage, P. F., and DeClerk, J. L. (**1969**). ''On the motor control of coarticulation in CVC monosyllables,'' J. Acoust. Soc. Am. **45**, 1217–1233.

McGowan, R. S. (**1995**). ''Recovering task dynamics from formant frequency trajectories: Results using computer ''babbling'' to form an indexed database,'' in *Producing Speech: Contemporary Issues*, edited by F. Bell-Berti and L. J. Raphael (AIP Press, New York).

Mermelstein, P. (**1972**). ''Articulatory model for the study of speech production,'' J. Acoust. Soc. Am. **53**, 1070–1082.

Niitsu, M., Kumada, M., Niimi, S., and Itai, Y. (**1992**). ''Tongue movement during phonation: a rapid quantitative visualization using tagging snapshot MR imaging,'' Ann. Bull. Res. Inst. Log. Phoniat. (Univ. Tokyo) **26**, 149–155.

Öhman, S. E. G. (**1966**). ''Coarticulation in VCV utterances: Spectrographic measurements,'' J. Acoust. Soc. Am. **39**, 151–168.

Okadome, T., and Honda, M. (**1922**). ''Trajectory formation in sequential arm movements,'' Proc. IEEE Intern. Conf. SMC, pp. 471–478.

Okadome, T., Kaburagi, T., and Honda, M. (**1998**). ''Trajectory formation of articulatory movements for a given sequence of phonemes,'' Proc. Int. Conf. Spoken Language Processing (ICSLP98), pp. 3131–3134.

Okadome, T., Kaburagi, T., and Honda, M. (**1999**). ''Articulatory movement formation by kinematic triphone model,'' Proc. IEEE Intern. Conf. SMC, WQ16-6.

Perkell, J. S., Cohen, M. H., Svirsky, M. A., Matthies, M. L., Garabieta, I., and Jackson, M. T. T. (**1992**). ''Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements,'' J. Acoust. Soc. Am. **92**, 3078–3096.

Saltzman, E. L. (**1979**). ''Levels of sensorimotor representation,'' J. Math. Psychol. **20**, 91–163.

Saltzman, E., and Munhall, K. G. (**1989**). ''A dynamical approach to gestural pattering in speech production,'' Ecological Psychol. **1**, 333–382.

Schönle, P. W., Gräbe, K., Wenig, P., Höhne, J., Schrader, J., and Conrad, B. (**1987**). ''Electromagnetic articulography: Use of alternating magnetic fields for tracking movements of multiple points inside and outside the vocal tract,'' Brain Lang. **31**, 26–35.

Sharf, D. J., and Ohde, R. N. (**1981**). ''Physiological, acoustic, and perceptual aspects of coarticulation: Implications for the remediation of articulatory disorders,'' in *Speech and Language: Advances in Basic Research and Practice*, Vol. 5, edited by N. J. Lass (Academic, New York).

Stevens, K. N., and House, A. S. (**1963**). ''Perturbations of vowel articulations by consonantal context: An acoustical study,'' J. Speech Hear. Res. **6**, 111–128.

Strang, G. (**1976**). *Linear Algebra and its Applications* (Academic, New York).

Takano, S., and Honda, K. (**2001**). ''Muscle length measurement during vowel production based on magnetic resonance images'' (in Japanese), Technical report of IEICE, SP2000-150, pp. 59–66.

Vatikiotis-Bateson, E., Hirayama, M., Honda, K., and Kawato, M. (**1992**). ''The articulatory dynamics of running speech: Gestures from phonemes?,'' Proc. Int. Conf. Spoken Language Processing (ICSLP92), pp. 887–890.

Generation of articulatory movements by using a kinematic triphone model

Takeshi Okadome and Masaaki Honda
NTT Laboratories/CREST, JST, 3-1 Morinosato-Wakamiya, Atsugi-shi, Kanagawa-pref., 243-0198 Japan

(Received 4 April 2000; accepted for publication 11 April 2001)

The method described here predicts the trajectories of articulatory movements for continuous speech by using a kinematic triphone model and the minimum-acceleration model. The kinematic triphone model, which is constructed from articulatory data obtained from experiments using an electro-magnetic articulographic system, is characterized by three kinematic features of a triphone and by the intervals between two successive phonemes in the triphone. After a kinematic feature of a phoneme in a given sentence is extracted, the minimum-acceleration trajectory that coincides with the extremum of the time integral of the squared magnitude of the articulator acceleration is formulated. The calculation of the minimum acceleration requires only linear computation. The method predicts both the qualitative features and the quantitative details of experimentally observed articulation. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1377633]

PACS numbers: 43.70.Aj, 43.70.Bk, 43.70.Jt [AL]

I. INTRODUCTION

A. Motivation and background

Speech acoustic patterns of an uttered phoneme, in continuous speech, vary depending on the phonemic context. The contextual variability of an acoustic pattern originates in temporal overlapping of phoneme-specific articulatory behaviors, called coarticulation.

An articulatory parameter, the position of the articulator, changes more smoothly in time compared with an acoustic parameter. Furthermore, when uttering a phoneme, each articulator (such as the jaw, lips, tongue, velum, and Adam's apple) has its own phonemic-specific features.

We have developed a method for producing articulatory trajectories for continuous speech. It can be used by an articulatory-domain speech-synthesis system. Articulatory-domain speech synthesis requires high-fidelity generation of articulatory behavior and, thus, the error between the observed trajectory and that predicted by using the method must be small. Furthermore, an articulatory-domain speech-synthesis system must compute the trajectory in a short time.

The method uses phoneme-specific tasks based on a classical context-sensitive coding model. The context-sensitive coding model has been criticized on several grounds, one of which was computation of the number of extrinsic allophones that would need to be stored (for example, see Kent, 1983). None of the criticisms, however, has resulted in a quantitative evaluation of the model, because of a lack of articulatory data to evaluate the context-sensitive coding model. Our strategy minimizes the smoothness criterion to determine the articulator movements uniquely among an infinite number of the trajectories that are compatible with phoneme-specific tasks assigned in order of time. This strategy is along the same line of that of arm movement studies (for example, Flash and Hogan, 1985). We show that the context-sensitive coding model, together with the strategy of minimizing the smoothness criterion, is sufficiently powerful to determine the articulatory movements.

B. Related work

1. Strategies for coping with contextual variability

Trajectory formation models for articulatory movements, just like acoustic-domain speech-synthesis models for continuous speech, deal with problems associated with the contextual variability in articulatory behavior (for example, Vatikiotis-Bateson *et al.*, 1992). They are classified into two classes: proximal and distal control models.

The proximal control model uses a classical context-sensitive coding method (Wickelgren, 1969) and different phonemic tasks according to the phonemic context to generate the contextual variability. The proximal control model requires a large amount of articulatory data, because it uses a phoneme-specific task defined for each phonemic context.

On the other hand, the distal control model defines phoneme-specific tasks that are independent of phonemic contexts. The model uses the excess degrees of freedom in determining the trajectory of the coordinated articulatory movements. It thus requires only a small number of tasks defined by using formant frequencies (Bailly *et al.*, 1991; Stevens, 1989; Perkell *et al.*, 1993), vocal tract features (Saltzman and Munhall, 1989), and kinematic features of the articulator (Honda and Kaburagi, 1994, 1996; and Kaburagi and Honda, 1996). This is a merit of the distal control model.

Browman and Goldstein (1992) offered a distal control model that is couched in gestural terms, where gestures are abstractions rather than movement trajectories. Although, in their model, the context variability appears as an inevitable consequence of coproduction of articulatory gestures, the values of schwa that their model produces are much less variable than in observed speech. Saltzman and Munhall's (1989) task-dynamics model also could not sufficiently recover the context variability by only using excess degrees of freedom. Thus, they had to bring phonemical rules into the model as well.

Kaburagi *et al.*'s (1999) method did not use any phonemical rules. The error between an observed and predicted

trajectory by their method, however, was much larger than that using the proximal control model. In particular, the method using the distal control model failed to reproduce the characteristics of stop consonants such as /t/ or /d/ whose pronunciation requires quick movement of the tongue tip. This is because the smoothness criterion in their method was kept constant over the whole sentence.

Although the proximal control model requires a large amount of articulatory data, we use it because it is simple and can predict the trajectory of an articulatory movement with a small error.

2. Smoothness criteria: Trajectory generation of the articulation

Inspired by Jordan's work on motor control (for example, Jordan, 1989), Bailly *et al.* (1991) used a forward model of the articulatory plant that had a neural net structure. In the learning process of their model, the network units' output values are compared with the desired trajectories and the errors are propagated back in time, while modifying the parameters of the network according to their partial contribution to the total error. An additional cost, kinetic energy, was added to these backpropagated partial errors to cope with constraints specific to each unit. Although the model predicts the vocalic gesture well, it is not known whether or not it can predict articulatory movements for continuous speech.

Hirayama *et al.* (1992) generated input commands to the muscle-skeletal system of the articulatory organs by minimizing the criterion related to the change in input commands. They also constructed a neural-network model of the musculo-skeletal system (the articulator). The network learned the correlations between position, velocity, EMG at time t, and the velocity and acceleration of the articulator at the next time sample $t+1$. The network used the initial articulator position and velocity and the continuous EMG motor command input to generate trajectories. Furthermore, they used a cascade neural network to generate motor commands and articulator trajectories from phoneme-specific articulatory targets and speaking rate. This network was used to determine the appropriate setting of the smoothness constraint: minimum motor command change. Note first that they constructed only the model of the lips and jaw and, thus, that their trajectory-formation method predicted only the movements of the lips and jaw. Note second that the calculation of articulatory movements by the model required nonlinear optimization and, thus, it took much time to compute articulatory trajectories.

To solve the kinematic redundancies, Kaburagi and Honda (1996) used the cost function of Okadome and Honda (1992), which is the time integral of a weighted sum of the torque change and angular velocity around the joints. The trajectory formation model based on the cost function predicts general sequential movements more accurately than do other criteria relating to jerk, torque changes, or kinetic energy. If the dynamical system is nonlinear, however, the calculation of trajectories based on the cost function requires nonlinear optimization. Thus, Kaburagi and Honda had to assume that the articulatory system was represented by a linear system, which may not be physiologically plausible.

Okadome *et al.* (1998) developed a method that could be used to calculate the trajectories of articulatory movements. Their method used the minimum-jerk model (Flash and Hogan, 1985). The model predicts trajectories well; however, to produce trajectories, the minimum-jerk model requires, in addition to the position and velocity, the acceleration of the articulator as one of the phoneme-specific features. This may affect accuracy because the acceleration was numerically calculated from the observed data.

C. Outline and characteristics of the developed model

The model for forming trajectories of articulatory movements proposed here uses a classical context-sensitive coding method to specify phoneme-specific tasks. It determines trajectories uniquely by minimizing a cost function. As the classical context-sensitive coding method, we develop a kinematic triphone model. The cost function is the time integral of the square of the magnitude of *acceleration* of the articulator. The triphone model is characterized by intervals between two successive phonemes in a triphone and by three kinematic features associated with the triphone; each of the kinematic features is associated with each of the phonemes contained in the triphone. A kinematic feature associated with a phoneme (*the kinematic feature of a phoneme*) is represented by the position and velocity of an articulator. Furthermore, the calculation of the minimum-acceleration trajectory does not require any knowledge of the dynamical property of the articulator. The method produces trajectories simply by extracting kinematic features and by using linear computation. In addition, we can construct the kinematic triphone model in a short computational time by using observed data because we need only a straightforward computation.

The following sections describe the kinematic triphone model, the minimum-acceleration model that calculates the trajectory of the articulatory movement for a given sequence of phonemes, and an evaluation of that model. This article deals only with the production of speech at normal speed.

II. KINEMATIC TRIPHONE MODEL

The developed triphone model, which is called a *kinematic triphone model*, is characterized by three kinematic features associated respectively with three phonemes in a triphone and by the time intervals between two successive phonemes in the triphone; each of the kinematic features is defined for each of the phonemes contained in the triphone. A *kinematic feature* is represented by the position and velocity of an articulator.

A. Experiments

To construct the kinematic triphone model, we used articulatory data obtained in experiments using an electromagnetic articulographic system, Carstens AG 100 (Schönle *et al.*, 1987; Perkell *et al.*, 1992). We observed the following nine points on the midsagittal plane: jaw (J), upper lip (UL),

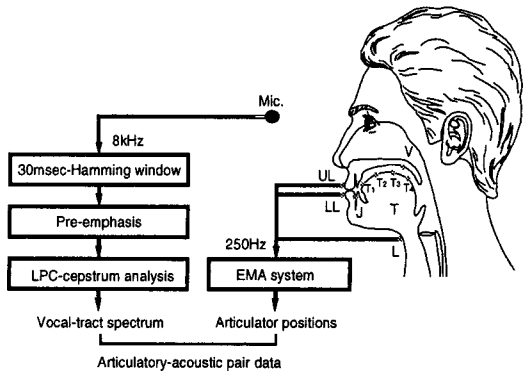


FIG. 1. Experimental system.

lower lip (LL), tongue points (T1, T2, T3, and T4), velum (V), and Adam's apple (L) (see Fig. 1). The jaw receiver was attached to the center of the lower incisors. The most anterior receiver on the tongue was about 5 mm from the tip of the tongue, the most posterior one was almost under the uvula, and the other two were placed at roughly equal intervals between the anterior and posterior coils. The receiver coil on the Adam's apple is for monitoring the height of the larynx. Both the vertical and horizontal orientations of each point were measured at a sampling rate of 250 Hz. (For details of the measurement, see Kaburagi and Honda, 1994.) Two additional receivers placed on the nose and the upper incisors were used for the correction of head movements. We used two receivers attached to a plate in order to record the occlusal plane by having the subject bite on the plate during recording. All data were subsequently corrected for head movements and rotated to bring the occlusal plane into coincidence with the horizontal axis. The speech material in the experiment consisted of 354 sentences that were read at normal speed by three male subjects (HM, OT, and MT).

For each experiment datum, we assigned an articulatory timing to each phoneme. This was done by marking the time at which the ''articulatory feature'' of each phoneme is most prominent. Table I shows the condition that the time at which we put a marker on the time axis to each phoneme satisfies. We first put a marker on the time axis for each phoneme manually, then correct the marker position automatically to satisfy the condition in the table. If we find some minimal points around a manually determined point, we select the middle point among them. For about 50% of vowel occurrences except /a/ and /a:/, there is no minimal point. In such a case, we put the marker in the middle between the previous and successive markers. We call the time assigned to a phoneme the *(observed) articulation timing for the phoneme*. Note that the articulation timing for each phoneme is assigned ''synchronously'' for the articulators. The mean utterance velocities of subjects HM, OT, and MT were, respectively, 17.18, 17.26, and 15.88 phonemes per second.

B. Construction of the triphone model

Using the data taken from reading 338 sentences of the 354, we calculated the position and velocity of the articulator at each observed articulation timing.

TABLE I. The condition that the time at which we put a marker to each phoneme satisfies.

Phonemes	Condition
/a/, /a:/	The vertical velocity of the lower jaw is minimal.
/i/, /i:/, /i̥/, /y/	The vertical velocity of T2 is minimal.
/u/, /u:/, /u̥/, /w/	The vertical velocity of T3 is minimal.
/e/, /e:/	The vertical velocity of T2 is minimal.
/o/, /o:/	The horizontal velocity of T4 is minimal.
/k/, /g/, /ʔk/, /ʔg/, /ŋ/	The distance between the palate and T4 is minimal if the successive phoneme is /o/ or /o:/, otherwise that between the palate and T3 is minimal.
/t/, /d/, /ʔt/, /ʔd/, /n/, /r/, /tʃ/, /ʔtʃ/	The vertical velocity of T4 is minimal.
/s/, /z/, /ʔs/, /dʒ/, /ʃ/, /ʔʃ/	The vertical velocity of T1 is minimal.
/m/, /p/, /b/, /ʔp/, /ʔb/	The distance between the low and upper lips is minimal.
/h/	The same condition as the marker condition that the successive vowel satisfies.
/ɸ/	The horizontal velocity of the upper lip is minimal.
/ɴ/	The vertical velocity of velum is minimal.

We used 40 kinds of phonemes and four special symbols, which represented the onset and release of utterances and the short- and long-silence intervals, respectively. Figure 2 lists the phonemic labels. The 338 sentences used as training sentences contained a total of 11 154 phonemes and 2460 triphones. We evaluated our method by using the remaining 16 sentences. These test sentences contained 507 triphones, 31 of which were not in the training sentences (coverage rate = 93.89%).

To construct a kinematic triphone model, we calculate the average of the positions and that of the velocities of the articulator for each triphone at the observed articulation timing.

III. PRODUCING TRAJECTORIES

A. Kinematic feature extraction

Our method for producing trajectories extracts three kinematic features for each phoneme in a given sequence of phonemes because each phoneme in the sequence is contained in three successive triphones. The kinematic feature of each phoneme in the sequence is determined to be the weighted average of the three kinematic features. Let $\mathbf{x}_i^* = (x^*, v^*)$ be the pair consisting of the predicted position and velocity of the articulator. Note that x^* and v^* consist,

vowel: /a/, /o/, /u/, /e/, /i/
long vowel: /a:/, /o:/, /u:/, /e:/, /i:/
unvoiced vowel: /i̥/, /u̥/
semi-vowel: /w/, /y/
nasal: /m/, /n/, /ŋ/
plosive: /b/, /d/, /g/, /p/, /t/, /k/ /ʧ/, /ts/
fricative: /z/, /h/, /s/, /ɸ/, /dʒ/, /ʃ/
n: /ɴ/
dan-onn: /r/
glottal stop: /ʔt/, /ʔd/, /ʔʃ/, /ʔk/, /ʔp/, /ʔs/, /ʔʧ/.

FIG. 2. Phonemic labels.

respectively, of horizontal and vertical components, x^*_{hori}, x^*_{vert} and v^*_{hori}, v^*_{vert}. For a sequence of phonemes $p_0 p_1,\ldots,p_n$, the kinematic feature of a phoneme $p_i, 0 \leqslant i \leqslant n$, is given by

$$\mathbf{x}_i^* = \frac{\alpha \cdot \mathbf{x}_i^L + \beta \cdot \mathbf{x}_i + \gamma \cdot \mathbf{x}_i^R}{\alpha + \beta + \gamma},$$

where $\mathbf{x}_i^L = (x^L, v^L)$, $x^L = (x^L_{\text{hori}}, x^L_{\text{vert}})$, $v^L = (v^L_{\text{hori}}, v^L_{\text{vert}})$, is the kinematic feature of p_i in triphone $p_{i-2}p_{i-1}p_i$ in the kinematic triphone model, $\mathbf{x}_i = (x, v)$, $x = (x_{\text{hori}}, x_{\text{vert}})$, $v = (v_{\text{hori}}, v_{\text{vert}})$, is that of p_i in $p_{i-1}p_i p_{i+1}$, $\mathbf{x}_i^R = (x^R, v^R)$, $x^R = (x^R_{\text{hori}}, x^R_{\text{vert}})$, $v^R = (v^R_{\text{hori}}, v^R_{\text{vert}})$, is that of p_i in $p_i p_{i+1} p_{i+2}$, and constants α, β, and γ are, respectively, fixed to 1/6, 4/6, and 1/6. Incidentally, we tried to optimize α, β, and γ in such a way that, for each subject, the errors between the observed and predicted trajectories were minimized. The average distance between the observed trajectories and those predicted by using the parameters optimized for each subject is 1.208 mm and that predicted by using the fixed parameters (1/6, 4/6, and 1/6) is 1.209 mm. This difference is quite small; thus, we used constants for all the subjects.

If one or more of three successive triphones containing the phoneme are not in the triphone model, the method uses the remaining triphones, diphones, or a uniphone. That is, for each phoneme in the sequence, if there exist three triphones containing the phoneme, we use the kinematic feature of the phoneme in the triphones. Otherwise, we try to use those for one or two triphones. If there are no triphones, we use the kinematic feature of the phoneme in one or two diphones. If no diphones exist, we use that of the uniphone (phoneme).

The intervals between successive phonemes in the sequence are also determined on the basis of the triphone model. That is, the interval between p_i and p_{i+1}, $0 \leqslant i \leqslant n-1$, is given by the expression $t_i^* = (t_i + t_i')/2$, where t_i is the time interval between p_i and p_{i+1} in triphone $p_{i-1}p_i p_{i+1}$ in the kinematic triphone model; and t_i' is the time interval between p_i and p_{i+1} in triphone $p_i p_{i+1} p_{i+2}$ in the kinematic triphone model.

Let $p_1 p_2,\ldots,p_n$ be a given sequence of phonemes. Using the estimated intervals t_i^* between p_i and p_{i+1}, $0 \leqslant i \leqslant n-1$, we can calculate the *estimated articulation timing* for each phoneme. That is, the estimated articulation timing for p_i, $0 \leqslant i \leqslant n$, is defined by $t_{\text{onset}} + \Sigma_{j=0}^{i-1} t_j^*$, where t_{onset} is the time of the articulation onset denoted by p_0.

B. Minimum-acceleration trajectories

Using each kinematic feature as a constraint, we can formulate the trajectory by calculating the *minimum-acceleration trajectory* of each point on the articulator which coincides with the extremum of the following cost function:

$$\frac{1}{2}\int_0^{t_f}\left(\left(\frac{d^2x}{dt^2}\right)^2 + \left(\frac{d^2y}{dt^2}\right)^2\right)dt, \tag{1}$$

where (x,y) are the time-varying Cartesian coordinates on the sagittal plane of the point on the articulator and t_f is the end time of the movement. The minimum-acceleration model does not need the acceleration of a phoneme in each triphone and, thus, the kinematic triphone model contains only the position and velocity of a phoneme in each triphone.

To determine the trajectory that optimizes the cost function, we use variational calculus and dynamic optimization theory (Pontryagin *et al.*, 1962) in order to obtain a set of linear differential equations. Solving these equations gives us a piecewise polynomial function of time (for details, see the Appendix). The kinematic features are used in the linear computation, which determines the coefficients of the piecewise polynomial function. The trajectory formation method thus produces trajectories simply by extracting kinematic features and using linear computation.

TABLE II. Average distances (mm) between the observed trajectories and the minimum-acceleration trajectories predicted by the kinematic triphone model for the 16 test sentences. The times assigned for observed data are used as the articulation time for each phoneme in the sentences.

	Sentences							
Subject	s1	s2	s3	s4	s5	s6	s7	s8
HM	1.67	1.79	1.44	1.54	1.60	1.71	1.40	1.54
OT	1.92	1.95	2.06	1.76	1.84	1.51	1.55	1.50
MT	1.44	1.46	1.40	2.06	1.49	1.44	1.03	1.36
	s9	s10	s11	s12	s13	s14	s15	s16
HM	1.52	1.62	1.63	1.84	1.77	1.86	1.77	1.66
OT	1.62	1.86	1.50	1.48	1.84	1.91	1.95	1.87
MT	1.52	1.73	1.31	1.48	1.74	1.82	1.92	1.48

IV. EVALUATION OF THE METHOD

A. Distance between observed and predicted trajectories

As noted in Sec. III B, we used the 16 test sentences to evaluate the trajectory formation method. We first discuss the results of predictions calculated by the minimum-acceleration model using the observed articulation timing for each phoneme. Table II lists the average distances between the observed trajectory and that predicted by the kinematic triphone model. The average distances between the observed and predicted trajectories are in the range between 1.03 and 2.06 mm, and the mean of these average distances is 1.65 mm, which is compatible with the average distances between trajectories of articulatory movements observed when the subject reads a sentence twice. Figure 3 shows an example of the predicted and observed trajectories of the articulatory organs. (Note that this trajectory prediction example is calculated on the basis of the triphone model. In Fig. 7, we will show another prediction based on the uniphone model described later.) We can see that the predicted trajectory coincides with the observed one. Figure 4 shows the distances between the observed and predicted articulator positions averaged for each type of phoneme at the observed articulation timing. The distances corresponding to the vowels and the velar consonants are relatively larger than those corresponding to the alveolar and glottal-stop consonants.

We calculate the distances between the observed trajectory and that predicted by the minimum-acceleration model, in which we directly specify the position and velocity of each phoneme by using the observed data. The average dis-

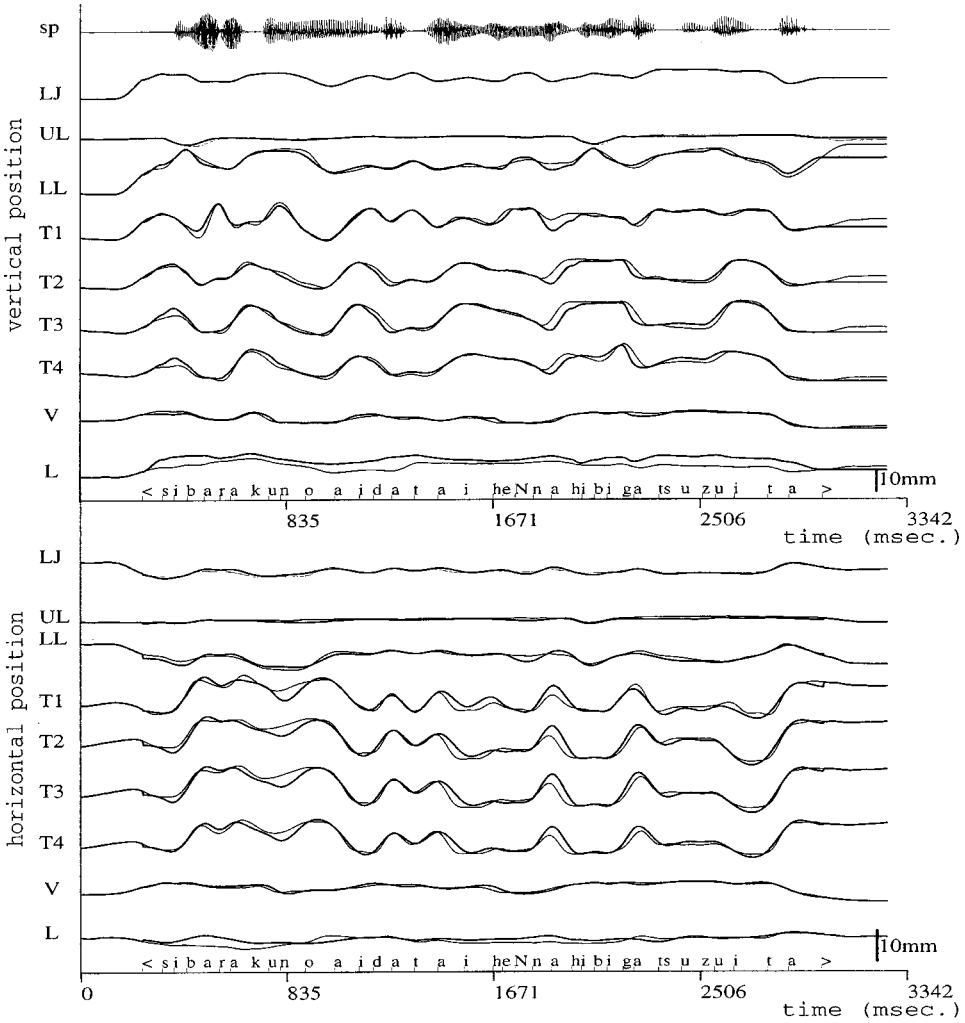


FIG. 3. Trajectories in the utterance of subject MT (thin lines) and the corresponding trajectories predicted by the kinematic triphone model (thick lines). The times assigned for the observed data are used as the articulation time for each phoneme in the sentences. The top trace is the speech waveform (sp), and the others show the movements of jaw (j), upper lip (UL), lower lip (LL), tongue points (T1, T2, T3, and T4), velum (V), and larynx (L).

tances between the minimum-acceleration and observed trajectories are between 0.32 and 0.86 mm, and their mean is 0.52 mm, which is much smaller than that between observed and predicted trajectories by the triphone and minimum-acceleration models. This fact reveals that the errors between the observed trajectory and those predicted by the kinematic triphone and minimum-acceleration models originate mainly in the errors predicted by the kinematic triphone model. Hence, if we can estimate the position and velocity of the articulator more precisely, we can use the minimum-acceleration model to predict articulatory movements with less error.

We now turn back to the trajectory production by the kinematic triphone model and further discuss the recovery of the contextual variability. Let $p_1 p_2,\ldots,p_n$ be a given sequence of phonemes. Assume that, for a phoneme p_i in the sequence, the kinematic triphone model does not contain the triphone $p_{i-1}p_i p_{i+1}$, because none of the 338 sentences contains the triphone. The kinematic triphone model may extract the less accurate kinematic feature of such a pho-

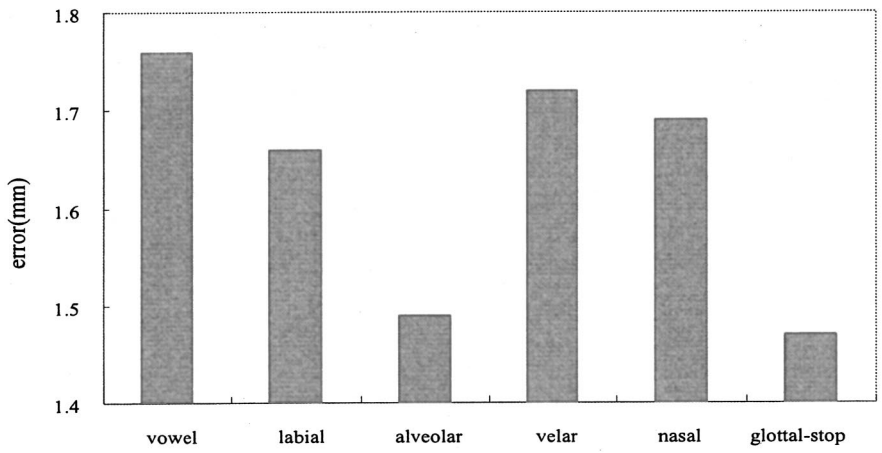


FIG. 4. Average distances (mm) between the observed trajectories and those predicted using by the observed articulation time for each type of phonemes.

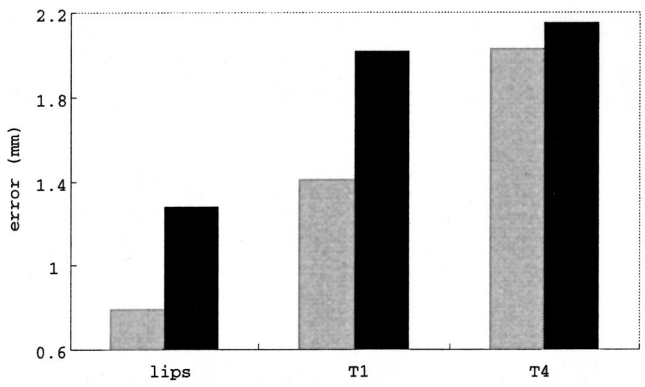


FIG. 5. Average distances between the observed and predicted trajectories at lips, tongue tip (T1), and tongue back (T4) at the observed articulation timings. The black bars show the averaged results for all phonemes. The gray bars indicate those from the labials (lips), the alveolars (T1), and the velars (T4).

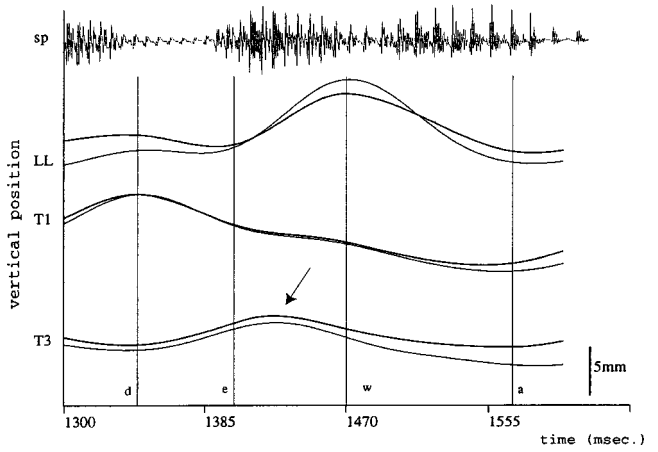


FIG. 6. Asynchronous property of articulators. The top trace is the speech waveform (sp) and the others show the movements of lower lip (LL) and tongue points (T1 and T3) for two production.

neme. We divide the phonemes in the test sentences into two categories: (a) those which a triphone in the triphone model contains and (b) those which any triphones in the triphone model do not contain. For each category of phonemes in the test sentences, we calculate the averaged difference between the observed and predicted positions at the articulation timings. The calculation shows that the position differences are 1.66 mm for the phonemes in category (a) and 2.15 mm for those in category (b), and the difference between the two is statistically significant. The result suggests that constructing the kinematic triphone model by using more sentences containing triphones which are not included in the 338 sentences enables us to improve the prediction accuracy of the method.

B. Qualitative aspects of phonological phenomena

Our method also predicts the qualitative features of the observed trajectories. For example, the method is good for predicting the articulator characteristics that are consistent for each consonant because the variability of the articulatory configuration is small. See, for example, the position of the tongue tip (T1) for uttering the consonants /t/ and /d/ and the position of the dorsum (T3 and T4) for uttering the consonants /k/ and /g/ in Fig. 3. Our method is also particularly good for predicting the fast motion in the release of occlusion for stop consonants. See the position of the tongue tip (T1) for uttering the consonants /t/ and /d/. The black bars in Fig. 5 show the average distances between the observed and predicted trajectories at lips, tongue tip (T1), and tongue back (T4) at the observed articulation timings. The gray bars in Fig. 5 show the average distances at the lips for uttering the labials, at the tongue tip for uttering the alveolars, and at the tongue back for uttering the velars. We can see that the latter distances are smaller than the former distances. Thus, Fig. 5 reveals that our method's predictions reflect articulator characteristics that are consistent for each consonant type.

We next discuss the ''asynchronous (out-of-phase) property'' of the articulators (see, for example, Löfqvist, 1999). That is, when the vertical or horizontal position of an articulator, say the lower lip, is at a minimal or maximal point, the vertical or horizontal position of another articulator, say the tongue tip, is sometimes not at a minimal or maximal point. Figure 6 shows the observed and predicted out-of-phase phenomena. We can see, for example, that when the vertical position of T3 is at a maximal point those of the lower lip and T1 are not at a minimal or maximal point. Our method can reconstruct the asynchronous property of the articulators, although it specifies the kinematic feature of the articulators ''synchronously'' at the articulation timing for each phoneme. This is because it uses the velocity of the articulator as one of the elements of the kinematic features.

C. Comparisons of the current method with other models

We can predict the trajectory of an articulatory movement by the kinematic *uniphone* model, not by the triphone model, by using the minimum-acceleration model. In the uniphone model, we specify the position and velocity of each articulator for each phoneme by using the average value of the position and that of the velocity of the phoneme calculated from the observed data. Figure 7 shows an example of the observed trajectories and the minimum-acceleration trajectories predicted by using the kinematic uniphone model. The average distances between the predicted and observed trajectories are between 2.12 and 3.26 mm, and their mean is 2.64 mm, which is larger than the corresponding value obtained when the triphone model is used (1.65 mm). We also calculate the mean of the distances between the observed trajectories and the minimum-acceleration trajectories predicted by using the kinematic *diphone* model. The calculation shows that the average distance is 1.93 mm. Figure 8 summarizes the average distances between the observed and predicted trajectories. We can see that the predictions of the triphone model are much better than those of the uniphone and diphone models. (Also compare Figs. 3 and 7.)

To calculate the minimum-acceleration trajectory, our method uses the position and velocity of the articulator at the articulation timing for each phoneme. That is, the position and velocity of the articulator are used as boundary conditions for computing the optimal path. We can also calculate the minimum-acceleration trajectory with only the position

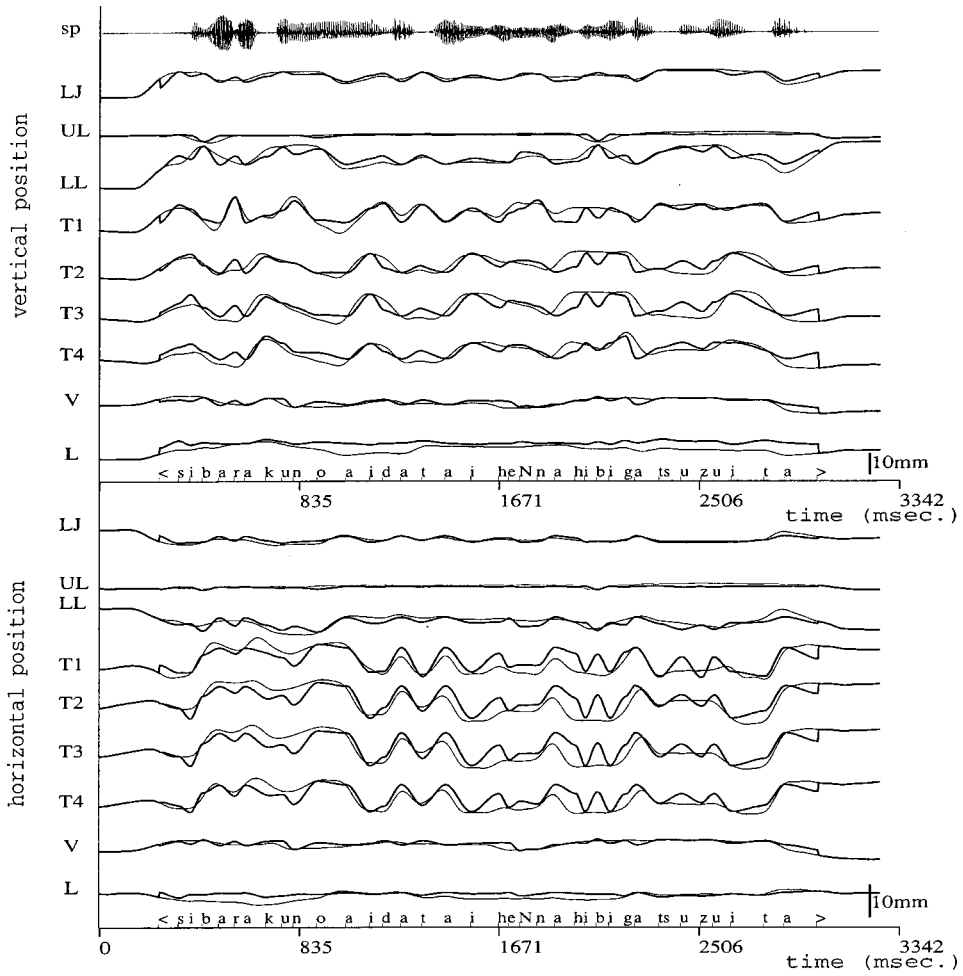


FIG. 7. Observed trajectories in the utterance of subject MT (thin lines) and the minimum-acceleration trajectories predicted by the kinematic uniphone model (thick lines). The observed articulation timings are used. The top trace is the speech waveform (sp), and the others show the movements of jaw (LJ), upper lip (UL), lower lip (LL), tongue points (T1, T2, T3, and T4), velum (V), and larynx (L).

of the articulation as a boundary condition. In this calculation, we assume that the velocity of the articulator is continuous at the boundary (at the articulation timing for each phoneme). The average distance between the observed and trajectories predicted by the modified model is 1.73 mm (subjects HM: 1.75 mm; OT: 1.82 mm; MT: 1.62 mm). The prediction of the original kinematic triphone model (1.65 mm) is thus better than that of the modified model (1.73 mm).

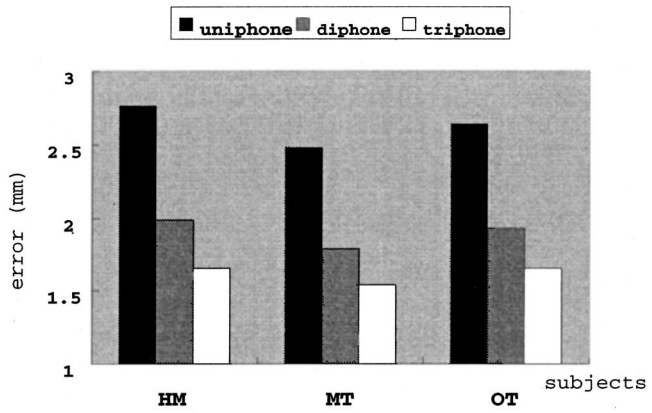


FIG. 8. Average distances between the observed trajectories and the minimum-acceleration trajectories predicted by the kinematic uniphone, diphone, and triphone models.

D. Articulation timing

Figure 9 shows an example of the observed trajectories and those predicted by the minimum-acceleration model, in which the estimated times were used as the articulation time for each phoneme in the sentences. The method predicts the qualitative features observed experimentally, although the predicted trajectory shifts to the right or left of the observed one in some parts. For each of two successive phonemes in the test sentences, we listed the observed interval times between the successive phonemes. Table III shows the percentages of the predicted intervals inside the permitted range defined by the observed interval ± 20 ms. It also lists those of the range between the observed interval ± 12 ms. We can see that about 85% of the predicted intervals are inside the permitted range.

A further analysis of the errors of the interval times between two successive phonemes reveals that the errors between the successive phonemes including the vowels, the long vowels, and the geminative consonants are relatively larger than those between the other consonants. Among these three types of phonemes in Japanese, only the vowels appear at the word end. We thus divide the vowels occurring in the test sentences into two categories: vowels at the word end and those not at the word end.

Table IV lists the standard deviations of the intervals between the articulation time of a vowel at the word end and

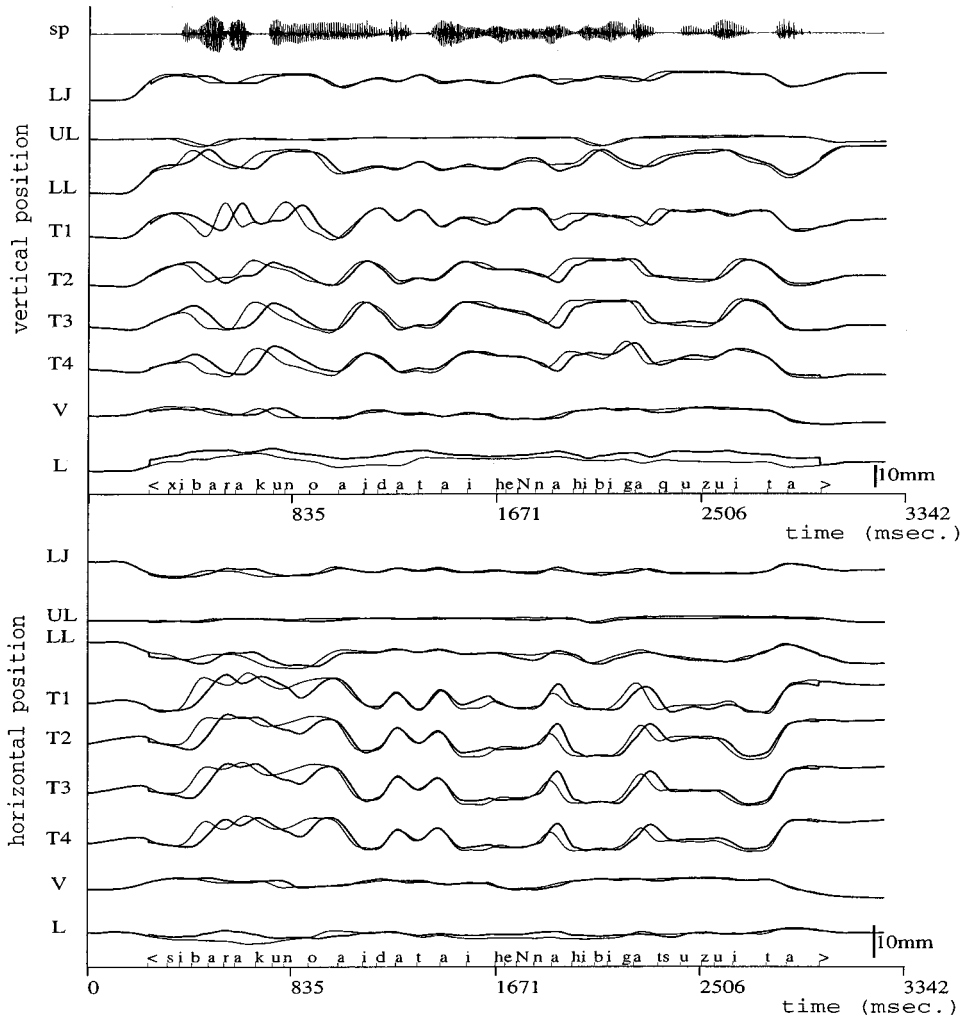


FIG. 9. Observed trajectories in the utterance of subject MT (thin lines) and the minimum-acceleration trajectories predicted by the kinematic triphone model (thick lines). The estimated articulation timings are used. The top trace is the speech waveform (sp), and the others show the movements of jaw (LJ), upper lip (UL), lower lip (LL), tongue points (T1, T2, T3, and T4), velum (V), and larynx (L).

that of the successive phoneme (word-final intervals). It also lists those for a vowel that is not the last phoneme in a word (word-inner intervals). We can see that the standard deviations of intervals for the word-final are larger than those for the word-inner. The differences between those standard deviations are all statistically significant or nearly significant. This result suggests that, in Japanese, a word is an utterance unit.

The above result also suggests that the word-final intervals predicted by the triphone model are relatively more inaccurate than the word-inner ones. We correct the word-final intervals predicted by the triphone model in the following way. We first classify the observed word-final intervals into five categories: very large, large, normal, small, and very small. We then calculate the respective averages and use the averaged values of the word-final intervals instead of the prediction by the triphone model. We predict intervals by the triphone model with this word-final intervals correction.

TABLE III. Percentages of predicted intervals inside the permitted range defined by the observed interval ±20 ms and those inside the range between the observed interval ±12 ms.

	HM	OT	MT
±20 ms	84.3%	86.9%	85.3%
±12 ms	62.8%	65.1%	64.2%

Table V lists the percentages of the intervals inside the permitted range defined by the observed interval ±20 ms and those for the range between the observed interval ±12 ms. This table shows the interval prediction is improved in comparison with that by the kinematic model without the word-final correction. This result implies that we can greatly improve the timing prediction of the method if we know the rule that determines to which category a word-final interval belongs.

E. Production of acoustics from predicted articulation

We next examine how much an error in articulator position affects acoustics. This is done by using a mapping from articulator position to acoustics on the basis of an articulatory-acoustic codebook. The articulatory-acoustic codebook is designed, again, using the simultaneous observed data of articulatory motions and speech acoustics through EMA experiments. The codebook contains pair data of acoustic and articulatory segments. That is, in the codebook, each articulatory segment is associated with an acoustic segment. The interval time of the segment spans a fixed period.

Let us denote an acoustic parameter at time t by c^t and an articulatory parameter by x^t. The parameter c^t is a vector consisting of 30 LPC cepstral coefficients and x^t is a vector

TABLE IV. Standard deviations (ms) of the intervals between the articulation time of a vowel at the word end and that of the successive phoneme (word-final intervals) and those of the intervals between the articulation time of a vowel which is not the last phoneme of a word and that of the successive phoneme (word-inner intervals).

/a/	HM	OT	MT
Word-inner	18.96	28.47	32.20
Word-final	67.93	54.91	55.84
$F(38,39)$	3.58	1.93	1.73
/u/	HM	OT	MT
Word-inner	20.40	29.05	40.44
Word-final	65.93	66.79	65.07
$F(70,61)$	3.23	2.30	1.61
/i/	HM	OT	MT
Word-inner	21.50	29.05	37.65
Word-final	59.83	56.44	52.52
$F(70,61)$	2.78	1.94	1.39
/e/	HM	OT	MT
Word-inner	26.61	30.91	40.83
Word-final	55.42	58.80	58.73
$F(110,92)$	2.08	1.90	1.44
/o/	HM	OT	MT
Word-inner	25.93	25.81	35.19
Word-final	60.08	58.62	64.84
$F(144,121)$	2.32	2.27	1.84

of 18th order whose elements are represented by the vertical and horizontal position of nine points on the midsagittal plane described in Sec. II A. Let a sequence $\mathbf{c}^t = c^{t-m}, c^{t-m+1}, \ldots, c^{t-1}, c^t, c^{t+1}, \ldots, c^{t+m}$ be an acoustic (parameter) segment at t whose length M is $2m+1$. Likewise, let $\mathbf{x}^t = x^{t-m}, x^{t-m+1}, \ldots, x^{t-1}, x^t, x^{t+1}, \ldots, x^{t+m}$ be an articulatory (parameter) segment at t whose length is $M = 2m+1$.

The articulatory-acoustic codebook $\mathcal{D}$ is represented by

$$\mathcal{D} = \{\langle \mathbf{x}^0, \mathbf{c}^0 \rangle, \langle \mathbf{x}^1, \mathbf{c}^1 \rangle, \ldots, \langle \mathbf{x}^T, \mathbf{c}^T \rangle\},$$

where $\langle \mathbf{x}^t, \mathbf{c}^t \rangle$ is a pair of code vectors and T is the number of codes in the codebook.

Now, we describe the algorithm which estimates a sequence of acoustic parameters from that of articulatory parameters. Let us denote an articulatory segment at t by $\mathbf{x}_{in}(t) = x_{in}(t-m), \ldots, x_{in}(t), \quad x_{in}(t+1), \ldots, x_{in}(t+m) \quad (M = 2m+1)$.

At every time t, we compare an articulatory segment with the articulatory segment part of the codes in the codebook, and we calculated the square sum error between them. We then selected a fixed number of candidates of acoustic segments from the articulatory-acoustic pair code segments

TABLE V. Percentages of the predicted intervals inside the permitted range defined by the observed interval ±20 ms and those inside the range defined by the observed interval ±12 ms. The prediction is based on the kinematic model with the word-final correction.

	HM	OT	MT
±20 ms.	86.7%	89.9%	88.5%
±12 ms.	71.6%	72.8%	72.3%

TABLE VI. The averaged LPC cepstrum distances (dB) between the observed and estimated acoustic parameters. The estimated acoustic parameters are produced from trajectories predicted by the uniphone, diphone, and triphone models. Those produced from observed trajectories are also shown.

	HM	OT	MT
Uniphone	3.38	3.52	3.63
Diphone	2.91	3.09	3.12
Triphone	2.74	2.95	2.98
Observation	2.66	2.95	2.84

in order of increasing the square sum error. That is, for each time $t = 0,1,\ldots,n$, we select a set of pairs of code vectors $S(t) = \{\langle \mathbf{x}_1, \mathbf{c}_1 \rangle, \langle \mathbf{x}_2, \mathbf{c}_2 \rangle, \ldots, \langle \mathbf{x}_l, \mathbf{c}_l \rangle\}$, such that $\langle \mathbf{x}_1, \mathbf{c}_1 \rangle, \ldots, \langle \mathbf{x}_l, \mathbf{c}_l \rangle$ is in $\mathcal{D}$ and

$$\max_{\mathbf{x} \in \{\mathbf{x}_1, \ldots, \mathbf{x}_l\}} e(\mathbf{x}, \mathbf{x}_{in}(t)) < \min_{\langle \mathbf{x}', \mathbf{c}' \rangle \in \mathcal{D}'} e(\mathbf{x}', \mathbf{x}_{in}(t))$$

is satisfied, where $\mathcal{D}' = \mathcal{D} - S(t)$, l denotes the number of candidates, and $e(\mathbf{x}, \hat{\mathbf{x}})$ is the square sum error between $\mathbf{x}$ and $\hat{\mathbf{x}}$ expressed by

$$e(\mathbf{x}, \hat{\mathbf{x}}) = \sum_{i=1}^{M} \sum_{j=1}^{p} (x_{ij} - \hat{x}_{ij})^2.$$

From $S(t)$, we can obtain candidate acoustic segments $C(t) = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_l\}$.

We then extract the center of the acoustic segments in $C(t)$ and thus we obtain l LPC cepstral coefficient vectors for each time. That is, let $\mathbf{c}_i = \{c_{i1}, c_{i2}, \ldots, c_{il}\}$, $i = 0,\ldots,M$. We can construct $A(t) = \{c_1, c_2, \ldots, c_l\}$, $t = 0,\ldots,n$, where $c_i = c_i M/2$. Finally, we calculate the centroid of the elements in $A(t)$. We denote the centroid of the elements in $A(t)$ by $A_c(t)$. The sequence $A_c(0), A_c(1), \ldots, A_c(n)$ is of the estimated acoustic parameters.

Again, using the data taken from reading the same 338 sentences, we construct the articulatory-acoustic codebook and we use the data for the same 16 test sentences to evaluate the estimated acoustics. Table VI shows the averaged LPC cepstrum distances between the observed and estimated acoustic parameters. The averaged LPC cepstrum distance between two acoustics represented by p-order LPC-cepstrum $(c_{1_1}, \ldots, c_{1_p})$ and $(c_{2_1}, \ldots, c_{2_p})$, is defined by

$$CD = \frac{10 \cdot \sqrt{2 \cdot \Sigma_{j=1}^{p} (c_{1_j} - c_{2_j})^2}}{\ln 10}.$$

To calculate the distances shown in this table, we fix the number of candidate vectors to 25 and the length of articulatory and acoustic segments to 56 ms. In Table VI, estimated acoustic parameters are calculated from the observed and predicted articulatory positions by the uniphone, diphone, and triphone models using observed articulation timings.

The averaged LPC cepstrum distance between observed acoustic parameters and those estimated from articulatory positions predicted by the triphone model is 2.89 dB (the averaged distance between observed and predicted trajectories is 1.65 mm); that by diphone model is 3.04 dB (the averaged distance between observed and predicted trajectories is 1.93 mm); that by uniphone model is 3.51 dB (the

averaged distance between observed and predicted trajectories is 2.64 mm). Furthermore, we can see that distances between observed acoustic parameters and those estimated from articulatory positions predicted by the triphone model are the same as those from the observed positions (the average: 2.89 dB vs 2.82 dB). These results suggest that the averaged position error of 1.65 mm is good enough to reproduce acoustics from a trajectory of articulation and that of 2.0 mm is a critical point for production of acoustics.

V. CONCLUDING REMARKS

We have developed a method of calculating the trajectory of articulator movements for continuous speech utterances at normal speed. Each phoneme-specific task is specified by a kinematic triphone model constructed on the basis of experimental data obtained by an electro-magnetic articulographic system, and the trajectories of articulator movements are determined by minimizing the acceleration of each point on the articulator organs. The method is accurate and quick and predicts both the qualitative features and the quantitative details experimentally observed.

We show that the errors between the observed trajectory and those predicted by the the kinematic triphone and minimum-acceleration models originate mainly in the errors originating from the kinematic triphone model. We thus believe that we can predict articulatory movements by the minimum-acceleration model with less error, if we can estimate the position and velocity of the articulator more precisely, for example, by using more triphones constructed from a larger amout of articulatory data.

Using the articulatory-acoustic database, we produced acoustic parameters from a predicted articulatory movement. The acoustic distance analysis reveals that reproduced acoustics from articulatory positions predicted by the triphone model are similar to those from the observed positions.

Although our method accurately predicts the articulatory movements, it has some problems. One problem is that it predicts the trajectory only of utterances at normal speed. The reduction of articulatory movements for uttering vowels can be observed as the utterance speed increases (for example, Lindblom, 1963). We described a relation between utterance speed and the amount of articulatory reduction and proposed a linear-regression model that predicts reduced articulatory movements (Okadome *et al.*, 1999). To predict the articulatory movement for fast or slow utterances, we plan to improve the linear-regression model and to combine it with the minimum-acceleration model.

The second problem concerns the structure of languages. Our method uses the kinematic triphone model to predict articulatory movements for a given phonemic sequence in Japanese. A Japanese phonological unit is considered to be a CV syllable. On the other hand, in languages such as English, French, or Spanish, a variable-length syllable is considered to be a unit (for example, Fujimura, 1994; Krakow, 1999). We thus speculate that, with the minimum-acceleration model, a kinematic ''syllable model'' which uses syllables instead of triphones may be better for predicting the trajectory of the articulatory movements for such languages.

We finally point out that the method can be applied to predict actual articulator movements. And, the predicted trajectory of articulatory movements can be directly compared with the observed trajectory. It is thus possible to quantitatively evaluate the capability of the method.

APPENDIX: CALCULATION OF THE MINIMUM-ACCELERATION TRAJECTORY

Let $\xi=(x,y)$ be the time-varying Cartesian coordinates of a point p. We determine the trajectory which minimizes

$$\frac{1}{2}\int_0^{t_f}\left(\left(\frac{d^2x}{dt^2}\right)^2+\left(\frac{d^2y}{dt^2}\right)^2\right)dt,$$

where the time interval $[0,t_f]$ is divided into $t=t_0,t_1,t_2,\ldots,$ and $t_n=t_f$ and at each t_i, $i=0,...,n$, the position (x_i,y_i) and the velocity $(\dot{x}_i,\dot{y}_i)$ are given. In general, for a cost function $L[t,\xi,\dot{\xi},\ldots,d^n\xi/dt^n]$, the trajectory $\xi(t)$ which minimizes

$$\int_{T_1}^{T_2}L\left[t,\xi,\dot{\xi},\ldots,\frac{d^n\xi}{dt^n}\right]dt$$

satisfies the following Euler–Poisson equation:

$$\frac{\partial L}{\partial\xi}-\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\xi}}\right)+\cdots+(-1)^n\frac{d^n}{dt^n}\left(\frac{\partial L}{\partial\xi^{(n)}}\right)=0,$$

where $\xi^{(n)}=d^n\xi/dt^n$. When $L=\frac{1}{2}((d^2x/dt^2)^2+(d^2y/dt^2)^2)$, we obtain

$$\frac{d^2}{dt^2}\left(\frac{\partial(x^{(2)})^2}{\partial x^{(2)}}\right)+\left(\frac{\partial(y^{(2)})^2}{\partial y^2}\right)=0,$$

and thus

$$\frac{d^4x}{dt^4}=0,\quad\frac{d^4y}{dt^4}=0.$$

Solving these equations gives the following time-varying functions:

$$x(t)=a_0+a_1t+a_2t^2+a_3t^3,$$

$$y(t)=b_0+b_1t+b_2t^2+b_3t^3.$$

If we give, as constraints, $x(T_1)$, $\dot{x}(T_1)$, $x(T_2)$, $\dot{x}(T_2)$, $y(T_1)$, $\dot{y}(T_1)$, and $y(T_2)$, $\dot{y}(T_2)$ then we can determine the coefficients $a_0,\ldots,a_3,b_0,\ldots,$ and b_3 uniquely. Hence, for each interval $[t_i,t_{i+1}]$, a trajectory which satisfies (x_i,y_i) and $(\dot{x}_i,\dot{y}_i)$ at each t_i, $i=0,...,n$, and minimizes $L=\frac{1}{2}((d^2x/dt^2)^2+(d^2y/dt^2)^2)$ is determined uniquely. Clearly, the trajectory constructed by piecewisely joining these trajectories is the solution to our problem.

Bailly, G., Laboissiere, R., and Schwartz, J. L. (**1991**). ''Formant trajectories as audible gestures: An alternative for speech synthesis,'' J. Phonetics **19**, 9–23.

Browman, C. P., and Goldstein, L. (**1992**). ''Targetless schwa: An articulatory analysis,'' in *Papers in Laboratory Phonology II*, edited by G. J. Docherty and D. R. Ladd (Cambridge University Press, Cambridge), pp. 26–67.

Fujimura, O. (**1994**). ''C/D Model: A computational model of phonetic implementation,'' in *Language and Computations*, edited by E. S. Ristad (American Math Society, Providence), pp. 1–20.

Flash, T., and Hogan, N. (**1985**). ''The coordination of arm movements: An experimentally confirmed mathematical model,'' J. Neurosci. **5**, 1688–1703.

Hirayama, M., Vatikiotis-Bateson, E., Kawato, M., and Honda, K. (**1992**). ''Neural network modeling of speech motor control,'' in Proceedings of the Second International Conference on Spoken Language Processing, pp. 883–886.

Honda, M., and Kaburagi, T. (**1994**). ''A dynamical articulatory model using potential task representation,'' in Proceedings of the Third International Conference on Spoken Language Processing, pp. 179–182.

Honda, M., and Kaburagi, T. (**1996**). ''Statistical analysis of a phonemic target in articulatory movements,'' J. Acoust. Soc. Am. **100**, 2598(A).

Jordan, M. (**1989**). ''Indeterminate motor skill problems,'' in *Attention and Performance*, edited by M. Jeannerod (Erlbaum, Hillsdale, NJ), Vol. XIII.

Kaburagi, T., and Honda, M. (**1994**). ''Determination of sagittal tongue shape from the positions of points on the tongue surface,'' J. Acoust. Soc. Am. **96**, 1356–1366.

Kaburagi, T., and Honda, M. (**1996**). ''A model of articulatory trajectory formation based on the motor tasks of vocal-tract shapes,'' J. Acoust. Soc. Am. **99**, 3154–3170.

Kaburagi, T., Honda, M., and Okadome, T. (**1999**). ''A trajectory formation model of articulatory movements using multidimensional phonemic task,'' in *Proceedings of the Sixth European Conference on Speech Communication and Technology* (ESCA, Budapest), Vol. 1, pp. 121–124.

Kent, R. D. (**1983**). ''The segmental organization of speech,'' in *The Production of Speech*, edited by P. F. MacNeilage (Springer, New York).

Krakow, R. A. (**1999**). ''Physiological organization of syllables: A review,'' J. Phonetics **27**, 23–54.

Lindblom, B. (**1963**). ''Spectrographic study of vowel reduction,'' J. Acoust. Soc. Am. **35**, 1773–1781.

Löfqvist, A. (**1999**). ''Interarticulator phasing, locus equations, and degree of coarticulation,'' J. Acoust. Soc. Am. **106**, 2022–2030.

Okadome, T., and Honda, M. (**1992**). ''Trajectory formation in sequential arm movements,'' in Proceeding of the IEEE International Conference SMC, 471–478.

Okadome, T., Kaburagi, T., and Honda, M. (**1998**). ''Trajectory formation of articulatory movements for a given sequence of phonemes,'' in Proceedings of the Fifth International Conference on Spoken Language Processing (Sydney), Vol. 7, pp. 3131–3134.

Okadome, T., Kaburagi, T., and Honda, M. (**1999**). ''Relations between utterance speed and articulatory movements,'' in *Proceedings of the Sixth European Conference on Speech Communication and Technology* (ESCA, Budapest), Vol. 1, pp. 137–140.

Perkell, J. S., Cohen, M. H., Svirsky, M. A., Matthies, M. L., Garabieta, I., and Jackson, M. T. T. (**1992**). ''Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements,'' J. Acoust. Soc. Am. **92**, 3078–3096.

Perkell, J. S., Matthies, M. L., Svirsky, M. A., and Jordan, M. I. (**1993**). ''Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot motor equivalence study,'' J. Acoust. Soc. Am. **93**, 2948–2961.

Pontryagin, L. S., Boltyanski, V. G., Gamkrelidze, R. V., and Mischenko, E. F. (**1962**). *The Mathematical Theory of Optimal Processes* (Interscience, New York).

Saltzman, E., and Munhall, K. G. (**1989**). ''A dynamical approach to gestural pattering in speech production,'' Ecological Psychol. **1**, 333–382.

Schönle, P. W., Gräbe, K., Wening, P., Höhne, J., Schrader, J., and Conrad, B. (**1987**). ''Electromagnetic articulography: Use of alternating magnetic fields for tracking movements of multiple points inside and outside the vocal tract,'' Brain Lang. **31**, 26–35.

Stevens, K. N. (**1989**). ''On the quantal nature of speech,'' J. Phonetics **17**, 3–46.

Vatikiotis-Bateson, E., Hirayama, M., and Kawato, M. (**1992**). ''The articulatory dynamics of running speech: Gestures from phonemes?,'' in Proceedings of the Second International Conference on Spoken Language Processing, pp. 887–890.

Wickelgren, W. A. (**1969**). ''Context-sensitive coding, associative memory, and serial order in (speech) behavior,'' Psychol. Rev. **76**, 1–15.

Sex-specific fundamental and formant frequency patterns in a cross-sectional study

Sandra P. Whiteside[a)]
Department of Human Communication Sciences, University of Sheffield, Sheffield S10 2TA, United Kingdom

(Received 13 July 2000; accepted for publication 18 April 2001)

An extensive developmental acoustic study of the speech patterns of children and adults was reported by Lee and colleagues [Lee *et al.*, J. Acoust. Soc. Am. **105**, 1455–1468 (1999)]. This paper presents a reexamination of selected fundamental frequency and formant frequency data presented in their report for ten monophthongs by investigating sex-specific and developmental patterns using two different approaches. The first of these includes the investigation of age- and sex-specific formant frequency patterns in the monophthongs. The second, the investigation of fundamental frequency and formant frequency data using the critical band rate (bark) scale and a number of acoustic-phonetic dimensions of the monophthongs from an age- and sex-specific perspective. These acoustic-phonetic dimensions include: vowel spaces and distances from speaker centroids; frequency differences between the formant frequencies of males and females; vowel openness/closeness and frontness/backness; the degree of vocal effort; and formant frequency ranges. Both approaches reveal both age- and sex-specific development patterns which also appear to be dependent on whether vowels are peripheral or nonperipheral. The developmental emergence of these sex-specific differences are discussed with reference to anatomical, physiological, sociophonetic, and culturally determined factors. Some directions for further investigation into the age-linked sex differences in speech across the lifespan are also proposed. [DOI: 10.1121/1.1379087]

PACS numbers: 43.70.Ep, 43.70.Gr [AL]

I. INTRODUCTION

Sex differences in the formant frequency values of adults are well established and widely documented (e.g., Childers and Wu, 1991; Deterding, 1997; Peterson and Barney, 1952; Wu and Childers, 1991). In addition, the nonuniform patterns of sex differences across different formant frequencies and vowels are well-established and have been observed across different languages (Fant, 1975; Traunmüller, 1984, 1988). These nonuniform sex differences highlight the nonlinear sex differences in the vocal tract dimensions of men and women (Fant, 1966, 1975; Fitch and Giedd, 1999) and explain the developmental emergence of nonlinear sex differences in the formant frequency data of preadolescent children (e.g., Bennett, 1981; Busby and Plant, 1995; Eguchi and Hirsh, 1969; White, 1999; Whiteside and Hodgson, 2000).

There is some acoustic-phonetic evidence which suggests that women have more peripheral vowel spaces compared to men (e.g., Deterding, 1997; Henton, 1995; Traunmüller, 1988), and this seems to hold true across a number of different languages including General American English, Italian, Japanese, British English (Middle Northern), German, Swedish, Standard Dutch, and French [e.g., see Henton (1995) and Rosner and Pickering (1994) for details and summaries of these data]. The more peripheral nature of vowel spaces exhibited by women has been attributed to sociophonetic factors which determine the different speech styles adopted by men and women (Henton, 1995). However, an alternative view is that the emergence of sex differences in the acoustic-phonetic characteristics of vowels can be explained by physiological factors and anatomical constraints, which are due to maturational differences between males and females (Traunmüller, 1984, 1988).

A recent and extensive contribution to the developmental literature on speech characteristics comes from Lee and colleagues (Lee *et al.*, 1999), who report on speech data collected from 436 children aged 5 to 18 years, and 56 adults (aged 25–50 years). Their study includes data on fundamental frequency and formant frequency data from ten monophthongs of American English. In their presentation and discussion of their data, the authors mention the need for a more detailed and thorough investigation of their data.

This paper aims to reexamine and elaborate on some of the fundamental and formant frequency data presented by Lee *et al.* (1999) by adopting a sex-specific developmental perspective, and exploring some of the factors that may be responsible for age-linked sex differences in these acoustic-phonetic parameters. The fundamental frequency and formant frequency data presented by Lee *et al.* (1999) are reexamined using a number of approaches. The first of these involves the application of *kn* factors (Fant, 1966, 1975), and the second, the investigation of vowel spaces and acoustic-phonetic parameters using the critical band rate (bark) scale (Traunmüller, 1988, 1990). Both approaches adopt an age- and sex-specific perspective.

In order to investigate within-sex patterns as a function of development, *kn-age* factors are derived for each of the male and female formant frequency data following Fant (1966, 1975). The motivation here is to chart and gauge the

a)Electronic mail: s.whiteside@sheffield.ac.uk

 0001-4966/2001/110(1)/464/15/$18.00

developmental patterns of formant frequencies of males and females separately, as a function of chronological age. Further, to investigate the emergence of sex differences across the lifespan, *kn-sex* factors are examined by gauging the male–female differences in formant frequencies across all vowels and selected vowels for the different age groups. The findings of this reanalysis of the formant frequency data reported by Lee *et al.* (1999) are presented and discussed with reference to developmental patterns previously reported for formant frequencies (e.g., Bennett, 1981; Busby and Plant, 1995; Eguchi and Hirsh, 1969; White, 1999), and developmental patterns in the morphology of the human vocal tract (e.g., Fitch and Giedd, 1999).

The formant frequency data reported by Lee *et al.* (1999) are also investigated using an approach which determines the distances of formant frequency values from speaker centroids as a measure of vowel space (e.g., Deterding, 1997). Acoustic-phonetic parameters expressed in the critical band rate (bark) scale have been shown to be effective in highlighting the ontogeny of sex-specific variation in the vowel quality of peripheral vowels (Traunmüller, 1988). Given that the ten monophthongal vowels reported by Lee *et al.* (1999) included both peripheral and nonperipheral vowels, the acoustic-phonetic quality of peripheral and nonperipheral vowels is examined separately using a number of acoustic-phonetic parameters based on the critical band rate (bark) scale from a sex-specific and developmental perspective. The patterns that emerge from this reexamination are discussed with reference to the sex-linked developmental trends in the data, and whether they shed any light on any factors that may be responsible for influencing the emergence of speaker sex differences in the phonetic quality of vowels (e.g., Henton, 1995; Rosner and Pickering, 1994; Traunmüller, 1988).

II. METHODS

A. *Kn*-age and *kn*-sex factors

The mean formant frequency values ($F1$ to $F3$) for the monophthongs of all age groups from age 7[1] years to those of the adults, reported by Lee *et al.* (1999) were examined for vowel-specific patterns, and related sex-specific and developmental changes. The ten monophthongs reported by Lee *et al.* (1999, p. 1456) are as follows: aa (*pot*); ae (*bat*); ah (*but*); ao (*ball*); eh (*bet*); er (*bird*); ih (*bit*); iy (*bead*); uh (*put*); uw (*boot*).[2]

Following Fant (1966, 1975) scaling factors were calculated for each of the formant frequency values ($F1$,$F2$,$F3$), and across all three formant frequencies ($(F1+F2+F3)/3$) to examine both within-sex-age-linked developmental patterns, and sex-specific developmental changes. Within-sex-age-linked developmental patterns were examined by applying the formula (1) to give two sets of *kn-age* values for the male and female speakers. These *kn-age* values allowed the examination of developmental changes in formant frequency values for each sex group, with reference to the adult formant frequency values. In addition, within-age sex-linked developmental patterns for formant frequency values were examined by using formula (2) to give one set of *kn-sex* factors. These *kn-sex* factors allowed the tracking of developmental patterns in within-age sex differences with increasing chronological age. Both *kn-age* and *kn-sex* factors were examined in more detail for a selected group of vowels which represented a range of acoustic-phonetic vowel quality dimensions of openness, constriction, rhoticity, frontness, and backness. These selected vowels were aa, ah, er, ih, iy, and uw.

$$Kn = Fn/\text{Ref } Fn. \quad (1)$$

Fn: Female or male formant frequency value,

Ref Fn: Formant frequency value of adult female (in the case of female Fn value) or adult male (in the case of male Fn value) formant frequency value. For example, the mean adult male value of $F1$ (vowel aa, 723 Hz) serves as the reference value for the corresponding $F1$ value for the male 18-year-olds (vowel aa, 737 Hz), to derive a $k1$ age factor value of 1.02 for the male 18-year-olds.

$$Kn = Fn/\text{Ref } Fn. \quad (2)$$

Fn: Female formant frequency value,

Ref Fn: Formant frequency value of male age peer formant frequency value. For example, the male value of $F1$ (vowel aa, 723 Hz) for the 18-year-old group serves as the reference value for the corresponding $F1$ value for the female 18-year-old group (vowel aa, 894 Hz). This therefore gives a $k1$ sex factor of 1.26 for the 18-year-olds.

B. Conversion of fundamental frequency and formant frequency values from hertz to bark to examine vowel spaces and critical band rate (bark) distances

The aim here was to investigate sex-linked developmental patterns in the vowel spaces of the male and female speakers of each age group. This reexamination was carried out as follows. First, the fundamental frequency ($F0$) and formant frequency data (in hertz) for the ten vowels reported by Lee *et al.* (1999) were converted to the Bark auditory scale, using the formula described by Traunmüller (1988, 1990). This formula for the Bark scale is an alternative to that documented by Zwicker and Terhardt (1980), where Z is the frequency in the Bark scale and F is the frequency in hertz as given in formula (3). In addition, all fundamental frequency values less than 2.0 bark were corrected using formula (4) (Traunmüller, 1990).

$$Z = [26.81F/(1960+F)] - 0.53. \quad (3)$$

$$\text{For calculated } Z < 2.0 \text{ bark: } Z' = Z + 0.15(2-Z). \quad (4)$$

Subsequently, vowel spaces of the 13 age groups were determined for the males and females of each age group using the critical band rate (bark) values. This was done by plotting the difference between $F1$ and $F0$ ($F1-F0$) as a speaker-independent index of degree of openness (Traunmüller, 1981), against the difference between $F2$ and $F1$ ($F2-F1$), as a general index of anterior/posterior position of constriction (Ladefoged and Maddieson, 1990). This method was used to normalize for sex differences, and was

therefore chosen to highlight any developmental sex differences within the defined $F1-F0/F2-F1$ vowel space of the ten vowels after normalization.

In order to determine tonotopic distances between vowels in the vowel space defined by $F2-F1/F1-F0$, the speaker centroids of the vowel space for the male and female speakers were calculated separately for each age and sex group from the mean $F1-F0$ and $F2-F1$ values of all ten vowels, formula (5). In (5) D refers to the distance of vowel V from the speaker centroid of each age and sex group vowel space, *xcentroid* to the $F2-F1$ coordinate of the speaker centroid, *ycentroid* to the $F1-F0$ coordinate of the speaker centroid, xV to the $F2-F1$ value for vowel V, and yV to the $F1-F0$ value for vowel V. The application of formula (5) gleaned a total of 26 sets of speaker centroid values and vowel distances [2 (males and females)×13 (age groups)].

$$D=\sqrt{((xV-xcentroid)^2+(yV-ycentroid)^2)}. \quad (5)$$

C. Investigating critical band rate (bark) distances and variation in vowel quality in peripheral and nonperipheral vowels from a sex-linked developmental perspective

Using a selection of methods, critical band rate (bark) distances and the variation of the ten vowels were also investigated in greater detail from a sex-specific developmental perspective by examining sex-specific developmental patterns before puberty (ages 7 to 12 years), during puberty (ages 13 and 14 years), after puberty (ages 15 to 18 years), and in adulthood (age 25 to 50 years). A number of acoustic-phonetic dimensions were investigated for both the peripheral (aa, ae, ao, eh, iy, and uw) and nonperipheral (ah, er, ih, uh) vowels as a function of age group and sex. Details and the motivation for these dimensions follow: (i) $F1$ and $F0$ have been identified as major cues to the perceived phonetic openness of a vowel. Therefore, vowel openness expressed as a function of $F0$ (in bark), and the degree of standard deviation in the critical band rate of $F1$ (as an index of the extent of variation in the open–close dimension of vowel quality) was examined (Traunmüller, 1988); (ii) The critical band rate of $F3$ can be taken to represent speaker size as it decreases with increasing vocal tract length. The standard deviation values of the critical band rates of $F1$ index the open–close dimension of vowel quality, those of $F2$, the front–back dimension of vowel quality, and those of $F3$, the degree of rhoticity for the vowel er, for example. The development of signaling these different dimensions of vowel quality would be reflected in the variation in $F1$, $F2$, and $F3$. Therefore, the dispersion (standard deviation) values of the critical band rates of $F1$, $F2$, and $F3$ expressed as a function of the critical band rate of $F3$ (vocal tract length and speaker size) were examined (Traunmüller, 1988), (iii) $Z3-Z0$ is approximately the same in the speech of men, women, and children, but it decreases with increasing vocal effort. Therefore, it can be taken to represent vocal effort while $Z3$ can be taken to represent speaker size, as it decreases with increasing vocal tract length. Therefore, the critical band rate (bark) difference between the third formant and $F0$ $(Z3-Z0)$ was examined as a function of the critical band rate of $F3$; (iv) The range values for $F1$, $F2$, and $F3$ expressed as the difference between the maximum and minimum values of each formant frequency provide information on vowel quality and highlight the role of individual formant frequencies in shaping the vowel quality of peripheral and nonperipheral vowels. For example, peripheral vowels have a greater range of values in $F1$ and $F2$ compared to nonperipheral vowels, whereas a nonperipheral vowel such as er would be expected to show a greater range in $F3$ values due to its rhotacized phonetic quality, which is signaled by lowered values in $F3$ (e.g., Alwan *et al.*, 1997; Dalston, 1975; Espy-Wilson *et al.*, 2000). The range in formant frequencies was therefore examined for $F1$ to $F3$.

Variations in developmental and adult sex differences as a function of both vowel and formant frequency have been reported (e.g., Bennett, 1981; Busby and Plant, 1995; Fant, 1966, 1975). The differences between female and male formant frequency values were therefore examined from an auditory perspective for each vowel as a function of age group using critical band rate values for $F1$ to $F3$ (Traunmüller, 1988).

III. RESULTS

A. *Kn-age* values

The *kn-age* factor values for all ten vowels combined, are depicted in Fig. 1 by age and sex for the formant frequencies $F1$ to $F3$ [Figs. 1(a)–(c)], and for the overall mean of $F1$, $F2$, and $F3$ [Fig. 1(d)]. The general developmental trend for both males and females indicates a decrease in the *k-age* values for all three formant frequencies, as both the males and females between the ages of 7 and 18 years approach the formant frequency values of the adult men and women speakers, respectively. What is worth noting here is that the *k1-age* values for the females are higher than those of the males from age 15 onwards. In addition, the younger females have higher $F1$ values compared to the women speakers.

There are sex-specific differences in both the degree and rate of developmental change. For example, when the *k-age* values were averaged across $k1$, $k2$, and $k3$ ($k_{1,2,3\text{-age}}$), the difference in the degree of change is marked by a shift from a value of 1.34 at age 7 years, to 1.01 at age 18 years, in the males' data [see Fig. 1(d)]. This compares with $k_{1,2,3\text{-age}}$ values of 1.18 and 1.05 for the 7- and 18-year-old females, respectively [see Fig. 1(d)]. The sex-linked differences with respect to the extent of developmental change show that although both the males and females display maturational patterns during puberty (age 10 to 15 years), these are less prominent for the females.

B. *Kn-sex* values

The mean *k-sex* values for all vowels combined are given in Fig. 2(a) by formant frequency and age group. The overall *k-sex* values in Fig. 2(a) depict discernible sex differences before puberty, and the emergence of differences at age 10, where we see small increases in $k2$ and $k3$ factors until age 12. Thereafter, between the ages of 12 and 18 years, the pattern is one of substantial increases with a marked

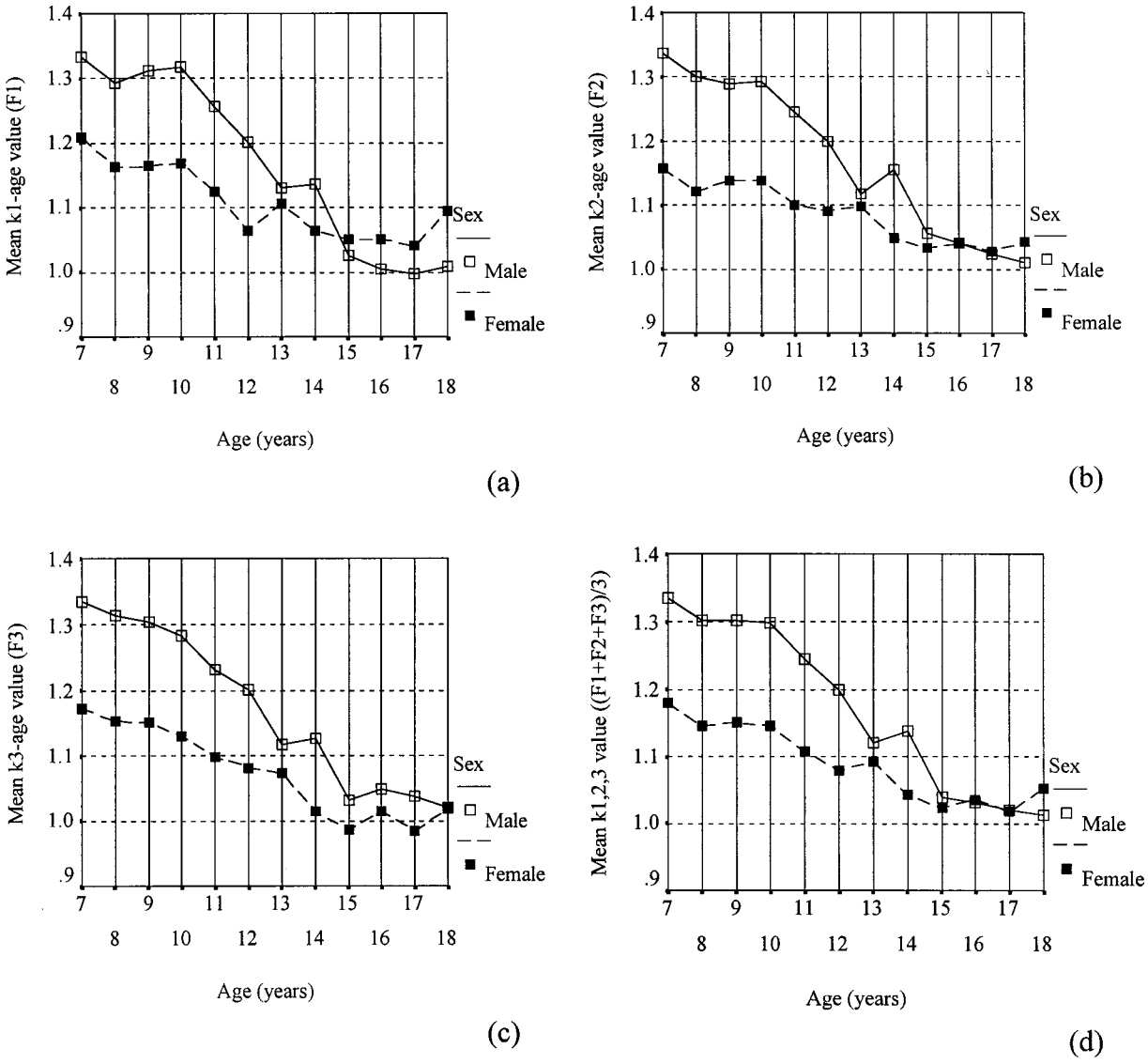


FIG. 1. (a) mean *k1-age* ($F1$) values across all ten vowels by age and sex; (b) mean *k2-age* ($F2$) values across all ten vowels by age and sex; (c) mean *k3-age* ($F3$) values across all ten vowels by age and sex; (d) mean *k*-values averaged for $k1$, $k2$, and $k3$ $[(F1+F2+F3)/3]$ across all ten vowels by age and sex.

decrease (this was a deviation in the general developmental trend) at age 14, which was due to the males displaying an increase in formant frequencies at age 14 (Lee *et al.*, 1999). The *kn-sex* factors depicted in Fig. 2(a) show parallel increases for all three formant frequencies from age 14 to 16 years, with $F1$ having the highest values, and $F3$ the lowest. From age 16 to 17 years there is no change for $k1$ ($F1$) and $k2$ ($F2$), but $k3$ ($F3$) displays a decrease from 1.13 to 1.10. Between age 17 and 18 years marked increases for $k1$ (1.25 to 1.30), $k2$ (1.20 to 1.24), and $k3$ (1.10 to 1.16) are observed. After this point, $k1$ and $k2$ display decreases ($k1$: 1.30 to 1.20, $k2$: 1.24 to 1.20) with no change being observed for $k3$.

When the data for all the children and adults are divided into the four age groups, namely prepuberty (age 7 to 12 years), puberty (13 to 14 years), postpuberty (15 to 18 years), and adults (25 to 50 years), we are able to see the net effect of puberty on sex differences in terms of *k-sex* factors across all three formant frequencies of the ten vowels. This net effect is illustrated in Fig. 2(b), which depicts sex-linked developmental trends. Although there is evidence to suggest that there are sex differences before puberty, these become more marked both during and after puberty. The data given in Fig. 2(b) also suggest that while a substantial degree of sex differences emerges after puberty for some vowels [e.g., ae, ao, iy, and uw in particular—see Fig. 2(b)], others show a lower proportional increase in *kn-sex* values after puberty (e.g., aa, ah, er, and uh). The observation that some vowels display greater sex differences in adulthood [e.g., overall *k-sex* values for aa, ah, eh, and uh—Fig. 2(b)], suggest that some sex-specific patterns continue to unfold during adulthood.

C. *Kn-age* and *kn-sex* values for selected vowels

Figure 3 depicts the mean *k-age* values averaged across all three formant frequencies by age and sex for the selected vowels aa, ah, er, ih, iy, and uw, and Fig. 4, the *k-sex* values by formant frequency [$F1$ ($k1$), $F2$ ($k2$), and $F3$ ($k3$)] for the same group of selected vowels. The developmental trend

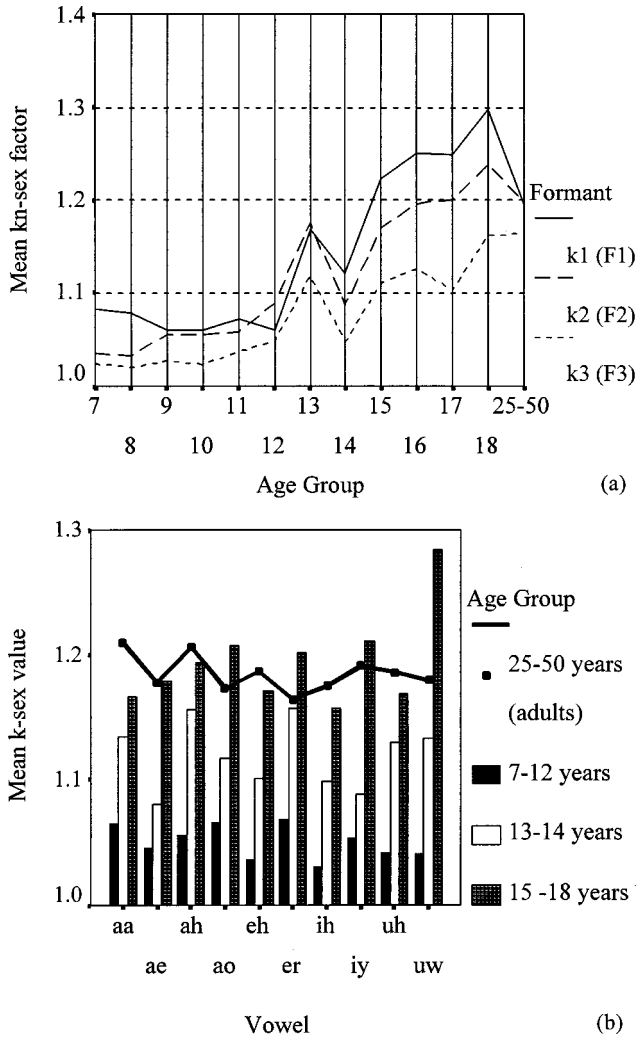


FIG. 2. (a) Mean *kn-sex* values for the first three formant frequencies ($F1$, $F2$, and $F3$) averaged across all vowels by age group. (b) Mean *k-sex* values averaged across all three formant frequencies by vowel for the four age groups (7–12 years, 13–14 years, 15–18 years, and 25–50 years).

in the *k-age* factors across all the six selected vowels is a general decrease in values with age (see Fig. 3), and for the *k-sex* factors a general increase with age (see Fig. 4). However, if we examine the data on a vowel-by-vowel basis from a sex-linked developmental perspective, we are able to note sex differences in the reduction of *k-age* values as a function of both age and vowel. For example, there are general decreases in the males' *k-age* values from age 7 to 13 years, and between age 14 and 15 years for the vowels aa [Fig. 3(a)], ih [Fig. 3(d)], iy [Fig. 3(e)], and uw [Fig. 3(f)], which are more marked when compared to the females data. In addition, after the age of 15, the *k-age* values for males show a negligible decrease, and therefore only slight decreases in formant frequencies during this period. This contrasts with the female values which display variable patterns of increases and decreases in the *k-age* values for different vowels after the age of 15 years (see Fig. 3). For example, the overall *kn-age* values for the females decreases gradually from age 7 and reaches a minimum at age 16 for the vowel aa [see Fig. 3(a)]. After this, the values increase for the vowel aa until age 18. The *k-sex* values for this vowel indicate increases in values after age 14, with the largest of these occurring for $F1$ [see Fig. 4(a)]. In addition, the females' *k-age* values averaged across all three formants for the vowel iy [see Fig. 3(e)] display an increase at age 16, followed by a decrease at age 17, and finally an increase at age 18. The *k-sex* values for iy [see Fig. 4(e)] also suggest as for the vowel aa, increases in values from age 14 years, with again the highest and lowest values being observed for $F1$ and $F3$, respectively.

Further sex-specific differences are exemplified by the vowel uw, which displays the highest *kn-sex* values for the postpuberty group [see Fig. 2(b) and Fig. 4(e)]. This pattern is reflected in the females' *k-age* values and *k-sex* values for the vowel uw [Fig. 4(f)]. If we focus on the male patterns for the *kn-age* values of uw in Fig. 3(f), these are generally similar to those patterns for the other vowels depicted in Figs. 3(a)–(e). That is, with the exception of the increase at age 14, there is a decrease with age towards the adults' values, reaching the adults values at age 17. This pattern is rather different from that of the females, who show the lowest *kn-age* values for uw at age 17, which never approaches 1.0 after this point, but in fact increases markedly [see Fig. 3(f)]. This pattern suggests that the formant frequencies of the 18-year-old females are significantly higher than those values for the adult females, and also explains the high *k-sex* values for uw observed for $k1$ (1.42) and $k2$ (1.55) at age 18, compared to the much lower values for the adults [$k1-1.20$; $k2-1.17$—see Fig. 4(f)]. The net effect of this greater sex difference in the postpuberty group compared to the adult group is depicted in Fig. 2(b).

D. Vowel spaces and critical band rate (bark) distances

The mean distances of all vowels of each sex and age speaker centroid in the $F1-F0/F2-F1$ vowel space are illustrated in Fig. 5(a) by age and sex. As seen in Fig. 5, the mean distances for the males and females both show an overall developmental trend of decreasing mean distances with age. However, upon closer examination of the data, we are able to observe that there are some sex-specific developmental differences in the mean distances from the vowel-group centroid. For example, the mean distances from the vowel-group centroids are similar for both the males and females for the age groups: 7, 10, 16, and 18 years. However, for the 8-year-olds, 17-year-olds, and the adults, the females display larger mean distances from the vowel-group centroids than the males. This contrasts with the data for the 11-, 12-, and 14-year-olds, which display larger mean distances from the vowel centroids for the males. If we examine the mean distances from the speaker centroids in the prepuberty, puberty, postpuberty, and adult groups, sex-specific effects are observed as a function of age group. For example, the males display a gradual decrease in the mean distance from the speaker centroid, whereas the females display a decrease from prepuberty to puberty, but subsequent increases thereafter [Fig. 5(b)].

The distances of each vowel from the speaker centroid for the males and females in the $F1-F0/F2-F1$ vowel space are given in Table I, and depicted in Fig. 6 for the

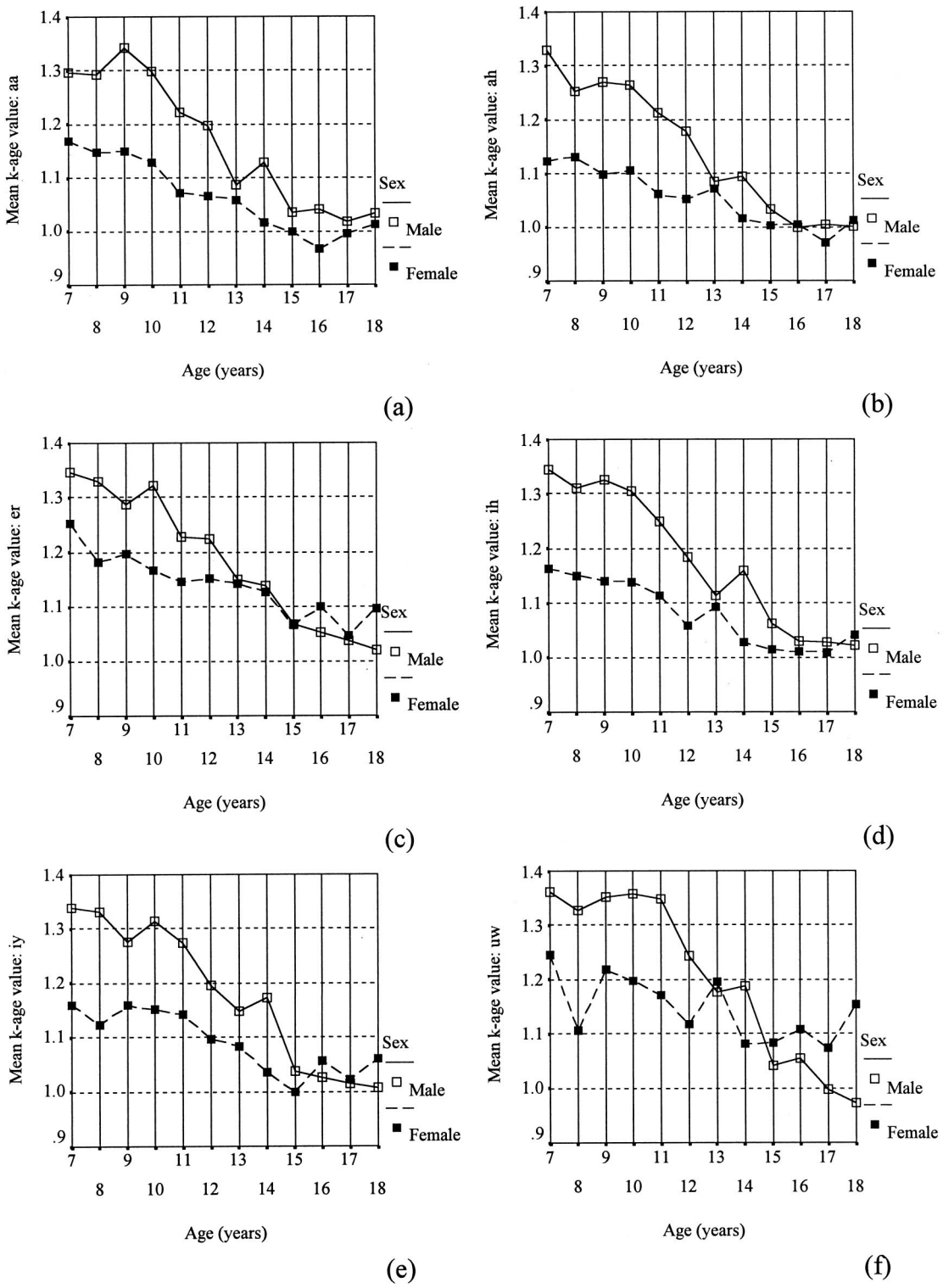


FIG. 3. Mean *k-age* values averaged across three formant frequencies $[(F1+F2+F3)/3]$ by age and sex for the selected vowels (a) aa; (b) ah; (c) er; (d) ih; (e) iy; and (f) uw.

adults (25 to 50 years). What is apparent here is that the women display a more peripheral vowel space than the men. If we define the outer bounds of the monophthongal vowel space using the peripheral vowels iy, uw, ao, ae, and aa, the women occupy a larger acoustic space than the men in the $F2-F1/F1-F0$ dimension. For example, the women display a greater degree openness for the vowel aa, and a greater degree of frontness and constriction for the vowel iy. The female adults also display greater distances from the vowel centroid than their male peers, for nine out of the ten vowels and for all vowels combined, which was found to be significant for the ten vowels using a paired t-test $[t(9)=3.570,\ p<0.01]$ (see Table I).

The results of the vowel openness parameter expressed

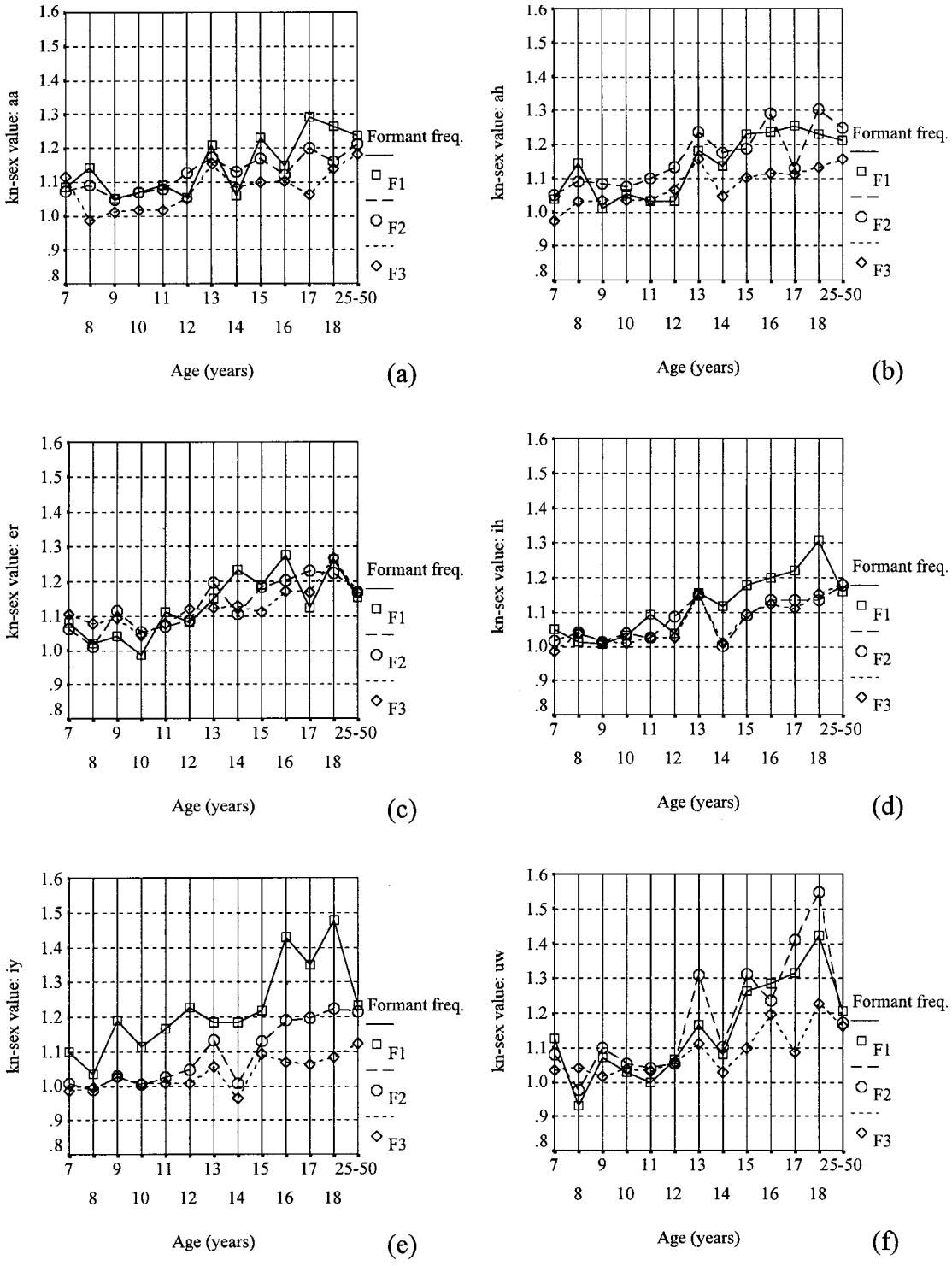


FIG. 4. Mean *k-sex* values by age group and formant frequency ($F1$,$F2$,$F3$), for the selected vowels (a) aa; (b) ah; (c) er; (d) ih; (e) iy; and (f) uw.

as a function of the fundamental frequency (in bark) and the degree of standard deviation (dispersion) in the critical band rate of $F1$ is illustrated as a function of age group and sex in Fig. 7, for the peripheral vowels (left), and the nonperipheral vowels (right). The females display a decrease in $F1$ dispersion between prepuberty and puberty with subsequent increases between postpuberty and adulthood for both vowel sets. This contrasts with the males, who show a decrease in $F1$ dispersion between prepuberty and puberty, with only slight increases for the same age intervals. These results agree with the developmental data reported by Traunmüller (1988). In addition, the peripheral vowel set displays higher dispersion values compared to the nonperipheral vowels.

The degree of variation in vowel formants as a function of age group and sex, expressed as the dispersion (standard deviation) of the critical band rates of $F1$, $F2$, and $F3$ and the critical band rate of $F3$ for both the peripheral and nonperipheral vowels is depicted in Fig. 8. All three formant frequencies display sex-specific developmental differences [Figs. 8(a), (b), and (c)]. For example, both the males and

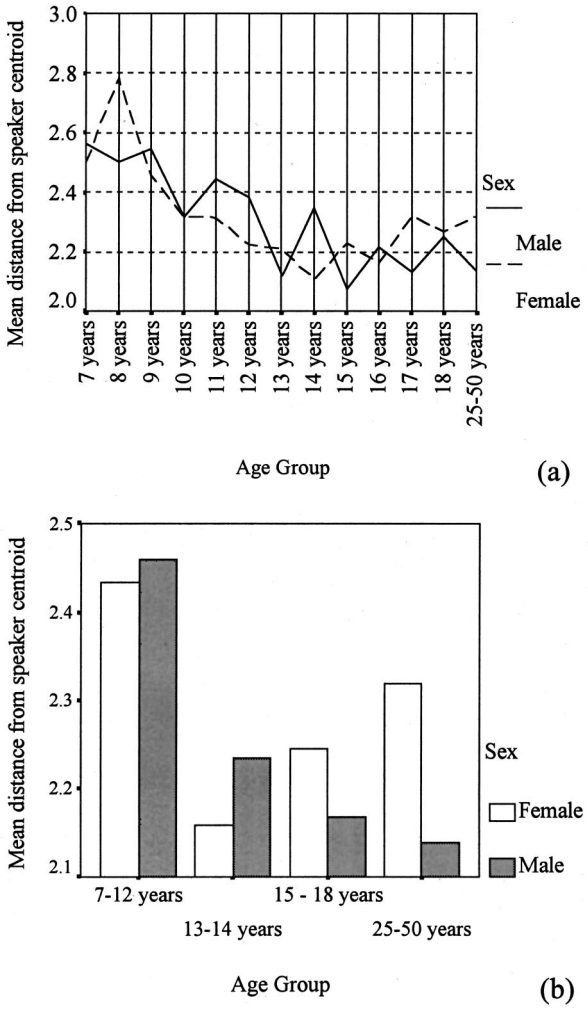


FIG. 5. (a) Mean distance of vowels from the speaker centroid of the male and female speakers by age; (b) Mean distance of vowels from the speaker centroid by sex for the four age groups (7–12 years, 13–14 years, 15–18 years, and 25–50 years).

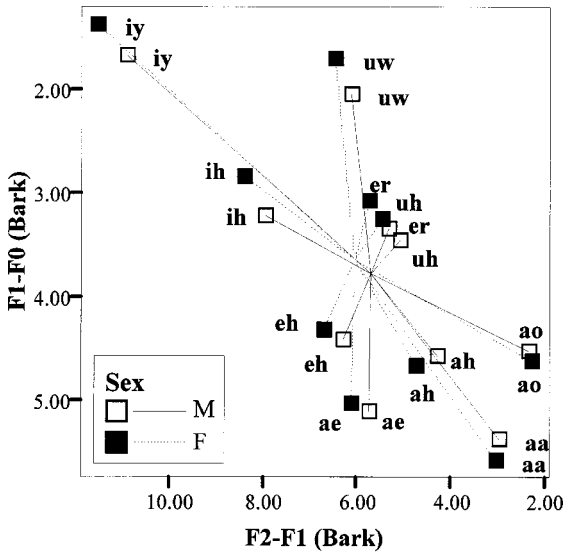


FIG. 6. Vowel spaces [$F1-F0$ (bark) versus $F2-F1$ (bark)] for the men and women adults (25 to 50 years).

females show a decrease in the critical band rate dispersion (standard deviation) from prepuberty to puberty for both $F1$ and $F2$. However, between postpuberty and adulthood, the females exhibit more marked increases in the critical band rate dispersion than the males for both the peripheral and nonperipheral vowels. The women also display higher mean critical-band rate dispersion values for $F1$ compared to the men for both vowel sets [Fig. 8(a)], and the pattern for the peripheral vowels replicates those described by Traunmüller (1988) for a set of peripheral Japanese vowels. The women also display a larger increase in the dispersion of $F2$ values than the men from postpuberty to adulthood for both vowel sets [Fig. 8(b)], and the ontogenetic development of sex differences in the $F2$ dispersion patterns for the peripheral vowels are also similar to the peripheral Japanese vowels reported by Traunmüller (1988). For $F3$ [see Fig. 8(c)] there are also sex-specific developmental differences for both vowel sets, but the pattern of these differences varies with the vowel set in question. For example, the peripheral vowels indicate that both the males and females display slight increases in the critical band-rate dispersion between prepuberty and puberty, and between puberty and postpuberty, and slight decreases between postpuberty and adulthood. This contrasts with the nonperipheral vowels which show more sex-specific differences. For example, the males show a decrease, but the females an increase, between prepuberty and puberty. In addition, although both the females and males show an increase in the dispersion values of $F3$ between postpuberty and adulthood, this is more marked for the females [see Fig. 8(c)]. What is worth noting at this point is that the dispersion values of $F3$ are higher for the nonperipheral vowels compared to the values of the peripheral vowels, due to the inclusion of the rhotacized vowel er in the nonperipheral vowel set.

The relationship between speaker size and vocal effort of the ten vowels expressed as a function of the third formant ($Z3$: bark) and the difference between the third formant and $F0$ ($Z3-Z0$: bark) (after Traunmüller, 1988) is depicted in Fig. 9 by age group and sex for both the peripheral and nonperipheral vowel sets. An increase in vocal tract length is indexed by a decrease in $F3$, and an increase in vocal effort, by a decrease in the value of $Z3-Z0$ (bark). These data

TABLE I. Distances from $F2-F1/F1-F0$ monophthong speaker centroid for the adult group (25 to 50 years) for each vowel by sex as a measure of vowel space. Paired t-test for all ten vowels to test for sex differences (F–M): $t(9)=3.570$, $p<0.01$.

Centroid value for $F2-F1(x)/F1-F0(y)$ by sex	Sex	aa	ae	ah	ao	eh	er	ih	iy	uh	uw	All vowels
F: $x=6.05$, $y=3.65$	F	3.57	1.39	1.66	3.90	0.95	0.64	2.49	5.92	0.69	1.99	2.32
M: $x=5.69$, $y=3.77$	M	3.15	1.32	1.63	3.45	0.87	0.56	2.32	5.62	0.69	1.78	2.14

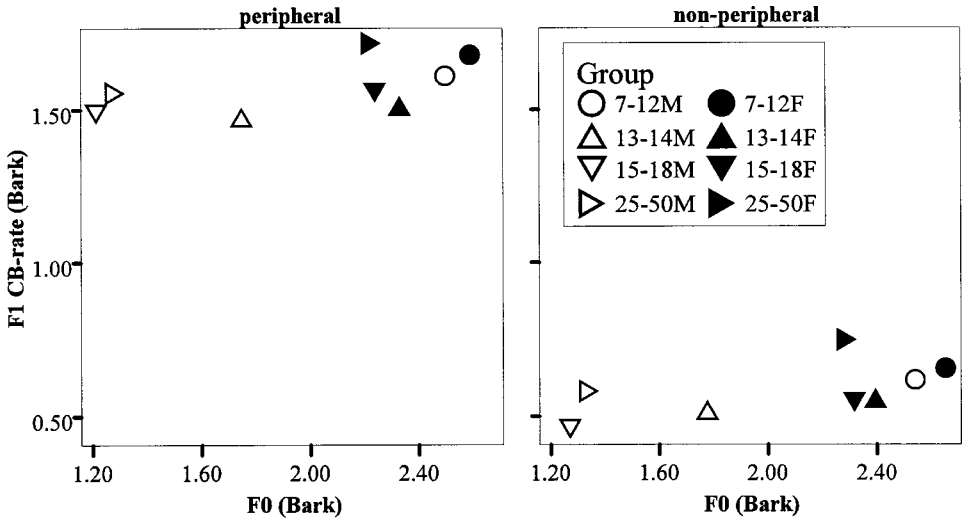


FIG. 7. Critical-band rate standard deviation of $F1$ plotted against $F0$ (bark) for the peripheral vowels (aa, ae, ao, eh, iy, uw) (left), and nonperipheral vowels (ah, er, ih, uh) (right). Values are plotted for the four age groups (7–12 years, 13–14 years, 15–18 years, and 25–50 years) by sex.

illustrate sex-specific developmental differences in both vocal tract length ($F3$: bark) and vocal effort ($Z3-Z0$: bark). For both vowel sets, the females display a pattern of increase in both the dimensions of speaker size and vocal effort from pre- through to postpuberty, the net result of which is more marked for the nonperipheral vowel set that includes er. Although the males also display a net increase in vocal effort from pre- to postpuberty, this developmental increase is less marked than that observed for the females. In addition, between postpuberty and adulthood, the males display higher increases in vocal effort compared to the females for both the peripheral and nonperipheral vowels (see Fig. 9). When the vowel sets are compared, levels of vocal effort for all three age groups are greater for the nonperipheral vowel set that includes er. With respect to vocal tract length, the nonperipheral vowels also display both marked decreases in $F3$ (bark) and therefore, a marked increase in vocal tract length, compared to the peripheral vowels.

The ranges in formant frequency values ($F1$ to $F3$ in bark) across all vowels are illustrated in Fig. 10 for both vowel sets for the four age groups (prepuberty: 7 to 12 years; puberty: 13 and 14 years; postpuberty: 15 to 18 years; adults: 25 to 50 years) by sex. These data show larger range values for both $F1$ and $F2$ for the peripheral vowels. This contrasts with the data for $F3$, which show larger range values for the nonperipheral vowels, a pattern that once again can be explained by the inclusion of er in this vowel set.

The tonotopic distances between female and male formant frequencies ($F1$ to $F3$ in bark) across all vowels are depicted in Fig. 11 for the four age groups (7 to 12 years; 13 and 14 years; 15 to 18 years; 25 to 50 years). These data show that, although there are sex-specific developmental patterns, these are dependent upon both the formant frequency and the vowel. For example, all three formant frequencies display some sex differences before puberty, but some of these differences are greater for specific formants and vowels {e.g., $F1$ of iy [Fig. 11(a)] and $F3$ of er [Fig. 11(b)]}. In addition, although an increase in male–female differences continues to occur from puberty to postpuberty for $F1$, $F2$, and $F3$, these sex differences appear to become less marked between postpuberty and adulthood for some of the data [e.g., $F1$ as shown in Fig. 11(a)].

IV. DISCUSSION

The aim of this paper was to reexamine the fundamental frequency and formant frequency data presented by Lee *et al.* (1999) from a sex-specific developmental perspective. The reexamination of the data adopted two basic approaches; examining sex-linked developmental formant frequency differences as a function of age and sex using formant scaling (Fant, 1966, 1975); and investigating sex-specific developmental patterns in fundamental frequency and formant frequencies using a number of acoustic-phonetic dimensions based on the critical-band rate (Bark) scale (Traunmüller, 1988, 1990). Both approaches revealed a range of sex differences which were developmental in nature. These are discussed below.

A. Formant scaling *k-age* and *k-sex* factors

A number of points emerge from this reexamination of Lee *et al.*'s (1999) data using *k-age* and *k-sex* values based on Fant's method (1966, 1975). The mean *k-age* values for $F1$, $F2$, and $F3$ showed a general decrease with age, with the males displaying greater decreases than the females (see Fig. 1). This trend reflects the general pattern of developmental sex differences in the maturation of the vocal tract (Fitch and Giedd, 1999). In addition, the patterns in Fig. 1 indicate that from age 10, both the males and females display a more marked decrease in formant frequencies, which coincides with the onset of the peripubertal stage at age 10.3 years, identified by Fitch and Giedd (1999). Notably, the decrease in *k-age* values appears to be less marked after age 15 years. This suggests that in the postpubertal stage ($\sim$15 years to $\sim$18 years), overall changes in formant frequencies are of a smaller magnitude than those which occur during puberty.

The *k-age* values also show evidence of sex-specific differences for the formant frequencies of the vowels. For example, the male *k-age* values decrease with age, and by age 18 years, they are close to a scale factor of 1.0 for $F1$ (1.01),

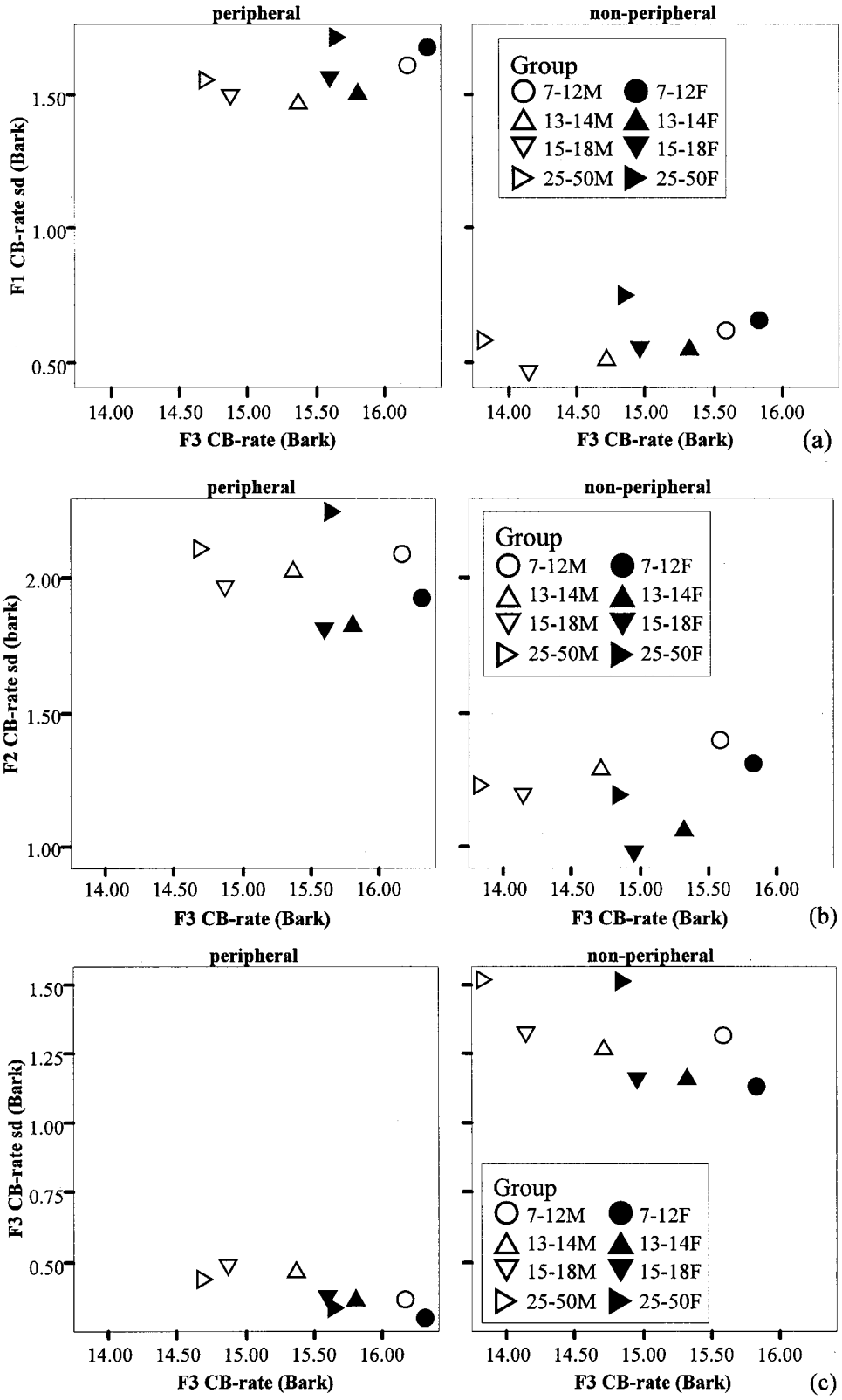


FIG. 8. Critical-band rate standard deviation for (a) first formant frequency ($F1$); (b) second formant frequency ($F2$); (c) third formant frequency ($F3$) plotted against the critical-band rate of $F3$ for the peripheral vowels (aa, ae, ao, eh, iy, uw) (left), and nonperipheral vowels (ah, er, ih, uh) (right). Values are plotted for the four age groups (7–12 years, 13–14 years, 15–18 years, and 25–50 years) by sex.

$F2$ (1.01), and $F3$ (1.02). This contrasts with the data for the females, who also show this decrease with age, but by age 18 years, only $k3$ is close to 1.0 (1.02), whereas $k1$ and $k2$ have values of 1.10 and 1.04, respectively. The *k-age* data for the selected vowels in Fig. 3 suggest that this pattern is the result of the 18-year-old females' higher *k-age* values for er, iy, and uw relative to the adult female group. Such marked differences at age 18 years cannot be explained solely in terms

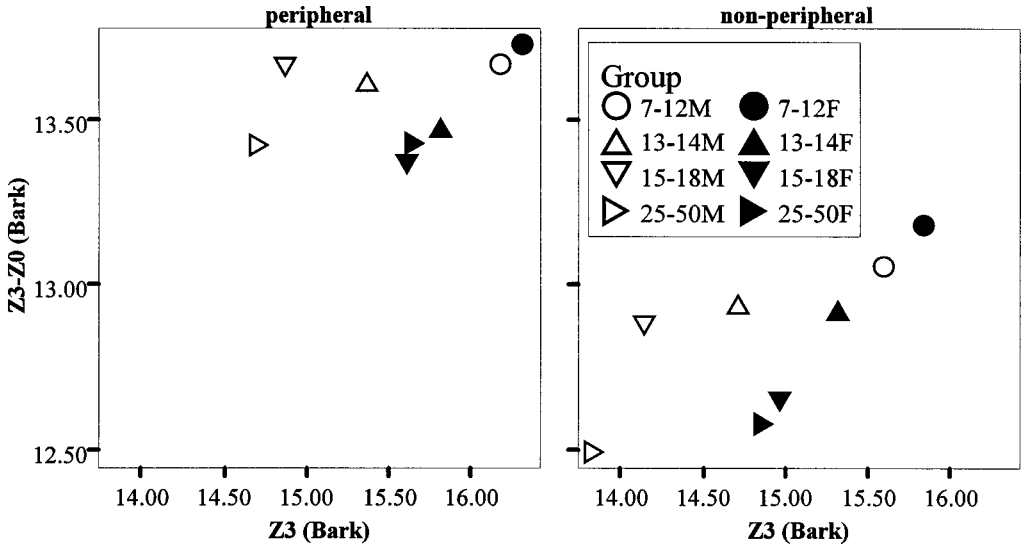
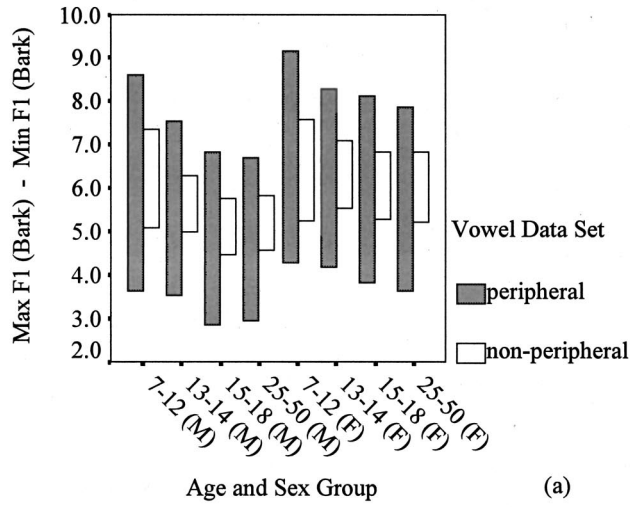


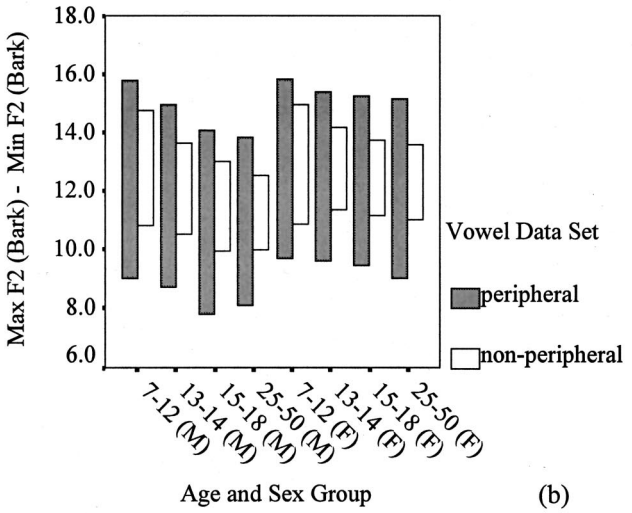
FIG. 9. Critical-band rate (bark) difference between $F3$ and $F0$ plotted against critical-band rate (bark) of $F3$ for the peripheral vowels (aa, ae, ao, eh, iy, uw) (left), and nonperipheral vowels (ah, er, ih, uh) (right). Values are plotted for the four age groups (7–12 years, 13–14 years, 15–18 years, and 25–50 years) by sex.

of maturational differences of the vocal tract alone. It is therefore more likely that the phonetic quality of uw samples produced by the 18-year-olds was different, and on average, more palatalized, less rounded, and more open than those of the adult women. This raises the role of sociophonetic factors as possible key influences in determining the phonetic quality of vowel formant frequencies (e.g., Byrd, 1992, 1994; Henton, 1995; Lee *et al.*, 1999). It is suggested that the phonetic quality of the vowel uw may be an example of sociophonetic and/or accent variation in the 18-year-old females. An additional observation worth noting here is that the 18-year-old females also displayed the highest fundamental frequency values within the postpubertal and adult female groups. The question, therefore, is whether the adults (25 to 50 years) simply came from a different accent group, or whether there were variations in stylistic conventions between the postpubertal and adult groups. Unfortunately, we are not provided with a detailed age, gender, and accent breakdown to ascertain this, and therefore it would be an interesting factor to explore further. Furthermore, because the adult group spans 25 years, it is not unreasonable to suggest that there may be further age- and sex-specific differences within this group. Given this, it would be worth investigating whether age-related sex differences are present between the younger adults (e.g., 25–35 years) and older adults (e.g., 40–50 years). If age-related differences were found, they would supplement evidence reported for changes in speech and voice characteristics during the adult lifespan (e.g., Decoster and Debruyne, 1997, 2000; Xue *et al.*, 1999).

The presence of sex differences before puberty replicates the findings of other studies (e.g., Bennett, 1981; Busby and Plant, 1995; Eguchi and Hirsh, 1969; White, 1999). These sex differences become more marked after puberty [see Fig. 2(b)], despite the drop at age 14 years, which is due to the males displaying higher formant frequencies [see Fig. 2(a)]. Lee *et al.* (1999) explain this drop as being the consequence of maturational processes. The developmental trend of *k-sex* values reflects sex-specific patterns in the maturation of the vocal tract (Fitch and Giedd, 1999), and the *k-sex* values as noted in the Results above (subsections B and C) are both formant frequency dependent [see Figs. 2(a), (b), (c)] and vowel dependent [see Fig. 2(b) and Fig. 4]. These patterns suggest the emergence of nonuniform sex differences in the vocal tract morphology of males and females which therefore affects the degree of variation in formant frequencies as a function of sex and vowel context. The *k-sex* factors for $F1$, $F2$, and $F3$ [see Fig. 2(a)], for example, reflect the emergence of sex differences in the pharynx, oral cavity, and total vocal tract length. This is supported by Fitch and Giedd (1999), who report significant sex differences in the relative length differences between the oral and pharyngeal cavities, with greater mean sex differences in the postpubertal subjects (12.9 mm) compared to those of the peripubertal subjects (7.5 mm). This, therefore, provides some indirect evidence to explain why front vowels such as ae and iy in Lee *et al.*'s data exhibit the bulk of their increases in *kn-sex* values after puberty. The degree of the sex differences exhibited by the vowel uw, in contrast to the other selected vowels, however [see Fig. 2(b)], suggests that vocal tract length alone is not sufficient in explaining the extent of some sex-specific effects, and that other factors related to accent, speaking style, or sociophonetic influences may be responsible for some speaker sex differences. A similar finding is reported by Traunmüller (1988), who observed variations in vowel quality for some women speakers in a Japanese data set.

On the basis of evidence which suggests that the male vocal tract continues to go through maturational changes (Fitch and Giedd, 1999), one might expect to find more dramatic drops in the formant frequencies of the males from age 15 in Lee *et al.*'s (1999) data, and therefore larger *kn-sex* values than those reported here [see Figs. 2(a) and (b)]. A marked lowering of the male formant frequencies after age 15 years does not occur, which is a point that Lee *et al.*





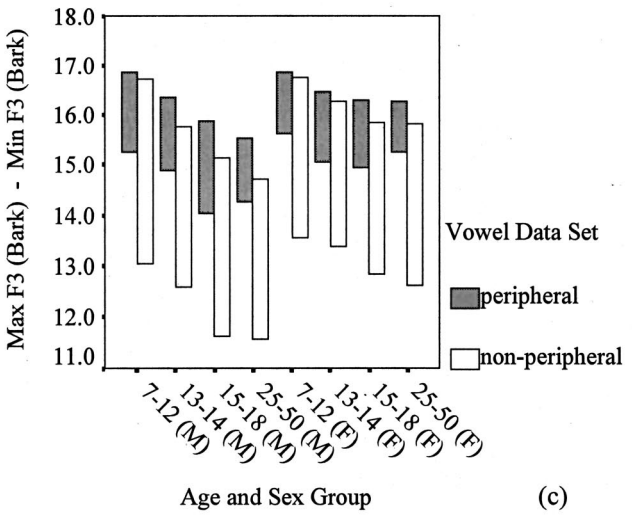


FIG. 10. Ranges in formant frequency values expressed as a function of the difference between the maximum and minimum formant frequency values for (a) $F1$ (bark); (b) $F2$ (bark); and (c) $F3$ (bark). Values are given for the peripheral and nonperipheral vowel sets by age group (7–12 years, 13–14 years, 15–18 years, and 25–50 years), for both females and males.

(1999) raise in their paper. This suggests once again that the physical length of the vocal tract alone, may not be sufficient in explaining the male formant frequency patterns of the ten vowels after age 15 years, and that other factors (e.g., physiological, sociocultural, stylistic conventions) may be responsible for these and other sex-specific patterns (e.g., Byrd, 1992, 1994; Hasek *et al.*, 1980; Henton, 1995; Lee *et al.*, 1999; Mattingly, 1966; Traunmüller, 1984, 1988).

B. Fundamental frequency and formant frequency patterns using the critical-band rate (bark) scale

The examination of fundamental and formant frequency patterns using a selection of acoustic phonetic parameters in the Bark scale reveals the ontogeny of sex differences in the phonetic quality of the ten monophthongs reported by Lee *et al.* (1999). For example, the vowel spaces defined by $F2-F1$ versus $F1-F0$ display evidence of sex-specific patterns, which are developmental in nature. Different sex-specific patterns are observed before, during, and after puberty, and by adulthood, the women display a more peripheral vowel space and therefore a greater phonetic distinctiveness in vowel quality than the men [see Table I, Figs. 5(a) and (b), and 6]. The ontogeny of this sex difference in phonetic distinctiveness is explained if we examine the acoustic-phonetic dimensions of vowel quality that were investigated using the Bark scale. For example, the degree of vowel openness expressed as a function of fundamental frequency (in bark) and the degree of standard deviation in the critical-band rate of $F1$ showed the adult women displaying an increase in the dispersion of $F1$ (bark) for both peripheral and nonperipheral vowels compared to the 15- to 18-year-olds [Figs. 7 and 8(a)]. This increase in the dispersion of $F1$ is not observed for the men speakers and cannot be accounted for as a function of either $F0$ (Fig. 7) or $F3$ [Fig. 8(a)], which therefore suggests that the women in Lee *et al.*'s study must have been producing their vowels with greater acoustic-phonetic distinctiveness, by may be adopting more extreme articulatory postures than the men (e.g., greater openness and closeness). This increased phonetic distinctiveness also explains the higher range values observed for the women's $F1$ values [Fig. 10(a)], and the greater openness of aa depicted in Fig. 6. Similarly, if we examine the phonetic quality of the vowels in terms of the front–back dimension, the adult women display increases in the dispersion values for $F2$ [Fig. 8(b)], and increases in the range values for $F2$ [Fig. 10(b)] compared to the 15- to 18-year-olds. These patterns are particularly marked for the peripheral vowels, and again suggest that the women are displaying a greater degree of phonetic distinctiveness for this group of vowels compared both to the 15- to 18-year-old females and the men within this dimension of vowel quality. The contribution of $F3$ to vowel quality with respect to both dispersion and range values appears to be more significant for the nonperipheral vowels which is largely due to the inclusion of the rhotacized vowel er, which is characterized by lower $F3$ values (Alwan *et al.*, 1997; Dalston, 1975; Espy-Wilson

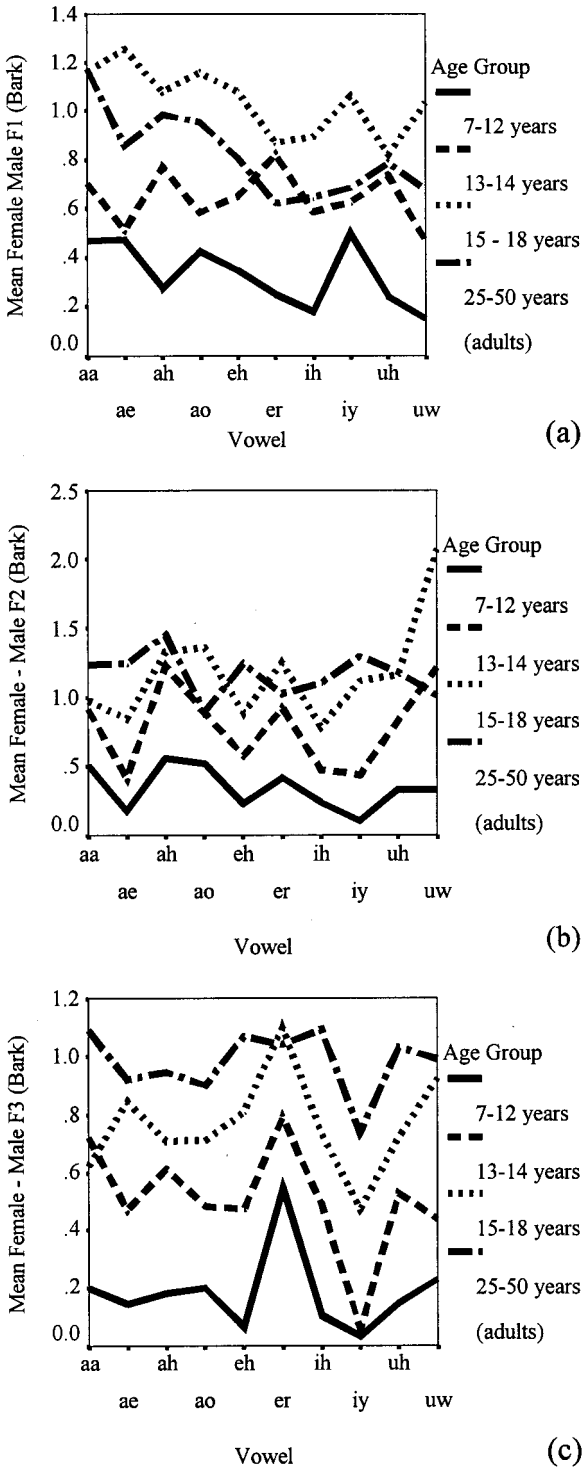


FIG. 11. Critical-band rate (bark) differences between female and male formant frequency values for (a) $F1$; (b) $F2$; and (c) $F3$ by vowel and age group (7–12 years, 13–14 years, 15–18 years, and 25–50 years).

et al., 1997, 2000). What is interesting to note is that although the adult men and women both display increases in the dispersion values of $F3$, the extent of this increase between postpuberty and adulthood is more marked for the adult women [Fig. 8(c)]. This is further evidence to suggest that the adult women are producing vowels which are more distinct in their acoustic-phonetic dimensions. What is interesting to note at this point is that the data for the $F1$, $F2$, and $F3$ dispersion values for the peripheral vowels are similar to those reported by Traunmüller (1988) for a Japanese data set. These cross-linguistic similarities in sex-specific patterns in speech therefore suggest that at least some linguistic behavior in males and females can be explained by underlying physiological differences between the sexes.

The data for $F1$ and $F2$ discussed above go some way in explaining both the greater distances from the speaker centroid, and the more peripheral vowel spaces for the women speakers in terms of the open–close acoustic-phonetic dimension of vowels (Table I, Fig. 6). The pattern of larger and more peripheral vowel spaces for women replicates previous findings (e.g., Deterding, 1997; Henton, 1983, 1995; Rosner and Pickering, 1994; Traunmüller, 1988), but more importantly, the reexamination of Lee *et al.*'s (1999) data reveals that sex differences in vowel space appear to emerge with development, from preadolescence to adulthood. Physiological factors and anatomical constraints due to sex-specific maturational differences may be instrumental in shaping the more peripheral vowel spaces displayed by women (Traunmüller, 1984, 1988). However, sociophonetic factors may also be playing a part in influencing the development of learned sex-specific speech behaviors. This suggestion is supported by evidence which suggests that the auditory space of men and women varies across languages, and that there are therefore language-specific stylistic factors that may determine some habitual speech settings [see Henton (1995) and Rosner and Pickering (1994) for examples of data from a variety of languages]. The extent to which sociophonetic influences appear to be culturally determined should therefore be acknowledged in light of these cross-language data.

An issue related to sociophonetic factors and stylistic convention is the extent to which speech behavior(s) are shaped by a particular scenario. For example, a cross-language study of American, Russian, and Swedish (Kuhl *et al.*, 1997) found evidence of mothers producing ''more extreme'' vowels in infant-directed samples, compared to those in adult-directed samples. In addition, studies by Byrd (1992, 1994) report evidence to suggest that women may adopt a speech style that displays less phonetic reduction in more formal contexts, such as experimental settings. Therefore, the extent to which the experimental scenario influenced the speech style of the women in Lee *et al.*'s study (1999), to produce their more peripheral vowel spaces, remains both an important and interesting question. What emerges from this discussion, however, is that a combination of anatomical, physiological, sociophonetic/cultural, and idiosyncratic factors are all likely to play some role in determining speaker sex differences, and their developmental patterns.

The relationship between vocal tract length and vocal effort of the ten vowels expressed as a function of the third formant ($Z3$: bark) and the difference between the third formant and $F0$ ($Z3-Z0$: bark) show sex-specific developmental patterns for both vowel sets. Although both the females and males display increases in the vocal effort parameter with age for both vowel sets as decreases in $Z3-Z0$, the males display more significant increases in vocal effort from

postpuberty to adulthood, which is largely due to their lower $F3$ values compared to the females. In addition, the increases observed for vocal effort were found to be more marked for the nonperipheral vowels, which can be explained by the inclusion of the rhotacized vowel er in this vowel set which has lowered $F3$ values. The dimension of speaker size indexed as an inverse relationship to $F3$ shows that the development of rhoticity, and lower $F3$ values, is therefore instrumental in contributing to decreases in $Z3-Z0$, and therefore increases in ''speaker size'' [see Fig. 9 (right side)]. Of particular note is the markedly lower $F3$ values for the adult men (13.84 bark) compared to the adult women (14.87 bark) for the nonperipheral vowel set that includes er. The distinctive ''dip'' or lowered $F3$ values for the rhotacized vowel is likely to be the result of similar articulatory configurations that are typically reported for /ɹ/ (e.g., Alwan *et al.*, 1997; Espy-Wilson *et al.*, 2000) and replicates previously reported speaker sex differences in adults for the formant frequency values of /ɹ/ (Westbury *et al.*, 1998). The articulatory configurations for /ɹ/ include pharyngeal, palatal, and labial constrictions (e.g., Alwan *et al.*, 1997; Espy-Wilson *et al.*, 1997, 2000), and the creation of a sublingual cavity anterior between the palatal and lip constriction (e.g., Espy-Wilson and Boyce, 1999). The net effect of this sublingual cavity is to increase the volume of the oral cavity, which therefore lowers the frequency of $F3$. On this basis, the presence of a sex difference before puberty, for the mean value of $F3$ (bark) for the vowel er (see Fig. 11) suggests that there may already be sex differences before puberty in the volume of the oral cavity which includes the sublingual cavity, which continue to increase during and after puberty. In addition, the presence of speaker sex differences in the lengths of the lip segments of prepubertal boys and girls (Fitch and Giedd, 1999) may also help to explain the presence of this sex difference before puberty.

The sex differences in the MRI data of vocal tract morphology reported by Fitch and Giedd (1999) also highlight the sex differences in vocal tract length and the proportionately longer pharynx of males after puberty. Nonuniform sex differences in the vocal tract are capable of explaining the nonlinear increase in the tonotopic distance between female and male formant frequency values of different vowels. For example, greater female–male tonotopic distances of $F1$ for aa, ae, ao, and ah after puberty [see Fig. 11(a)] could be attributed to marked growth in the pharyngeal cavities of postpuberty males, and a similar explanation could be proposed for the marked increase in female–male differences for $F2$ from puberty to postpuberty for the palatal vowel iy. The fronted quality of uw (produced by the 17- and 18-year-olds), which was suggested to be a consequence of accent/sociophonetic factors, may also be explicable in these terms. However, in addition to nonuniform sex differences in vocal tract length, sex differences in vocal tract volume also require some investigation. Sex-specific differences in vocal tract volume may provide us with additional information on the emergence of some of the more marked sex differences in specific formants of specific vowels, such as $F3$ in er [see Fig. 11(c)]. What is worth highlighting at this stage is that even before puberty, there is an appreciable tonotopic distance between the $F3$ values of females and males [Fig. 11(c)], an observation which further stresses that vocal tract length alone may not be responsible for all the sex differences observed for vowel formant frequencies. The fact that $F1$ shows a decrease in the frequency differences between the adult men and women compared to the postpubertal (15–18 years) males and females [see Fig. 11(a)] is due to decreases in $F1$ values for the adult women, and suggests that there may be age-specific changes occurring in the vocal tracts of the older females in this group. This suggestion is speculative but is not unreasonable in the light of recent evidence which shows that there are differences between the vocal tract configurations and resulting formant frequencies of 33–48-year-old and 50–66-year-old women (Xue *et al.*, 1999). This further highlights the need for further information on the demographic profile of the adult men and women reported in Lee *et al.* (1999).

Further research into human vocal tract morphology and acoustic correlates, together with a longitudinal perspective of speech development, will provide valuable insights into age- and sex-specific developmental formant frequency patterns. Such a longitudinal perspective would also assist in shedding further light on specific aspects of individual differences in the development and maturational process of speech characteristics (Smith and Kenney, 1998). In addition, a longitudinal perspective may go some way in explaining some of the between-subject variability observed in cross-sectional studies, and the instability of speech patterns during periods of accelerated maturation and growth around and during puberty (Lee *et al.*, 1999), and highlight those changes that may be occurring across the adult lifespan (e.g., Decoster and Debruyne, 1997, 2000; Xue *et al.*, 1999). Such a perspective may also reveal the extent to which physiological, sociophonetic, cultural, and stylistic conventions are responsible for sex differences in fundamental and formant frequencies and the acoustic-phonetic quality of vowels which cannot be explained by developmental sex-specific differences in vocal tract morphology alone (e.g., Busby and Plant, 1995; Fant, 1966, 1975; Fitch and Giedd, 1999; Lee *et al.*, 1999; Traunmüller, 1988, 1990).

ACKNOWLEDGMENTS

I would like to thank Professor Anders Löfqvist, Professor Hartmut Traunmüller, and two anonymous reviewers for their thorough and constructive comments.

[1]The data from the 5- and 6-year-olds were excluded from this reexamination, as the vowel data were based on words produced in isolation. This contrasted with the source of other age groups, where vowel data were based on words spoken in carrier phrases. The exclusion of these data therefore meant that data from 13 age groups were included in this study [7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18-year-olds, and the adults (25 to 50 years)].

[2]These alphabetic symbols are also used to represent the monophthongs (vowels) in all tables and figures; however, the equivalent IPA symbols for reference to readers are as follows: aa (/ɑ/ as in *pot*); ae (/æ/ as in *bat*); ah (/ʌ/ as in *but*); ao (/ɔ/ as in *ball*); eh (/ɛ/ as in *bet*); er (/ɝ/ as in *bird*); ih (/ɪ/ as in *bit*); iy (/iː/ as in *bead*); uh (/ʊ/ as in *put*); uw (/uː/ as in *boot*). The term *vowel* will be used to complement the term ''monophthong'' from this point onwards.

Alwan, A., Narayanan, S., and Haker, K. (**1997**). ''Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. II. The rhotics,'' J. Acoust. Soc. Am. **101**, 1078–1089.
Bennett, S. (**1981**). ''Vowel formant frequency characteristics of preadolescent males and females,'' J. Acoust. Soc. Am. **69**, 231–238.
Busby, P. A., and Plant, G. L. (**1995**). ''Formant frequency values of vowels produced by preadolescent boys and girls,'' J. Acoust. Soc. Am. **97**, 2603–2606.
Byrd, D. (**1992**). ''Preliminary results on speaker-dependent variation in the TIMIT database,'' J. Acoust. Soc. Am. **92**, 593–596.
Byrd, D. (**1994**). ''Relations of sex and dialect to reduction,'' Speech Commun. **15**, 39–54.
Childers, D. G., and Wu, K. (**1991**). ''Gender recognition from speech. II. Fine analysis,'' J. Acoust. Soc. Am. **90**, 1841–1856.
Dalston, R. M. (**1975**). ''Acoustic characteristics of English /w, r, l/ spoken correctly by young children and adults,'' J. Acoust. Soc. Am. **57**, 462–469.
Decoster, W., and Debruyne, F. (**1997**). ''Changes in spectral measures and voice-onset time with age: A cross-sectional and a longitudinal study,'' Folia Phoniatr. Logop. **49**, 269–280.
Decoster, W., and Debruyne, F. (**2000**). ''Longitudinal voice changes: Facts and interpretation,'' J. Voice **14**, 184–193.
Deterding, D. (**1997**). ''The formants of monophthong vowels in standard southern British English pronunciation,'' J. Int. Phonetic Assoc. **27**, 47–55.
Eguchi, S., and Hirsh, I. (**1969**). ''Development of speech sounds in children,'' Acta Oto-Laryngol., Suppl. **257**, 1–51.
Espy-Wilson, C. Y., and Boyce, S. E. (**1999**). ''A simple tube model for American English /r/,'' Proceedings of the 14th International Congress of Phonetic Sciences, San Francisco, pp. 2137–2140.
Espy-Wilson, C. Y., Boyce, S. E., Jackson, M., Narayanan, S., and Alwan, A. (**2000**). ''Acoustic modeling of American English /r/,'' J. Acoust. Soc. Am. **108**, 343–356.
Espy-Wilson, C. Y., Narayanan, S., Boyce, S. E, and Alwan, A. (**1997**). ''Acoustic modeling of American English /r/,'' Proceedings of the 5th European Conference on Speech Communication and Technology (Rhodes, Greece), pp. 393–396.
Fant, G. (**1975**). ''Non-uniform vowel normalization,'' STL-QPSR, KTH-Stockholm 2-3/1975, 1–19.
Fant, G. (**1966**). ''A note on vocal tract size factors and non-uniform F-pattern scalings,'' STL-QPSR, KTH-Stockholm 4/1966 22-30.
Fitch, W. T., and Giedd, J. (**1999**). ''Morphology and development of the human vocal tract: A study using magnetic resonance imaging,'' J. Acoust. Soc. Am. **106**, 1511–1522.
Hasek, C. S., Singh, S., and Murry, T. (**1980**). ''Acoustic attributes of preadolescent voices,'' J. Acoust. Soc. Am. **68**, 1262–1265.
Henton, C. (**1983**). ''Changes in the vowels of received pronunciation,'' J. Phonetics **11**, 353–371.
Henton, C. (**1995**). ''Cross-language variation in the vowels of female and male speakers,'' Proceedings of the 13th Congress of Phonetic Sciences (Stockholm, Sweden), 4, 420–423.
Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., and Lacerda, F. (**1997**). ''Cross-language analysis of phonetic units in language addressed to infants,'' Science **277**, 684–686.
Ladefoged, P., and Maddieson, I. (**1990**). ''Vowels of the world's languages,'' J. Phonetics **18**, 93–122.
Lee, S., Potamianos, A., and Narayanan, S. (**1999**). ''Acoustics of children's speech: Developmental changes of temporal and spectral parameters,'' J. Acoust. Soc. Am. **105**, 1455–1468.
Mattingly, I. (**1966**). ''Speaker variation and vocal tract size,'' J. Acoust. Soc. Am. **39**, 1219.
Peterson, G. E., and Barney, H. L. (**1952**). ''Control methods used in the study of vowels,'' J. Acoust. Soc. Am. **24**, 175–184.
Rosner, B. S., and Pickering, J. B. (**1994**). *Vowel Perception and Production* (Oxford University Press, Oxford, UK).
Smith, B. L., and Kenney, M. K. (**1998**). ''An assessment of several acoustic parameters in children's speech production development: longitudinal data,'' J. Phonetics **26**, 95–108.
Traunmüller, H. (**1981**). ''Perceptual dimension of openness in vowels,'' J. Acoust. Soc. Am. **69**, 1465–1475.
Traunmüller, H. (**1984**). ''Articulatory and perceptual factors controlling the age- and sex-conditioned variability in formant frequencies of vowels,'' Speech Commun. **3**, 49–61.
Traunmüller, H. (**1988**). ''Paralinguistic variation and invariance in the characteristic frequencies of vowels,'' Phonetica **45**, 1–29.
Traunmüller, H. (**1990**). ''Analytical expressions for the tonotopic sensory scale,'' J. Acoust. Soc. Am. **88**, 97–100.
Westbury, J. R., Hashi, M., and Lindstrom, M. J. (**1998**). ''Differences among speakers in lingual articulation for American English /ɹ/,'' Speech Commun. **26**, 203–226.
White, P. (**1999**). ''Formant frequency analysis of children's spoken and sung vowel using sweeping fundamental frequency production,'' J. Voice **13**, 570–582.
Whiteside, S. P., and Hodgson, C. (**2000**). ''Speech patterns of children and adults elicited via a picture-naming task: An acoustic study,'' Speech Commun. **32**, 267–285.
Wu, K., and Childers, D. G. (**1991**). ''Gender recognition from speech. I. Coarse analysis,'' J. Acoust. Soc. Am. **90**, 1828–1840.
Xue, A., Jiang, J., Lin, E., Glassenberg, R., and Mueller, P. B. (**1999**). ''Age-related changes in human vocal tract configurations and the effects on speakers' vowel formant frequencies: A pilot study,'' Log. Phon. Vocol. **24**, 132–137.
Zwicker, E., and Terhardt, E. (**1980**). ''Analytical expressions for critical-band rate and critical bandwidth as a function of frequency,'' J. Acoust. Soc. Am. **68**, 1523–1525.

SIM—simultaneous inverse filtering and matching of a glottal flow model for acoustic speech signals

Matthias Fröhlich, Dirk Michaelis, and Hans Werner Strube
Drittes Physikalisches Institut, Universität Göttingen, Bürgerstr. 42-44, D-37073 Göttingen, Germany

(Received 30 October 2000; accepted for publication 12 April 2001)

A new method ''simultaneous *i*nverse filtering and *m*odel matching'' (SIM) is proposed that allows one to calculate voice source measures without any user interaction. It is based on the discrete all-pole modeling (DAP) technique for inverse filtering (IF), which is modified to include a model of the glottal flow as integral part [LF model, Fant *et al.*, STL-QPSR (Stockholm) **4/1985**, 1–13 (1986)]. As the correct LF parameters are initially unknown, they are estimated in an iterative procedure using multi-dimensional optimization techniques that are initialized according to the results of an exhaustive search. The error criteria applied reflect how well the IF is performed after the spectral contribution of the glottal flow has been removed. The resulting optimal LF parameter constellation serves as the basis to calculate 11 voice source measures. The performance was evaluated using synthesized signals and recordings of natural utterances. For the synthesized signals, the accuracy to reproduce the original parameters was high (correlations exceeding 0.88) for measures where the starting point of the glottal cycle did not enter explicitly. Errors were smaller compared to conventional estimation methods where the measures were estimated from the IF signal. The analysis of natural utterances indicates that problems still exist with regard to robustness, but that under advantageous conditions the open quotient, the speed quotient, the closing quotient, the parabolic spectral parameter, and the negative peak amplitude of the glottal flow derivative can indeed be determined automatically by the SIM method. [DOI: 10.1121/1.1379076]

PACS numbers: 43.70.Jt, 43.70.Gr, 43.72.Ct, 43.72.Ar [AL]

I. INTRODUCTION

The glottal flow is of fundamental importance in voice generation. In many applications—ranging from speech synthesis to phoniatric examinations—a detailed knowledge of the glottal flow signal is desirable. However, its direct assessment requires invasive measures that are only acceptable within specific scientific setups. In general practice, the glottal flow is estimated from a signal that is recorded noninvasively. Characteristic measures are derived from this estimate to quantitatively describe the vocal function.

A major problem in the extraction of glottal features is the noise present in the estimated glottal flow signal. Some researchers have dealt with this problem by determining the glottal features interactively (Alku and Vilkman, 1994, 1996; Hertegård and Gauffin, 1995; Södersten *et al.*, 1999). While this allows the analysis even of highly noisy signals, such an analysis procedure requires trained experts. Furthermore, it is labor intensive which limits the amount of data that can be analyzed in a reasonable time. Therefore, those methods are difficult to apply in examinations that are performed routinely. Other researchers have extracted source features in a completely automatic way (Childers and Lee, 1991; Södersten *et al.*, 1999), but the resulting measures were sometimes difficult to interpret in terms of the voice generation process (Schoentgen, 1982).

Generally, measures determined directly from the estimated glottal flow signal (e.g., open quotient, speed quotient) are sensitive to noise if they are based on certain time instances or thresholds within the cycle. In order to reduce this sensitivity, a parametric model may be fitted to the signal, so that the model parameters can then be used to characterize the signal. This approach reduces the noise sensitivity of the extracted features because the model fit is usually based on all samples of the the cycle (Milenkovic, 1986; Strik *et al.*, 1993). Major issues in this approach are the design of the fitting procedure and the choice of the glottal flow model. The particular properties that should be captured by the model depend on the specific task, which explains the number of different models described in the literature (e.g., Fant *et al.*, 1986; Milenkovic, 1986; Rosenberg, 1970; Titze, 1989).

The above considerations concern the parametrization of a given signal that represents the glottal flow. However, this signal has to be obtained by processing the recorded raw signal—a process with its own problems. The raw signal is usually either the oral flow—recorded by means of a specially designed flow mask (Rothenberg, 1973)—or the acoustic speech pressure waveform recorded in the free field. In the first case, the frequency response of the mask may complicate the interpretation of the results (Shadle *et al.*, 1999). In the latter case the signal is recorded by a microphone so that the acoustic conditions of the recording setup have to be controlled (frequency response of the microphone and of the other equipment, background noise, etc.). If the acoustic signal is used as input, only the ac component of the glottal flow can be estimated, whereas both ac and dc components can be assessed with the flow mask.

Researchers have basically followed two principal approaches to estimate the glottal flow. In the first, antiresonances are adjusted interactively, which has often been per-

formed for signals recorded with a flow mask (Gobl and Chasaide, 1992; Fant, 1993; Holmberg *et al.*, 1995; Hertegård and Gauffin, 1995; Sulter and Wit, 1996; Södersten *et al.*, 1999). As in the determination of glottal measures described earlier, the main drawbacks of this procedure are the labor intensiveness and the subjectivity of the results. In the second approach, the estimation of the antiresonance filters is performed automatically by applying optimization algorithms (Kasuya *et al.*, 1999; Milenkovic, 1986; Strik *et al.*, 1993).

The method that is proposed in this article—''*s*imultaneous *i*nverse filtering and *m*odel matching,'' referred to as the SIM method—belongs to this second category. The goal is to obtain characteristic features of the estimated glottal flow in a way that does not require any user interaction. The method is based on the source-filter theory (Titze, 1994), which describes the voicing process by a linear system of filters that is excited by a source signal. Although several limitations of such a linear description have been discussed in the literature (Fant, 1993; Kröger, 1991; Michaelis, 2000; Shadle *et al.*, 1999), it has been found appropriate for many applications (Alku and Vilkman, 1994; Childers and Lee, 1991; Childers and Ahn, 1995; Fant, 1993, 1995; Gobl and Chasaide, 1992; Ma *et al.*, 1994; Strik *et al.*, 1993; Titze, 1994).

According to the source-filter theory, the speech signal s is obtained by the convolution (denoted by $*$) of the impulse responses of the train of glottal flow pulses γ, the vocal tract resonance filter v, and the radiation into the free field w (Markel and Gray, 1976). The train of glottal flow pulses itself results from the convolution of a delta pulse train d and the impulse response of the glottal flow cycle g. The vocal tract contribution v is modeled by an all-pole filter. The radiation part w is usually approximated as differentiation in the time domain, but it can also be modeled to a higher complexity (Wakita and Fant, 1978). In the simple case where the radiation is modeled by differentiation, usually the glottal flow signal is differentiated: $e=dg/dt=g*w$. The time signal s is thus given by

$$s(t)=\gamma(t)*v(t)*w(t)=d(t)*g(t)*v(t)*w(t)$$
$$=d(t)*e(t)*v(t). \quad (1)$$

Equivalently, this can be expressed in the frequency domain by the product of the corresponding complex transfer functions (denoted by capital symbols):

$$S(\omega)=\Gamma(\omega)\cdot V(\omega)\cdot W(\omega)=D(\omega)\cdot G(\omega)\cdot V(\omega)\cdot W(\omega)$$
$$=D(\omega)\cdot E(\omega)\cdot V(\omega). \quad (2)$$

In order to obtain E (and from E the glottal flow signal g), V has to be estimated as accurately as possible. In the literature, this has often been performed by applying linear prediction (LP, Markel and Gray, 1976) techniques to the part of the signal that corresponds to the closed phase of the glottal cycle. The determination of the closed phase without the use of other, nonacoustic signals is a complex issue of its own right that is discussed elsewhere (Childers and Ahn, 1995; Ma *et al.*, 1994; Strube, 1974).

A different approach is followed for the SIM method that is similar to the one described by Alku and Vilkman (1994). From Eq. (2) it can be seen that a filter A^{-1} will yield D when applied to the signal $S'=S/GW=S/E$:

$$A^{-1}\frac{S}{E}=A^{-1}S'=D\Rightarrow A^{-1}=1/V. \quad (3)$$

Different methods are described in the literature to estimate the coefficients of such a filter A^{-1} that is the inverse of V. One method that is well suited for analyzing digitized voice data is the ''discrete all-pole modeling'' algorithm (DAP, El-Jaroudi and Makhoul, 1991). DAP iteratively optimizes the results obtained by ''conventional'' LP and provides a more accurate estimate of the filter coefficients than LP (see Sec. II B).

The major problem in applying Eq. (3) is how to arrive at a good estimate of E, because E is not known before the inverse filtering (IF) has been performed. However, if E is incorrect when setting $S'=S/E$, then the subsequent IF will not produce optimal results. By quantifying the success of the IF, a measure of the appropriateness of E can therefore be defined. Using such a measure, the SIM method applies an iterative procedure to obtain the optimal description of E. In this respect the method differs from similar approaches to remove the glottal flow shape prior to the inverse filtering (Alku and Vilkman, 1994), or to estimate the glottal flow parameters and the vocal tract filter within an ARX (autoregressive model with exogenous input) framework (Kasuya *et al.*, 1999).

In the SIM method, $e(t)$ is described by the LF model (Fant *et al.*, 1986), which constitutes a standard model of the glottal flow derivative with widespread application (Childers and Lee, 1991; Childers and Ahn, 1995; Fant, 1993, 1995; Gobl and Chasaide, 1992; Ma *et al.*, 1994; Strik *et al.*, 1993). Different LF parameter constellations are tested for their effect on the outcome of the IF. The particular constellations are chosen on the basis of a sequence of multidimensional optimization procedures. The LF constellation that leads to the best IF result is taken as optimal description of the glottal flow. Measures describing the glottal flow are therefore calculated from the LF model rather than from the IF signal in order to avoid the problems described in the initial part of this section.

In the first part of the article, the new method is described. After a brief review of the LF model (Sec. II A), the inclusion of the glottal flow model into the IF routine is described (Sec. II B). The multi-dimensional optimization of the model parameters is explained (Sec. II C), after which the measures are described that are calculated from the fitted model to characterize the estimated flow (Sec. II D).

The SIM method is tested using synthesized speech and natural utterances. For synthesized signals, the performance is assessed by testing to which accuracy the original parameter settings can be retrieved. Natural utterances are analyzed to find out whether the measures extracted from the matched LF model allow a consistent interpretation with regard to different phonation types. In this way the reliability of the results is evaluated. Those test results are presented in the second part of the article. After the description of the data

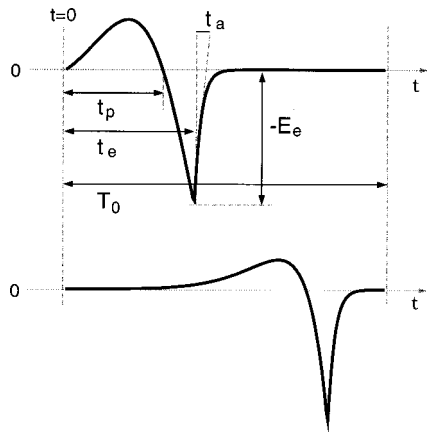


FIG. 1. Derivative of the glottal flow, parametrized by the LF model. Top: typical pulse illustrating the parameters used. Note that the negative peak value is specified by $-E_e$ so that $E_e>0$. Bottom: LF pulse synthesized with uncommonly high values of t_p and t_e ($t_p=0.75$, $t_e=0.82$). The glottal flow derivative therefore increases very gradually at the beginning of the cycle.

(Sec. III A), results are presented and discussed for synthetic speech (Sec. III B) and for natural utterances (Sec. III C). A general discussion follows in Sec. IV.

II. METHODS

A. LF model

The LF model (Fant *et al.*, 1986) is a four-parameter model of the glottal flow derivative $e(t)$. The time-normalized glottal cycle is modeled in two sections: an exponentially weighted sinusoid models the open phase until first collisional contact of the vocal folds, followed by an exponential return phase that prohibits an unrealistic abrupt termination of the flow:[1]

$$e(t)=\frac{dg(t)}{dt}=\begin{cases} E_0 e^{\alpha t}\sin\omega_g t, & 0\leqslant t\leqslant t_e, \\ -\dfrac{E_e}{\varepsilon t_a}[e^{-\varepsilon(t-t_e)}-e^{-\varepsilon(1-t_e)}], & t_e<t<1. \end{cases} \tag{4}$$

The continuity constraints between the two sections and the boundary condition $\int_0^1 e(t)dt=0$ lead to a reduction of the number of independent parameters. The time markers $t_p=\pi/\omega_g$, t_e, t_a together with E_e are used as set of independent parameters in this study. These parameters are illustrated in Fig. 1.

B. Inclusion of the glottal flow model into the IF procedure

For many years, the automatic estimation of the IF coefficients has been performed on the basis of linear prediction (LP, Markel and Gray, 1976). However, the ''discrete all-pole modeling'' method (DAP, El-Jaroudi and Makhoul, 1991) constitutes an improvement over standard LP approaches by taking into account the aliasing effect that inevitably occurs when the power spectrum P of the input signal s is sampled at discrete frequencies $\omega_m\,(m=1,...,N)$. In the DAP algorithm, the coefficients obtained by conventional LP are iteratively improved by minimizing the Itakura–Saito error ϵ_{IS} for discrete signals (El-Jaroudi and Makhoul, 1991):

$$\epsilon_{\mathrm{IS}}=\frac{1}{N}\sum_{m=1}^{N}\left(\frac{P(\omega_m)}{\hat{P}(\omega_m)}-\ln\frac{P(\omega_m)}{\hat{P}(\omega_m)}-1\right). \tag{5}$$

In this equation, $P=|S|^2$ indicates the power spectrum of the input signal [see Eq. (2)], $\hat{P}$ a model power spectrum. During the iteration, $\hat{P}$ is adapted to match P at N discrete frequencies ω_m [for details on the algorithm, see El-Jaroudi and Makhoul (1991)].

In the current study, the ω_m were chosen to reflect the spectral envelope as closely as possible. Each ω_m was to represent the mth harmonic of the fundamental frequency ω_0 and was defined in the following way. The local period lengths for each cycle in a given window were determined by the waveform matching algorithm (Titze and Liang, 1993; Parsa and Jamieson, 1999). The median period length $T_{0,\mathrm{med}}$ of the local window (consisting of ten consecutive periods) defined the reference normalized radian frequency $\tilde{\omega}_{0,\mathrm{ref}}=2\pi\cdot f_s^{-1}\cdot T_{0,\mathrm{med}}^{-1}$ (f_s: sampling frequency). For $m\geqslant 1$, the discrete normalized radian frequency closest to the mth harmonic $\tilde{\omega}_m=m\cdot\tilde{\omega}_{0,\mathrm{ref}}$ was chosen as ω_m to enter the DAP matching procedure (N_{spec}=total number of spectral samples):

$$\omega_m=\min(|m\cdot\tilde{\omega}_{0,\mathrm{ref}}-\omega_i|);\quad 1\leqslant i\leqslant N_{\mathrm{spec}}. \tag{6}$$

Values of ω_m were thus not equally spaced but represented the best sampling of the spectral envelope at the harmonics, even if the period length was not exactly the same for all ten periods of the local window. This was the case for the natural utterances where period lengths generally showed a variability (''jitter'') of the order 0.2%–0.4%.

The signal modifications $S'=S/E$ according to Eq. (3) were included into the DAP algorithm. This was realized by modifying the autocorrelation function (ACF) R of the source signal to which the ACF of the model $\hat{R}$ was matched. Originally, R is given by the inverse Fourier transform of P:

$$R(i)=\frac{1}{N}\sum_{m=1}^{N}P(\omega_m)e^{j\omega_m i}. \tag{7}$$

In a first modification with regard to the original algorithm (El-Jaroudi and Makhoul, 1991), this equation was replaced by the modified ACF R_{mod} where the power spectrum of the differentiated glottal flow $|E|^2$ is canceled from the signal spectrum before the ACF is calculated:

$$R_{\mathrm{mod}}(i)=\frac{1}{N}\sum_{m=1}^{N}\frac{P(\omega_m)}{|E(\omega_m)|^2}e^{j\omega_m i}. \tag{8}$$

A second modification concerns the initialization of the filter that was performed using the covariance method instead of the autocorrelation method (Markel and Gray, 1976). In a third modification, the termination criterion of the iteration was defined by a relative threshold instead of an absolute one. The complete modified algorithm (modifications with regard to the original algorithm are underlined) reads as

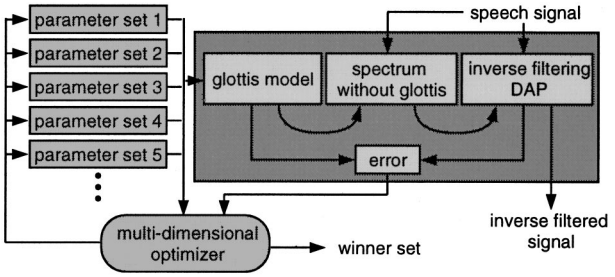


FIG. 2. Illustration of the design of the SIM method (for details, see text).

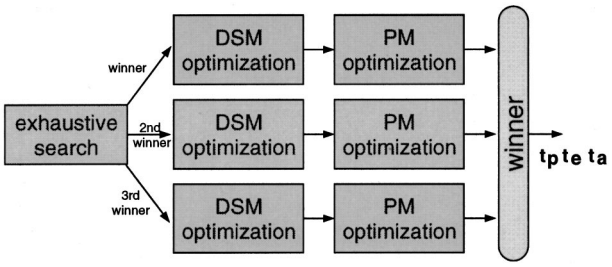


FIG. 3. The optimization procedure for the parameters t_p, t_e, t_a consists of three different runs of the DSM optimization followed by a PM optimization. The initializing simplices are determined on the basis of an exhaustive search.

(1) Calculate spectrum $E(\omega)$ of LF signal for a given parameter set (see Sec. II C).
(2) Calculate ACF $R_{\mathrm{mod}}(i)$ according to Eq. (8).
(3) Calculate initialization filter coefficients a_i by linear prediction (covariance method).
(4) Calculate $A(\omega_m)$, $1 \leqslant m \leqslant N$ as $A(\omega_m) = \sum_{k=0}^{p} a_k e^{-j\omega_m k}$.
(5) Calculate $p+1$ samples of the time-inverted impulse response of the prediction filter as $\hat{h}(-i) = (1/N)\sum_{m=1}^{N} e^{-j\omega_m i}/A(\omega_m)$.
(6) Calculate new filter coefficients a_i by solving $\sum_{k=0}^{p} a_k R_{\mathrm{mod}}(i-k) = \hat{h}(-i)$, $0 \leqslant i \leqslant p$ for a_i.
(7) Calculate Itakura–Saito error ϵ_{IS} according to Eq. (5).
(8) If $(\epsilon_{\mathrm{IS},\eta-1} - \epsilon_{\mathrm{IS},\eta})/(\epsilon_{\mathrm{IS},\eta}) > \xi$ (ξ=0.05: threshold, η: iteration step): increase η by 1, go to step 4.
(9) Normalize a_i so that $(1/N)\sum_{m=1}^{N} P(\omega_m)/\hat{P}(\omega_m) = 1$.
(10) End.

C. Multi-dimensional optimization

The modified DAP algorithm removes the glottal pulse shape and the radiation effect from the signal before inverse filtering. The true shape of the glottal flow pulse is not known. It is therefore replaced by an estimated model shape described by the LF model. However, the parameters of this estimated shape are also initially unknown so that an iterative multi-dimensional optimization of the LF model is performed. This is illustrated in Fig. 2.

In each iteration step of the optimization, the three parameters t_p, t_e, t_a (see Fig. 1) are taken as independent variables in a three-dimensional search space. They are estimated simultaneously as described below. During their optimization, E_e is set to a constant value. After t_p, t_e, t_a have been determined to the desired accuracy, E_e is estimated in a one-dimensional search. This two-step procedure results in an improved convergence compared to the simultaneous four-dimensional estimation of all LF model parameters. Furthermore, it considerably reduces the computational burden.

The most crucial issue in the design of an optimization procedure concerns the metric of the search space, in other words, the definition of the error ϵ. From many different error definitions tested in preliminary setups, the Itakura–Saito error ϵ_{IS} of Eq. (5) that is obtained during the IF procedure emerged as the most appropriate error in the optimization of t_p, t_e, t_a. This error is also adequate from a theoretical perspective to measure the success of the IF because it quantifies the flatness of the residual spectrum (El-Jaroudi and Makhoul, 1991).

The error ϵ in the three-dimensional optimization is calculated as follows: Each possible sequence of ten consecutive periods within the analysis frame (200 ms) is analyzed by the modified DAP algorithm.[2] For each sequence, ϵ_{IS} at the termination of the iteration is recorded as well as the corresponding filter coefficients. Finally, if at least one filter of the frame is stable, the smallest error value only of the stable filters is returned as ϵ, otherwise the smallest error of all (unstable) filters.[3]

Given the complexity of the relationship between ϵ_{IS} and the LF parameters t_p, t_e, t_a, multi-dimensional gradient methods that require the knowledge of the partial derivatives cannot be applied in the optimization of t_p, t_e, t_a using ϵ_{IS} as error criterion. Nevertheless, there exist several standard procedures that are suitable, such as the downhill simplex method (DSM) and the direction set method (Powell's method, PM) that do not require the knowledge of function derivatives. Both methods were implemented according to Press *et al.* (1988). They were applied as described below.

An adequate initialization of a given optimization procedure is of major importance if the global minimum is to be found. Parameters t_p, t_e, t_a are bounded, so that an exhaustive search of the three-dimensional data space $\{t_p, t_e, \log(t_a)\}$ can be performed. The volume of possible combinations is sampled at 61 ''nodes,'' and the corresponding errors are calculated in order to obtain good starting values without any interaction of the user [the term ''node'' is used to indicate a vector in the three-dimensional data space, for details, see Fröhlich (1999)]. The three nodes with the lowest error values ϵ supply the starting configurations for three independent optimization runs. Each run consists of a DSM optimization, after which the result is further improved by a subsequent PM optimization. In each DSM run, the initializing simplex is defined by the node itself and the three vectors obtained by changing each component by 10% one at a time [multiplying the particular parameter value by 1.1 (0.9) if the node value is in the lower (upper) half of the value range]. The complete procedure is illustrated in Fig. 3, top).

The fourth parameter E_e is determined in a one-dimensional optimization by Brent's method (Press *et al.*, 1988) after the optimal configuration of t_p, t_e, t_a has been obtained. An error criterion serves the squared difference

TABLE I. Overview of the different measures applied in this study. Top: list of the ''derived'' glottal measures calculated from the LF model parameters to characterize the estimated glottal flow. Bottom: reference measures used to assess phonation changes in the analysis of natural utterances.

	''Derived'' measures
open quotient	$OQ_1 = t_e - t_s(\sigma)$
open quotient	$OQ_2 = t_e + t_a - t_s(\sigma)$
closing quotient	$CQ_1 = t_e - t_p$
closing quotient	$CQ_2 = t_e + t_a - t_p$
speed quotient	$SQ_1 = \frac{OQ_1 - CQ_1}{CQ_1}$
speed quotient	$SQ_2 = \frac{OQ_2 - CQ_2}{CQ_2}$
parabolic spectral parameter	PSP=measure of the low-frequent spectral shape according to Alku *et al.* (1997)
	Reference measures
open quotient based on the EGG signal	OQ_{EGG}
signal energy	s

within the complete 200-ms frame between the IF signal and the LF signal that is synthesized using the appropriate period lengths, but with a constant LF parameter constellation. This error definition is the same as the one applied exclusively in other studies to estimate all four LF parameters (Milenkovic, 1986; Strik *et al.*, 1993). For optimal alignment between the LF model and IF signal, the error is calculated with a sample-by-sample increase of the lag for the LF signal (wrap-around) ranging from 0 to the length of one period. The minimum error value over all lags defines the final error in this part of the optimization procedure.

D. Measures of the glottal flow (''glottal measures'')

The matched glottal flow model can be characterized by different measures (referred to in the following as ''glottal measures'') to allow a numerical characterization of the voice generation. The first choice obviously is the LF model parameters themselves. However, other measures may offer an easier interpretation in terms of voice generation mechanisms. Those measures will be termed ''derived'' measures in order to distinguish them from the ''direct'' LF parameters t_p, t_e, t_a, E_e. The open quotient and the speed quotient are among the most prominent ''derived'' measures according to the literature (Alku and Vilkman, 1996; Childers and Lee, 1991; Childers and Ahn, 1995; Gobl and Chasaide, 1992; Hertegård and Gauffin, 1995; Sulter and Wit, 1996; Titze, 1994).

The open quotient (OQ) quantifies the duration of the open phase relative to the cycle length. In the LF model, it can therefore be defined as $OQ = t_e$ (Fant, 1995). However, if t_e is large, the exponential term in Eq. (4) leads to very small flow values at the beginning of a cycle despite its derivative being positive (see Fig. 1, bottom). This signal part should not be counted as part of the open phase because unrealistically large OQ values are obtained. A flow threshold σ is therefore introduced that has to be exceeded if the glottis is to be considered open.[4] The start of the open phase is then defined to occur at $t_s(\sigma)$. Additionally, the return phase may also be regarded as being part of the open phase:[5]

$$OQ_1(\sigma) = t_e - t_s(\sigma), \quad OQ_2(\sigma) = t_e + t_a - t_s(\sigma). \tag{9}$$

The choice of the definition depends on the point of view: physiologically, the time until the first contact between the vocal folds seems appropriate (OQ_1), while from a fluid dynamics approach the time until the flow has ceased should be given preference (OQ_2).

The closing quotient (CQ) quantifies the relative duration of the closing phase with respect to the cycle length, again either including or excluding the return phase:

$$CQ_1 = t_e - t_p, \quad CQ_2 = t_e + t_a - t_p. \tag{10}$$

The speed quotient (SQ) describes the temporal skewing between the opening and closing phase. It is also used in two definitions:

$$SQ_i = \frac{OQ_i - CQ_i}{CQ_i}, \quad i = 1,2. \tag{11}$$

A glottal measure introduced relatively recently is the parabolic spectral parameter (PSP, Alku *et al.*, 1997). It quantifies the low-frequency spectral shape of the estimated glottal flow spectrum by fitting a parabola $ax^2 + b$ to the low-frequency part of the pitch-synchronous power spectrum. The normalized coefficient a defines the PSP [for details, see Alku *et al.*, (1997)]. Differing from the original algorithm, the FFT length was not set to a fixed length in the PSP calculation, but chosen as eight times the local period length (Fröhlich and Alku, 2001). Table I gives an overview over the seven ''derived'' measures that were used in this study.

III. TESTS OF THE SIM METHOD

A. Data

The SIM method was tested using synthesized vowels and natural utterances. The synthetic signals were generated

according to the source-filter principle: for 504 LF configurations generated randomly, signals of 500-ms length were synthesized (sampling frequency $f_s = 10$ kHz, fundamental frequency $f_0 = 107$ Hz so that cycle lengths would be noninteger multiples of the sampling period $1/f_s$). Each signal was filtered by the same tenth-order all-pole filter representing the spectral shape of the vowel /a:/ (obtained by LP from the output of a speech synthesizer). Only the last 200 ms were used for analysis in order to guarantee that the analyzed signals were stationary.

The analysis of the natural utterances served as a first test to estimate the performance of the SIM method when used with ''real-world'' signals. Data were obtained in four recording sessions of one vocally healthy male. During each session, the subject phonated the vowel /ɛ:/ twice while changing the phonation gradually from ''hypofunctional'' (weak) to pressed and back again to hypofunctional, resulting in a total of eight utterances. Recordings took place in a reflection-free room, using a Sennheiser microphone (MKH 106T), a preamplifier (AXR Mic/Dat 2) and a DAT recorder (Pioneer D-07, $f_s = 48$ kHz). Mouth-to-microphone distance was 1 m. Period delimiters were determined automatically and checked by one of the authors (MF). Afterwards the signal was down-sampled to 10 kHz before further processing.

Simultaneous to the acoustic recordings, the electroglottographic signal (EGG) was recorded (EG8501, F-J Electronics Copenhagen). This EGG system features an automatic gain control so that amplitude-based measures could not be determined. The EGG signals were down-sampled to 10 kHz and high-pass filtered at 20 Hz with a second-order Butterworth filter.

B. Results for synthesized signals

1. Reference measures

The ''derived'' measures are commonly calculated from the waveform of the IF signal (Hertegård and Gauffin, 1995; Holmberg *et al.*, 1995; Sulter and Wit, 1996). For comparison purposes, they were therefore also estimated ''conventionally'' from the IF signal rather than from the fitted LF model.

For each period of the integrated IF signal (estimated glottal flow), index n_1 (indicating the instance of glottal opening) was defined as the location of the sample that first exceeded $0.1 \cdot v_i$ (v_i denotes the value range of period i). Index n_2 (indicating the instance of glottal closure) was defined as the location of the signal minimum in the case of OQ_1, to which in the case of OQ_2 the estimated value of t_a was added. The value of t_a was estimated by calculating the intersection point of the right-sided tangent at the minimum (approximated by a first-order finite difference) with the zero-axis. The value $(n_2 - n_1)/(T_0 \cdot f_s)$ defined the open quotient.

To calculate CQ_1 ''conventionally,'' the time difference between the location of the signal minimum and the zero crossing before the signal minimum was determined from the IF signal (corresponding to the glottal flow derivative). The resulting time span divided by $T_0 \cdot f_s$ defined CQ_1, while in the case of OQ_2 the estimated value of t_a was again added before the division. $SQ_{1|2}$ were calculated according to Eq. (11) using the ''conventional'' estimates of OQ and CQ. PSP was determined in the same way as for the SIM method, only this time based on the IF signal instead of the matched LF model.

TABLE II. Spearman's rank order correlation coefficient ρ and average relative difference δ between estimated results and original values. Data consisted of 504 synthetic signals that were randomly initialized. Left: results obtained by the SIM method (measures calculated from the LF model), right: results of the ''derived'' measures obtained conventionally from the IF signal.

	SIM		Conventional	
	ρ	δ (%)	ρ	δ (%)
t_p	0.76	15.9	⋯	⋯
t_e	0.69	19.0	⋯	⋯
t_a	0.76	13.0	⋯	⋯
E_e	0.92	[a]	⋯	⋯
OQ_1	0.99	7.0	0.82	58.0
OQ_2	0.99	6.2	0.79	60.1
SQ_1	0.88	12.3	0.87	105.3
SQ_2	0.93	13.4	0.87	61.6
CQ_1	0.97	12.2	0.97	13.6
CQ_2	0.99	13.3	0.95	18.3
PSP	0.99	21.9	0.37	605.4

[a]The δ value in the estimation of E_e was not determined because the estimated values approximated not the original but linearly transformed values that were caused by the filtering process during the generation of the pulses. Occasional outliers in the estimated values of this (unbounded) parameter would have imposed further restrictions on a meaningful redefinition of the average relative difference.

2. Reliability of estimation

The performance of the SIM method was assessed for the $N = 504$ synthetic signals by relating the estimated parameters to the original ones used during synthesis. Spearman's rank order correlation coefficient ρ (Press *et al.*, 1988) between the estimated (es) and original (or) parameter values was calculated as well as the average relative difference $\delta = (1/N) \sum_{i=1}^{N} |\mathrm{es}_i - \mathrm{or}_i| / \mathrm{or}_i$.[6] This was performed separately for the measures estimated by SIM and for the ''derived'' measures estimated conventionally. Results are stated in Table II. The values constitute conservative references with regard to natural speech, where deviations from the source-filter approach can be expected to lower the accuracy of the estimated parameters.

With the SIM method, correlations for the ''derived'' measures are high ($\rho = 0.88$ to 0.99). They clearly exceed the values of the time-based ''direct'' measures t_p, t_e, t_a ($\rho = 0.69$ to 0.76). If the ''derived'' measures were determined conventionally from the IF signal, high correlations were observed as well ($\rho = 0.79$ to 0.97), with the exception of PSP ($\rho = 0.37$). Nevertheless, the correlations obtained for the SIM-estimated measures exceed those values in almost all cases. Furthermore, δ values are much higher ($\delta > 58\%$) than for the corresponding SIM-estimated results, except for the CQ measures where δ values are comparable ($\delta < 20\%$). The closing quotient therefore seems to be the only ''derived''

TABLE III. Spearman's rank order correlation coefficient ρ between the glottal measures of Table I and the reference signals OQ_{EGG} and ς for all utterances 1–8. The number of frames n on which the correlation was based are stated for each utterance. Significances are corrected for multiple comparisons (Holm, 1979) within each utterance. Insignificant correlations are indicated by $^{\dagger}(p<0.05)$.

No.	n	Reference	t_p	t_e	t_a	E_e	OQ_1	OQ_2	SQ_1	SQ_2	CQ_1	CQ_2	PSP
1	84	OQ_{EGG}	$0.19^{\dagger}$	0.30	0.25	−0.74	0.63	0.64	−0.53	−0.56	0.71	0.74	0.70
		ς	−0.37	−0.46	−0.60	0.76	−0.72	−0.75	0.37	0.48	−0.66	−0.76	−0.81
2	80	OQ_{EGG}	0.45	0.54	$-0.26^{\dagger}$	−0.67	0.67	0.67	$-0.22^{\dagger}$	$-0.10^{\dagger}$	0.62	0.64	0.68
		ς	−0.57	−0.70	0.29	0.77	−0.85	−0.85	$0.23^{\dagger}$	$0.05^{\dagger}$	−0.84	−0.84	−0.88
3	55	OQ_{EGG}	$0.05^{\dagger}$	$0.33^{\dagger}$	−0.61	−0.52	$0.24^{\dagger}$	$0.24^{\dagger}$	$-0.24^{\dagger}$	$-0.13^{\dagger}$	$0.27^{\dagger}$	$0.24^{\dagger}$	$0.25^{\dagger}$
		ς	0.60	$0.12^{\dagger}$	$0.18^{\dagger}$	0.57	−0.82	−0.83	0.85	0.81	−0.87	−0.86	−0.86
4	58	OQ_{EGG}	$-0.09^{\dagger}$	$0.13^{\dagger}$	$0.06^{\dagger}$	−0.49	0.61	0.61	−0.60	−0.51	0.61	0.62	0.63
		ς	0.53	$0.28^{\dagger}$	$-0.33^{\dagger}$	0.85	−0.76	−0.78	0.80	0.79	−0.80	−0.81	−0.81
5	73	OQ_{EGG}	$-0.08^{\dagger}$	$-0.08^{\dagger}$	$-0.17^{\dagger}$	$-0.07^{\dagger}$	$0.03^{\dagger}$	$-0.06^{\dagger}$	$0.00^{\dagger}$	$0.05^{\dagger}$	$0.02^{\dagger}$	$-0.10^{\dagger}$	$-0.11^{\dagger}$
		ς	$0.30^{\dagger}$	$0.09^{\dagger}$	$-0.08^{\dagger}$	0.79	−0.52	−0.53	0.58	0.52	−0.58	−0.55	−0.53
6	59	OQ_{EGG}	$0.20^{\dagger}$	$0.29^{\dagger}$	$-0.20^{\dagger}$	−0.69	$0.28^{\dagger}$	$0.30^{\dagger}$	$-0.13^{\dagger}$	$0.02^{\dagger}$	$0.28^{\dagger}$	$0.20^{\dagger}$	$0.23^{\dagger}$
		ς	$-0.31^{\dagger}$	−0.42	$0.34^{\dagger}$	0.85	$-0.36^{\dagger}$	$-0.35^{\dagger}$	$0.18^{\dagger}$	$-0.03^{\dagger}$	$-0.35^{\dagger}$	$-0.23^{\dagger}$	$-0.27^{\dagger}$
7	69	OQ_{EGG}	$0.14^{\dagger}$	$0.08^{\dagger}$	$-0.34^{\dagger}$	−0.55	$-0.35^{\dagger}$	−0.36	$0.26^{\dagger}$	$0.24^{\dagger}$	$-0.34^{\dagger}$	$-0.33^{\dagger}$	$-0.35^{\dagger}$
		ς	$-0.12^{\dagger}$	$-0.12^{\dagger}$	$0.19^{\dagger}$	0.81	$0.29^{\dagger}$	$0.31^{\dagger}$	$-0.21^{\dagger}$	$-0.17^{\dagger}$	$0.29^{\dagger}$	$0.26^{\dagger}$	$0.29^{\dagger}$
8	73	OQ_{EGG}	−0.54	−0.40	0.37	−0.57	0.63	0.74	−0.67	−0.74	0.70	0.83	0.85
		ς	0.49	0.35	$-0.18^{\dagger}$	0.76	−0.57	−0.65	0.63	0.65	−0.65	−0.73	−0.73

measure that might be calculated from the automatically obtained IF signal by the conventional method with similar accuracy.

The effect of the variation of the period length (jitter) was tested by generating 49 additional LF configurations. For each configuration, signals were synthesized with randomly varying period lengths (Gaussian distributed, standard deviations set to 0%, 0.25%, 0.5%, 0.75%, 1.0%, 1.5%, 2.0%, 2.5%, 3.0%, 4.0%, 5.0% of the period length). Rank order correlations between the amount of jitter and the individual δ values were not significant for 538 out of the 539 cases (corrected for multiple comparisons, Holm, 1979). This supports that moderate jitter values should not be expected to systematically affect the accuracy of the estimated parameters in the analysis of natural utterances.

C. Results for natural utterances

1. Reference measures

For natural utterances, the ''correct'' model parameters are unknown. Therefore two ''external'' reference measures (i.e., measures that can be calculated independently of the SIM results) were determined in order to evaluate the consistency of the SIM results.

The first reference measure is the energy of each SIM analysis frame (i.e., nonoverlapping, rectangular frames of 200-ms length). The sequence of those frame-based energies defines the energy contour ς. It serves as a reference because it was observed that the energy varied with the realization of different phonation tasks (e.g., ''hypofunctional phonation'' was realized with a rather soft voice).

The second reference measure is the open quotient OQ_{EGG} based on the electroglottographic (EGG) signal that was recorded synchronously to the acoustic signal. The procedure is similar to the calculation of the open quotient described in other studies (Hertegård and Gauffin, 1995; Holmberg *et al.*, 1995): for each period *i*, the value range v_i of the EGG signal (i.e., the difference between maximum and minimum of the *i*th cycle) was determined using the period delimiters calculated for the acoustic signal. For this period, the quotient of the number of samples showing values greater than $0.6\,v_i$ above the minimum divided by the period length $T_{0,i}\cdot f_s$ in samples defined $OQ_{EGG}(i)$. This definition has to be regarded as an algorithmic definition of an ''open quotient'' rather than as a true approximation of the relative open time of the glottal cycle. Values were averaged within the 200-ms analysis frames to obtain one representative value for each frame.

Both measures were meant to supply a reference to the phonation changes for a first test using natural utterances. Both are relatively basic and have not been tuned for robustness under various conditions. Therefore, if the measures estimated by the SIM method show contours that are similar to OQ_{EGG} or ς (quantified by the correlation between the contours), they may be interpreted to reflect the changes in voice generation. On the other hand, if the contours are not correlated, it is difficult to decide from the data available whether this is due to inconsistencies in the estimation results or in the determination of OQ_{EGG} and ς.

2. Reliability of estimation

The correlations between the different estimated measures and the references OQ_{EGG} and ς are stated in Table III. While utterances 1 and 8 show significant correlations for almost all measures, utterances 6 and 7 show significant correlations only for E_e, and utterances 3 and 5 show significant correlations predominantly between the ''derived'' measures and ς. In one case (utterance 4), the ''derived'' measures show significant correlations to OQ_{EGG} although the direct LF parameters do not. E_e is in all cases more highly correlated to ς than to OQ_{EGG}. The correlations ρ between OQ_{EGG} and ς for utterances 1 to 8 are −0.64, −0.70, −0.29, −0.56, −0.17 (insignificant), −0.85, −0.60, −0.88, respectively.

The contours for utterance 8 are shown in Fig. 4. All measures except t_p, t_e, t_a show contours that start and end at high values and decrease noticeably in between, or vice versa. Parameters t_p and t_e show strongly elevated values in the range 4–9 s. However, the difference $t_e - t_p = CQ_1$ shows a perfectly inconspicuous curve with ρ values of 0.70

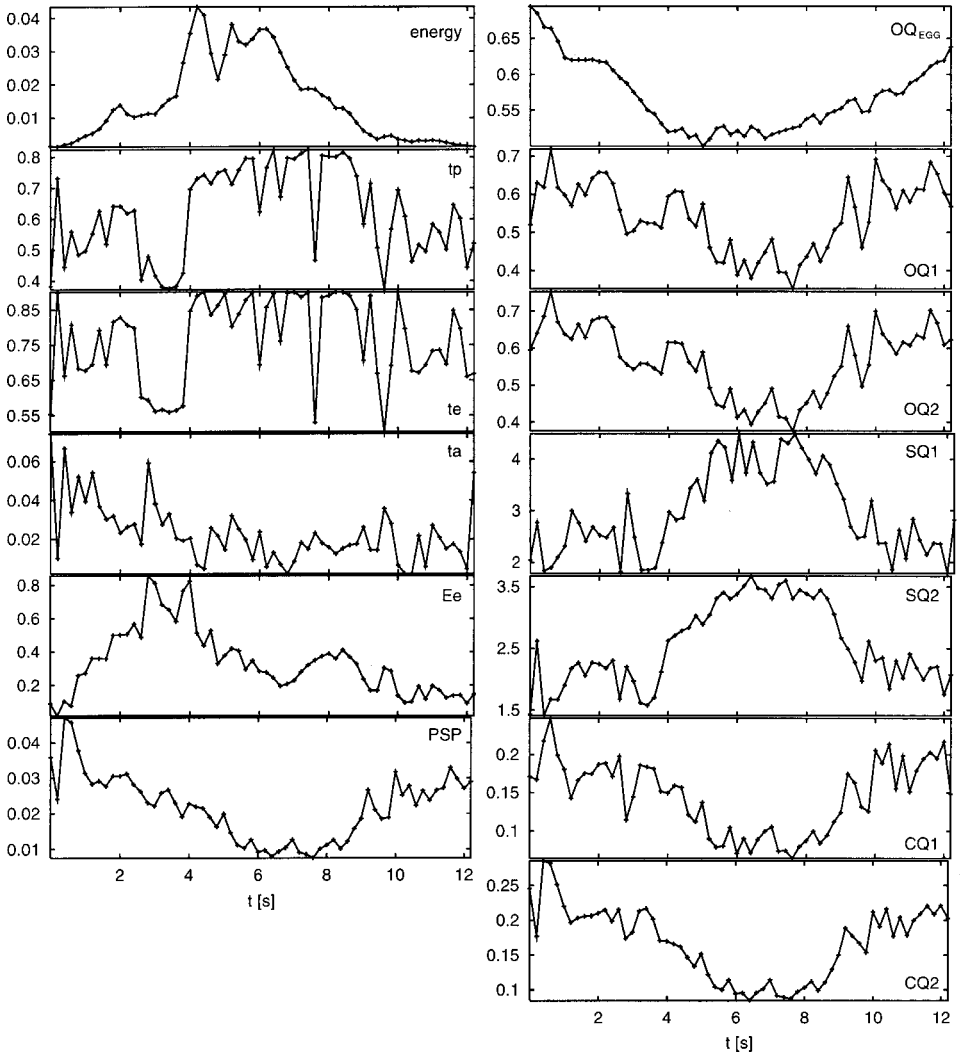


FIG. 4. Results for the different glottal measures during the gradual change from hypofunctional phonation to pressed phonation and back for utterance 8. The references OQ_{EGG} and the energy contour ς are shown at the top. The correlation coefficients between the contours and the references are stated in Table III.

(OQ_{EGG}) and −0.65 (ς). The contour of t_a does not seem to reflect the changes in voice generation, although its correlation to OQ_{EGG} (ρ=0.37) is significant. This indicates that the estimation of t_a was problematic for this utterance. E_e is remarkable because of its asymmetrical shape, with a maximum around 3 s. An asymmetry is also present in the energy contour ς, reflected by the highest correlation between ς and E_e of all glottal measures (ρ=0.76).

IV. DISCUSSION

The analysis of synthetic signals reveals that generally all measures can be estimated reliably, which shows in high individual correlations to the original values (ρ>0.69, see Table II). Average relative differences δ between estimated and original values are found within the range 6%–22%. The most problematic parameters appear to be the direct LF parameters t_p, t_e, t_a, for which correlation coefficients are comparatively low and δ values high. The ''derived'' measures mostly show extremely high correlations, whereas δ values are comparable to the ones for the ''direct'' LF parameters.

For t_p and t_e, those findings can be explained quite easily. If both t_p and t_e originally possessed high values, the glottal flow derivative increases very gradually at the beginning of the cycle (see Fig. 1). The open quotient measures that exclude this initial part of the cycle by applying the threshold σ can be estimated very reliably (ρ=0.99), which is also true for the difference $t_e - t_p = CQ_1$ (ρ=0.97) where the exact position of the beginning of the cycle is irrelevant. This implies that for some configurations the beginning of the glottal cycle cannot be reproduced reliably relative to the instance of glottal closure.

Problems in the estimation of t_a have already been reported in the literature (Fant, 1993). Possibly, the accuracy in the estimation of t_a could be improved by replacing ϵ_{IS} during the multi-dimensional optimization by an error criterion that applies a nonuniform spectral weighting. However, the fine tuning of the method to obtain improved estimates for certain measures remains a topic for future research.

Results for the natural utterances are somewhat ambiguous (see Table III). For some utterances, correlations between the estimated measures and the reference contours are insignificant. This may be caused by several factors: first, the SIM method may converge to ''wrong'' values; second, the reference measures may not have been calculated successfully; and third, the reference measures may be principally inadequate to reflect the change in phonation for some utterances. An indication to the presence of the second factor are

the correlations between ς and OQ_{EGG} that are low for some utterances (3 and 5). On the other hand, for some utterances (1 and 8) the correlations between the estimated measures and the reference contours are significant for almost all glottal measures. In those cases the estimated measures can be interpreted to reflect the change of phonation indicated by the two reference contours.

Looking closer at one example (utterance 8, see Table III and Fig. 4), the individual parameters are consistent with regard to findings reported in the literature. E_e shows low values for hypofunctional voice generation, which agrees with Gobl and Chasaide (1992). E_e also shows the highest (absolute) correlation to the signal energy ς of all measures, and for all other utterances a higher correlation to ς than to OQ_{EGG}. This reflects that E_e acts as a spectral scaling factor in the LF model (Fant, 1995). The estimated open quotient values decrease when changing voice generation from hypofunctional to pressed, which corresponds to the data of Alku and Vilkman (1996). Alku and Vilkman as well as Childers and Ahn (1995) also describe an increase in pulse skewing when changing from hypofunctional/breathy to pressed phonation, which can also be observed in Fig. 4. The value ranges correspond roughly to the values reported by Alku and Vilkman, although their open quotient values appear to be higher, their speed quotient values lower.

As the SIM method is potentially applicable in routinely performed examinations because it operates completely without supervision, it might be extended to apply different customized models of the glottal flow instead of the LF model. The choice of the LF model was motivated by its widespread use and by its ability to model different phonation types (Fant, 1993; Gobl and Chasaide, 1992). However, it may not offer the most appropriate parametrization of the glottal flow where voice disorders are concerned. One approach to develop more accurate models for specific vocal pathologies might be based on high-speed recordings of the vocal fold vibrations, despite the known problems in the mapping of the area function to the glottal flow (Hertegård and Gauffin, 1995). The model proposed by Titze (1989) might offer a possible starting point in the design of models customized for specific vocal pathologies, because it allows us to parametrize both the glottal flow and the glottal area function by the same parameters.

While having been tested just for acoustic signals recorded with a microphone, the proposed SIM method should also be applicable for the analysis of signals recorded with a flow mask. Additionally, the improved estimate of the vocal tract transfer function might be of interest in other contexts such as speech synthesis or speaker normalization.

V. CONCLUSION

A new inverse filtering method called SIM was proposed. It allows the calculation of measures describing the voice source in a way that does not require any user interaction. The LF model of the glottal flow derivative was integrated into the inverse filtering algorithm, thereby improving the estimation of the vocal tract resonance filter. The LF model parameters that optimally describe the estimated glottal flow were obtained in an iterative adaptation process. All glottal measures were calculated on the basis of the matched LF model.

The analysis of synthetic signals that were generated according to the source-filter approach revealed that the method is generally capable of reproducing the original values with high accuracy. The estimations of the LF parameters t_p, t_e, t_a were comparatively less accurate, which in the case of t_p and t_e can be attributed to constellations where t_p and t_e values are high so that the glottal flow derivative increases very gradually. The results estimated by the SIM method showed higher correlations to the original values than the values estimated conventionally from the IF signal.

Analysis results of natural utterances that represented gradual changes between different phonation types were encouraging, although the estimated parameter contours were interpretable only for some configurations with reference to the energy ς and the open quotient determined from the EGG signal. In those cases, measures that could be estimated most robustly were E_e, the open quotient, the speed quotient, the closing quotient, and the parabolic spectral parameter.

ACKNOWLEDGMENTS

We wish to express our gratitude to Professor E. Kruse for his encouragement and cooperation. Paavo Alku, Anders Löfqvist, and one anonymous reviewer provided many helpful comments on an earlier version of the manuscript. This research was part of projects funded by the Deutsche Forschungsgemeinschaft under Kr 1469/2-2 and Kr 1469/5-1.

[1]The implementation of the LF model used in this study was not corrected for aliasing that occured because $e(t)$ is not band-limited according to Eq. (4).

[2]For each analysis frame of the natural utterances, the median jitter was calculated from the jitter values of each ten-cycle sequence within the frame. The histogram of those median jitter values showed a unimodal distribution with a broad peak between 0.25% and 0.4% jitter. The variability of the period length was therefore well below 5%, which was the upper jitter limit tested for the synthesized pulses.

[3]The terms ''stable'' and ''unstable'' refer to the recursive infinite impulse response (IIR) filter V in Eq. (2). With regard to the *inverse* procedure, the finite impulse response (FIR) filter is stable in any case. Heuristic methods to transform an unstable IIR filter into a stable one as described by Press *et al.* (1988) resulted in a pronounced deterioration of the IF results. Therefore, the filter coefficients were used without modifications, even though their theoretical interpretation as the inverse of the original IIR filter was in this case unsatisfactory.

[4]The optimal value of σ was determined to be 0.2% of the value range of the period on the grounds of correlation analyses between open quotient values calculated by Eq. (9) and underlying ''true'' values. For details, see Fröhlich (1999).

[5]Strictly speaking, t_a is not the exact time t_{cease} until the flow has ceased, but the parameter controlling it (with $t_a < t_{cease}$). However, this difference is negligible with regard to the value of t_e ($t_{cease} - t_a \ll t_e$).

[6]Measure t_a was replaced by $\log(t_a)$ to determine the average relative difference δ.

Alku, P., and Vilkman, E. (**1994**). ''Estimation of the glottal pulseform based on discrete all-pole modeling,'' in *Proceedings of the International Conference on Spoken Language Processing 1994* (Yokohama), pp. 1619–1622.

Alku, P., and Vilkman, E. (**1996**). ''A comparison of glottal voice source quantification parameters in breathy, normal and pressed phonation of female and male speakers,'' Folia Phoniatr. Logop. **48**, 240–254.

Alku, P., Strik, H., and Vilkman, E. (**1997**). ''Parabolic spectral parameter—a new method for quantification of the glottal flow,'' Speech Commun. **22**, 67–79.
Childers, D., and Ahn, C. (**1995**). ''Modeling the glottal volume-velocity waveform for three voice types,'' J. Acoust. Soc. Am. **97**, 505–519.
Childers, D., and Lee, C. (**1991**). ''Vocal quality factors: Analysis, synthesis, and perception,'' J. Acoust. Soc. Am. **90**, 2394–2410.
El-Jaroudi, A., and Makhoul, J. (**1991**). ''Discrete all-pole modeling,'' IEEE Trans. Signal Process. **39**, 411–423.
Fant, G. (**1993**). ''Some problems in voice source analysis,'' Speech Commun. **13**, 7–22.
Fant, G. (**1995**). ''The LF-model revisited. Transformations and frequency domain analysis,'' Speech Transmission Laboratory—Quarterly Progress and Status Report (Stockholm) **2-3/1995**, 119–156.
Fant, G., Liljencrants, J., and Lin, Q. (**1986**). ''A four-parameter model of glottal flow,'' Speech Transmission Laboratory—Quarterly Progress and Status Report (Stockholm) **4/1985**, 1–13.
Fröhlich, M. (**1999**). ''Simultane Inversfilterung und Schätzung des glottalen Flusses aus akustischen Stimmsignalen,'' Ph.D. thesis, Georg-August Universität Göttingen; http://webdoc.sub.gwdg.de/diss/1999/froehlich/
Fröhlich, M., and Alku, P. (**2001**). ''Quantizing effects in the calculation of the spectral decay of the glottal source using the parabolic spectral parameter'', Speech Commun. (submitted).
Gobl, C., and Chasaide, A. N. (**1992**). ''Acoustic characteristics of voice quality,'' Speech Commun. **11**, 481–490.
Hertegård, S., and Gauffin, J. (**1995**). ''Glottal area and vibratory patterns studied with simultaneous stroboscopy, flow glottography, and electroglottography,'' J. Speech Hear. Res. **38**, 85–100.
Holm, S. (**1979**). ''A simple sequentially rejective multiple test procedure,'' Scand. J. Statist. **6**, 65–70.
Holmberg, E. B., Hillman, R. E., Perkell, J. S., Guiod, P. C., and Goldman, S. L. (**1995**). ''Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures,'' J. Speech Hear. Res. **38**, 1212–1223.
Kasuya, H., Maekawa, K., and Kiritani, S. (**1999**). ''Joint estimation of voice source and vocal tract parameters as applied to the study of voice source dynamics,'' in *Proceedings of ICPhS '99* (San Francisco), pp. 2505–2512.
Kröger, B. (**1991**). ''Zur Auswirkung der Glottis-Sprechtrakt-Kopplung auf die Stimmreinheit,'' Sprache-Stimme-Gehör **15**, 139–142.
Ma, C., Kamp, Y., and Willems, L. F. (**1994**). ''A Frobenius norm approach to glottal closure detection from the speech signal,'' IEEE Trans. Acoust., Speech, Signal Process. **2**, 258–265.
Markel, J., and Gray, A., Jr. (**1976**). *Linear Prediction of Speech*, Vol. 12 of *Communication and Cybernetics* (Springer-Verlag, Berlin).
Michaelis, D. (**2000**). ''Das Göttinger Heiserkeits-Diagramm—Entwicklung und Prüfung eines akustischen Verfahrens zur objektiven Stimmgütebeurteilung pathologischer Stimmen,'' Ph.D. thesis, Georg-August Universität Göttingen; http://webdoc.sub.gwdg.de/diss/2000/michaelis/
Milenkovic, P. (**1986**). ''Glottal inverse filtering by joint estimation of an AR system with a linear input model,'' IEEE Trans. Acoust., Speech, Signal Process. **ASSP-34**, 28–41.
Parsa, V., and Jamieson, D. G. (**1999**). ''A comparison of high precision F0 extraction algorithms for sustained vowels,'' J. Speech Lang. Hear. Res. **42**, 112–126.
Press, W., Flannery, B., Teukolsky, S., and Vetterling, W. (**1988**). *Numerical Recipes in C* (Cambridge U. P., Cambridge).
Rosenberg, A. (**1970**). ''Effect of glottal pulse shape on the quality of natural vowels,'' J. Acoust. Soc. Am. **49**, 583–590.
Rothenberg, M. (**1973**). ''A new inverse-filtering technique for deriving the glottal air flow waveform during voicing,'' J. Acoust. Soc. Am. **53**, 1632–1645.
Schoentgen, J. (**1982**). ''Quantitative evaluation of the discrimination performance of acoustic features in detecting laryngeal pathology,'' Speech Commun. **1**, 269–282.
Shadle, C. H., Barney, A., and Davies, P. (**1999**). ''Fluid flow in a dynamic mechanical model of the vocal folds and tract. II. Implications for speech production studies,'' J. Acoust. Soc. Am. **105**, 456–466.
Södersten, M., Håkansson, A., and Hammarberg, B. (**1999**). ''Comparison between automatic and manual inverse filtering procedures for healthy female voices,'' Logopedics Phoniatrics Vocology **24**, 26–38.
Strik, H., Cranen, B., and Boves, L. (**1993**). ''Fitting a LF-model to inverse filter signals,'' in *Eurospeech '93*, Vol. 1, pp. 103–106.
Strube, H. W. (**1974**). ''Determination of the instant of glottal closure from the speech wave,'' J. Acoust. Soc. Am. **56**, 1625–1629.
Sulter, A. M., and Wit, H. P. (**1996**). ''Glottal volume velocity waveform characteristics in subjects with and without vocal training, related to gender, sound intensity, fundamental frequency, and age,'' J. Acoust. Soc. Am. **100**, 3360–3373.
Titze, I. R. (**1989**). ''A four-parameter model of the glottis and vocal fold contact area,'' Speech Commun. **8**, 191–201.
Titze, I. R. (**1994**). *Principles of Voice Production* (Prentice Hall, Englewood Cliffs, NJ).
Titze, I. R., and Liang, H. (**1993**). ''Comparison of F_0 extraction methods for high-precision voice perturbation measurements,'' J. Speech Hear. Res. **36**, 1120–1133.
Wakita, H., and Fant, G. (**1978**). ''Toward a better vocal tract model,'' Speech Transmission Laboratory—Quarterly Progress and Status Report (Stockholm) **1/1978**, 9–29.

On the relationship between identification and discrimination of non-native nasal consonants

James D. Harnsberger[a)]
Department of Linguistics, University of Michigan, 1076 Frieze Building, Ann Arbor, Michigan 48109

(Received 16 August 2000; accepted for publication 20 March 2001)

To examine the relationship between the identification and discrimination of non-native sounds, nasal consonants varying in place of articulation from Malayalam, Marathi, and Oriya were presented in two experiments to seven listener groups varying in their native nasal consonant inventory: Malayalam, Marathi, Punjabi, Tamil, Oriya, Bengali, and American English. The experiments consisted of a categorial AXB discrimination test and a forced-choice identification test with category goodness ratings. The identification test results were used to classify the non-native contrasts as one of five ''assimilation types'' of the Perceptual Assimilation Model (PAM) that are predicted to vary in their relative discriminability: two-category (TC), uncategorizable–categorizable (UC), both uncategorizable (UU), category-goodness (CG), and single-category (SC). The results showed that the mean percent correct discrimination scores of the assimilation types, but not the range of scores, were accurately predicted. Furthermore, differences in category goodness ratings in the CG and SC assimilations that were predicted to correlate with discrimination showed a weak, but significant correlation ($r=0.31$, $p<0.05$). The implications of the results for models of cross-language speech perception were discussed, and an alternative model of cross-language speech perception was outlined, in which the discriminability of non-native contrasts is a function of the similarity of non-native sounds to each other in a multidimensional, phonologized perceptual space. [DOI: 10.1121/1.1371758]

PACS numbers: 43.71.An, 43.71.Hw [KRK]

I. INTRODUCTION

A. Background

Cross-language speech perception research has shown that listeners' abilities to discriminate some non-native contrasts can be constrained by the phonemic distinctions employed in their native language (Abramson and Lisker, 1970; Miyawaki *et al.*, 1975; Werker *et al.*, 1981). The effect of linguistic experience has also been shown to vary depending on the non-native contrast and listener group in question. For instance, Polka (1991) and Pruitt (1995) have demonstrated that the discriminability of Hindi dental-retroflex stop contrasts for American English listeners can vary significantly as a function of voicing or manner class, despite the fact that American English listeners have only a single native category /t/ or /d/ that corresponds to these contrasts. Several factors have been proposed to account for this variation in the discriminability of non-native contrasts, such as the psychophysical salience of the contrast (Sheldon and Strange, 1982; Burnham, 1986; Polka, 1991); a listener's general experience with features employed in the contrast (Werker *et al.*, 1981; Polka, 1992); the effect of allophonic variants, or phonetic realizations, of the native category (Henly and Sheldon, 1986; Ingram and Park, 1998); and the degree of similarity between a contrast and native perceptual categories (Best *et al.*, 1988; Best, 1995; Kuhl, 1991; Iverson and Kuhl, 1995, 1996).

These four factors have been cited in one form or another in the cross-language speech perception literature, but only degree of similarity and psychophysical salience have been incorporated in formal models, such as the Perceptual Assimilation Model (PAM) (Best, 1995), the Native Language Magnet Model (NLM) (Iverson and Kuhl, 1995), and the Speech Learning Model (SLM) (Flege, 1995). PAM and NLM both focus on the listener's native perceptual categories and how stimuli are filtered by these categories. NLM has been developed through the exploration of listeners' perceptual categories using synthetic stimuli based on native sounds. On the other hand, the research that inspired PAM has traditionally focused on the perception of novel, non-native, natural stimuli by naive listeners. SLM, like PAM and NLM, is concerned with cross-language speech perception and, more importantly, with second language acquisition in production and perception. These models are discussed in Sec. I B in terms of their capacity to predict the relationship between the identification and discrimination of non-native sounds.

B. Cross-language speech perception models

1. Perceptual assimilation model (PAM)

In PAM, non-native contrasts are perceived, if possible, in terms of their phonetic, specifically gestural, similarity to the phonological categories present in a listener's native language (L1). The degree of gestural similarity determines the matching between non-native phones and L1 categories, a process described in terms of six possible patterns of *assimilation* of members of the contrast. Three assimilation types concern non-native sounds that are consistently identified

[a)]Now at Speech Research Laboratory, Department of Psychology, Indiana University, Bloomington, IN 47405.

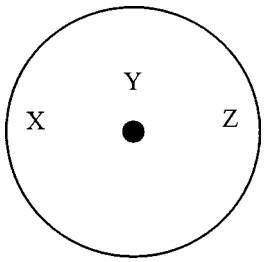


FIG. 1. A hypothetical category, represented as a circle, with its center represented as a black dot. *X*, *Y*, and *Z* are non-native stimuli that assimilate to this category.

with native categories: Two-category (TC), category-goodness (CG), and single-category (SC), listed in descending order of predicted discriminability. TC assimilations correspond closely to native phonemic contrasts. CG and SC assimilations both involve contrasts that are identified consistently with a single native category. CG and SC assimilations differ in how closely each segment in the contrast is identified with the single category. SC assimilations are ones in which both non-native sounds are either equally similar or equally dissimilar to a single native category. In contrast, CG assimilations describe two non-native sounds that differ from one another in category goodness. Category-goodness refers to the perceived similarity between a sound and the native category with which it is identified.

Within PAM, the discriminability of SC and CG assimilations is the product of the magnitude of the difference in category-goodness between the sounds that make up the non-native contrast: the greater the difference in category-goodness between the sounds of a contrast, the more discriminable the contrast. This aspect of PAM is the most testable and novel of the model. It is easily testable since it only requires measures of category-goodness and discriminability, as opposed to measures of the contrast's psychophysical or gestural differences.[1] It is a novel and rather interesting prediction that eliminates the role of psychophysical difference in predicting the discriminability of non-native contrasts that assimilate to single native categories. In fact, the discriminability of two contrasts that vary in psychophysical difference can be the opposite of that predicted by category-goodness difference. Such a hypothetical case is illustrated in Fig. 1, which shows an abstract, two-dimensional perceptual category (the circle) with a category center or prototype (the black dot). *X*, *Y* and *Z* in Fig. 1 are non-native stimuli that assimilate to this category. *X* and *Z* are judged as poor exemplars of the category, and thus appear in the category periphery; *Y* is judged to be a good exemplar of the category. According to PAM, the *X*-*Y* and *Y*-*Z* contrasts are more discriminable than *X*-*Z* because they involve two sounds that differ in category goodness. *X* and *Z*, while differing greatly from one another in psychophysical terms (as denoted by their Euclidean distance), should be harder to discriminate because they do not differ in their perceived category-goodness (they are both equally dissimilar to the category center).

The other three PAM assimilation types involve contrasts in which either one or both sounds are not consistently identified with any single native category (uncategorizable vs categorized—UC; both uncategorizable—UU; or both are not identified as speech sounds at all, nonassimilable—NA). The predicted discriminability of these three assimilation types is not clearly rank ordered with TC, CG, and SC. For UU contrasts, discriminability is expected to vary according to the segments' proximity to each other and to their proximity to L1 categories (Best, 1995). Since these two proximities are not clearly defined, or weighted with respect to one another, UU contrasts cannot be rank ordered. For UC contrasts, discriminability is predicted to be ''very good,'' which may correspond to the discriminability of CG assimilation types (Best, 1995, p. 195). Finally, NA contrasts are predicted to be discriminable as a function of their absolute psychophysical differences, given that non-native listeners are perceiving them in a ''nonspeech,'' or phonologically ''unfiltered,'' mode.

2. Native language magnet (NLM) model

The NLM model is the product of research by Patricia Kuhl and others into developmental change, in which infants begin life as language-general perceivers and, over the course of acquiring their native language, process speech in a language-specific mode (Kuhl, 1991; Iverson and Kuhl, 1995, 1996). Central to NLM is the concept of a category prototype, a good exemplar of a category that exerts a ''warping,'' or ''magnet,'' effect on nearby perceptual space, collapsing fine-grained phonetic distinctions near the prototype. The formation of such prototypes is argued to be the hallmark of language acquisition, a process that accounts for the perceptual difficulties listeners may encounter when discriminating a contrast from another language. The earliest work demonstrating the magnet effect involved synthetic [i] stimuli. Subsequent studies attempting to replicate Kuhl's work with [i] (Lively and Pisoni, 1997; Lotto *et al.*, 1998) and other vowels (Sussman and Gekas, 1997) have shown within-category variation in goodness ratings, but not a subsequent effect on the discrimination of stimuli pairs.

Like PAM, predictions based on NLM for non-native contrast perception are a product of category-goodness, that is, the perceived similarity between a stimulus and the prototype, or center, of a category. Unlike PAM, NLM predicts that the discriminability of two stimuli is a function of their psychophysical differences, weighted by their ''location'' in a category. When both stimuli fall close the prototype, the magnet effect reduces their discriminability. At the periphery, perceptual sensitivity is more of a function of their psychophysical differences. NLM differs from PAM in this latter prediction. In Fig. 1, NLM predicts that *X* and *Z* should be highly discriminable given that they are at the category periphery and that they differ from one another greatly in terms of their psychophysical differences. In contrast, *X*-*Z* is a single-category assimilation in PAM, since *X* and *Z* do not differ in their ''phonetic distance'' from the category prototype (i.e., their category goodness difference).

NLM is silent concerning a third possible pairing, that of a good, or prototypical, exemplar and a nonprototypical one. Because the nonprototypical exemplar is removed from the magnetic effects of the prototype, discrimination might be expected to be more accurate than for a pair of prototypical

exemplars. However, it is unclear how to order nonprototypical–prototypical and nonprototypical–nonprototypical pairs in terms of discriminability. Presumably, both pairs' discriminability is a function of their psychophysical differences.[2] For natural stimuli, though, it is unclear how this measure might be obtained. Category-goodness ratings alone only capture the perceptual ''distance'' between a stimulus and a prototype, not the distance between two stimuli that form a contrast. An acoustic analysis of natural stimuli could provide some measure of psychophysical difference. However, the complexity of natural stimuli make it problematic to assume that the particular measures taken (i.e., $F1$,$F2$, duration of vowels) constitute most or all of the perceptually relevant differences between the stimuli, particularly in experiments involving ''exotic'' contrasts that have not examined in much (or any) prior work. Thus in its current form, the NLM model is limited in its capacity to be tested with stimuli of greater ecological validity than synthetic sounds.

3. Speech learning model (SLM)

SLM has been proposed to account for the changes that occur in both the perception and production of the first and second language of an L2 learner. Flege (1995) sets out its hypotheses, and those pertaining to the relationship between the identification and discrimination of non-native sounds by naive listeners can be summarized as follows:

(1) The position-dependent allophonic level of articulatory/acoustic detail best characterizes L1 perceptual categories.
(2) New phonetic categories can be established given a certain degree of dissimilarity from L1 categories.
(3) *Equivalence classification* can occur, linking non-native phones to a single L1 category.
(4) Established L2 perceptual categories may differ from the equivalent categories of native speakers because of the need to maintain contrasts in the common L1/L2 space of the bilingual, or as a product of L1 patterns of weighing individual features in perception.

SLM differs from PAM and the NLM model in several respects. SLM does not specify the nature of the similarity relation that governs the identification of non-native sounds with native categories (see #2 above). Thus it could be acoustic/auditory or gestural in nature, as assumed in the NLM and PAM models, respectively. SLM is specific in its choice of a unit of processing in cross-language speech perception (#1). Unlike PAM,[3] SLM allows for the notion of language-specific weighting of the phonetic features of phonemes or allophones (#4), although the role it might play in cross-language speech perception is not defined. SLM resembles NLM and PAM in terms of its ''equivalence classification'' process (#3), which codifies the long-standing observation that some non-native contrasts are difficult to discriminate when they are strongly identified with a single native category. This observation is accounted for in NLM in terms of the perceptual magnet effect, while in PAM, such cases are classified as SC assimilations.

Like NLM, SLM is not specific enough to predict the discrimination of naturally produced non-native contrasts given a particular identification pattern. This limitation of SLM is due to the vagueness of the similarity metric referenced in the model. Moreover, it reflects the fact that SLM focuses more on the acquisition of L2 categories, and the link between L2 production and perception, rather than the initial perception of non-native sounds.

4. Summary

The three models of cross-language speech perception surveyed here share a number of commonalities, particularly in their focus on the internal structure of perceptual categories (NLM and PAM) and on the effect of the entire perceptual category inventory on the perception of non-native sounds (PAM and SLM). The results of this study have implications for all three models. However, of these three models, only PAM makes a number of specific, falsifiable claims concerning the relationship between the identification and discrimination of naturally produced speech sounds. NLM and SLM are not suited for testing the relationship between identification and discrimination, either because of the theoretical foci of the models (SLM), or because of limitations in our capacity to describe the characteristics of natural stimuli for the purposes of testing (NLM). Thus specific predictions for the discrimination of the contrasts examined in this study will only be provided for PAM (see Sec. II B).

C. Present study

The goal of this study was to examine the relationship between the identification and discrimination of non-native sounds, using both a broad range of non-native contrasts and listener groups. Non-native contrasts were elicited from speakers of three languages, representing a range of potentially challenging contrasts. These contrasts were presented in a discrimination and identification test to seven listener groups representing a variety of native phonological inventories deemed similar to the non-native contrasts in question. The identification test results have been reported and discussed in detail by Harnsberger (2000) in terms of the relationship between native phonological inventories and the identification of non-native sounds. This paper focuses on the discriminability of these non-native contrasts.

Specifically, this study employed nasal contrasts combining bilabial, dental, alveolar, and retroflex nasals, varying in talker, syllabic context, and vowel context. They were elicited from two talkers each of three languages: Malayalam, a Dravidian language spoken primarily in the southern Indian state of Kerala; Marathi, an Indo-Aryan language spoken primarily in Maharashtra and several other states in western India; and Oriya, an Indo-Aryan language spoken primarily in the state of Orissa in eastern India. A nasal series varying in place of articulation constituted a stimulus set that had not been examined in previous cross-language speech perception studies. Of the three languages, Malayalam is the only one to possess the full complement of nasals. However, this language exhibits a number of phonotactic constraints that impose limitations on the kind of stimuli

TABLE I. The demographics of the talkers. ''NL''=native language. ''Home''=home city or district within India. ''Years''=years outside of an environment where the native language is widely spoken.

Name	NL	Sex	Age	Home	Years	Other languages spoken
Ym	Malayalam	m	58	Malabar	29[a]	English, Hindi
Ys	Malayalam	f	47	Malabar	26[a]	English, Hindi, Tamil
Ms	Marathi	m	26	Mumbai	1	English, Hindi
Mv	Marathi	f	35	Mumbai	6.5	English, Hindi, Gujarati
Oc	Oriya	f	35	Cuttack	5	English, Hindi, Marathi, Bengali
Os	Oriya	f	30	Bhubaneswar	9	English, Telegu, Hindi

[a]While both Malayalam talkers had spent half or more of their lifetimes outside of their home state of Kerala, where Malayalam is spoken, both reported using Malayalam regularly at home with their Malayalam-speaking spouses. Moreover, their stimuli were consistently identified correctly by three native speakers of Malayalam.

that could be elicited from native speakers: all four nasals contrast only when they occur as intervocalic geminates. To ensure that the results were generalizable beyond Malayalam intervocalic geminates, nasals from Marathi and Oriya were also included. These languages allow some of these contrasts as singletons in medial and final position.

These stimuli were presented in a categorial AXB discrimination test and an identification test with category goodness ratings to seven listener groups varying in their coronal nasal consonant inventory: Malayalam (dental-alveolar-retroflex), Marathi and Punjabi (dental-retroflex), Tamil and Oriya (alveolar-retroflex), and Bengali and American English (bilabial-alveolar). Tamil is a Dravidian language spoken primarily in the Indian state of Tamil Nadu and the nation of Sri Lanka. Punjabi is an Indo-Aryan language spoken primarily in northern India, in the states of Punjab and Harayana, and India's capital, New Delhi. Bengali is also an Indo-Aryan language, spoken primarily in the Indian state of West Bengal and in the nation of Bangladesh. Multiple listener groups and non-native contrasts were used to ensure that the results were generalizable across other languages and contrasts and to increase the likelihood that a range of assimilation types was elicited for the purpose of evaluating PAM.

II. DISCRIMINATION AND IDENTIFICATION TESTS

A. Method

1. Stimulus materials

The stimulus materials for this experiment were identical to those described by Harnsberger (2000). Briefly, six talkers, two each of Malayalam, Marathi, and Oriya, were recorded reading from a list of real and nonsense words from their native language. The demographics of the six talkers are shown in Table I. The nasals of interest appeared in all syllable positions allowable by the individual languages, in an [a], [i], or [u] vocalic context. All of the stimuli recorded were evaluated by native speakers of the respective languages in an identification test in order to exclude any stimuli from use in the experiment that might be poor exemplars. Of the stimuli that were recorded and evaluated, a subset was used in the experiment: 4 exemplars (produced in isolation), 2 from each talker, of 18 types of stimuli. These stimulus types are listed in Table II.[4] They include bilabial, dental, alveolar, and retroflex nasals from Malayalam; dental and retroflex nasals from Marathi; and alveolar and retroflex nasals from Oriya.

2. Participants

The participants in this study were the same as those described by Harnsberger (2000). Native speakers of Malayalam, Marathi, Punjabi, Tamil, Oriya, Bengali, and American English were chosen for this experiment to represent a range of nasal consonant inventories. A description of the nasal consonant inventories of these seven groups, at the allophonic level of detail, appears in Table III. The studies referenced in the development of these inventories are described by Harnsberger (2000).

Fifteen to eighteen speakers of each group were recruited and tested. Some subjects were eventually excluded from the results analysis due to frequent lapses in answering within the test sequences, leaving 12 to 18 subjects per listener group for a given test. Table IV lists by gender the number of subjects that were recruited for each listener group, along with the mean age of each listener group.

All but the English listeners were tested in India, in order to recruit subjects who varied little in terms of age, dialect spoken, and overall linguistic experience. This last criterion was especially important. If listeners who belonged to a particular group also had experience in another language with a richer set of nasal contrasts, their identification results might be influenced by these contrasts. To limit the effect of

TABLE II. Stimuli and their source languages. The vocalic context was [a] for all but the underlined stimuli. Underlining indicates that the stimulus appeared in [i] as well as [a] contexts. Note: The dental nasal of one of the Malayalam talkers (Ym) was produced as an interdental.

		Nasal			
Language	Syllable	m	n̪	n	ɳ
Malayalam	VCV			√	√
Marathi			√		√
Oriya				√	√
Malayalam	VC:V	√̲	√̲	√̲	√̲
Marathi			√		√
Oriya					
Malayalam	VC				
Marathi			√		√
Oriya					

TABLE III. The perceptual category inventories for the seven listener groups, at the allophonic level of detail. ''Syllable''=syllabic context in which the nasal appears.

Group	Syllable	Perceptual Category			
		Bilabial	Dental	Alveolar	Retroflex
Malayalam	VCV	m		n	ɳ
	VC:V	m	n̪	n	ɳ
	VC	m		n	
Marathi	VCV	m	n̪		ɳ
	VC:V	m	n̪		ɳ
	VC	m	n̪		ɳ
Punjabi	VCV	m	n̪		ɳ
	VC:V	m	n̪		
	VC	m	n̪		ɳ
Tamil	VCV	m		n	ɳ
	VC:V	m		n	ɳ
	VC				
Oriya	VCV	m		n	ɳ
	VC:V				
	VC[a]	m		n	ɳ
Bengali	VCV	m		n	
	VC:V	m		n	
	VC	m		n	
English	VCV	m		n	
	VC:V				
	VC	m		n	

[a]Oriya allows/disallows final consonants, depending on the dialect spoken (see Harnsberger, 1998, for a summary).

the bi- or multi-lingualism of the listeners as a source of significant variability in the results, only those listeners were tested who had no experience in a language which employed a greater number of coronal nasal contrasts than those in the listener's first language. The linguistic background of the individual listeners was reported by Harnsberger (2000).

All subjects from India were recruited by posting flyers at local universities. The English listeners were recruited through introductory linguistics classes at the University of Michigan. The Malayalam, Oriya, and Marathi listeners were students attending national universities in India's capital, New Delhi. The Bengali listeners were university students who were recruited and tested in Calcutta, the capital of West Bengal, where Bengali is primarily spoken. The Punjabi listeners were recruited and tested in Amritsar, the cultural center of the Punjab state in northwestern India.

TABLE IV. The number of subjects in each listener group. ''N''=number of participants, ''M''=number of male participants, ''F''=number of female participants.

Listener group	N (M, F)	Mean age
Malayalam	18 (12, 6)	24
Marathi	17 (10, 7)	23
Punjabi	14 (13, 1)	22
Oriya	16 (12, 4)	24
Tamil	12 (9, 3)	22
Bengali	15 (9, 6)	24
American English	18 (3, 15)	19

3. Procedure

a. Identification test. The procedures for the identification test were reported by Harnsberger (2000). Stimuli were presented to subjects binaurally over Sony MDR-7506 headphones connected to a Sony TCD-D8 portable DAT recorder. Responses were made on photocopied answer sheets. The experiment consisted of an orthographic classification task in which listeners used response sets based on native phonemic categories. Specifically, the response choices were a closed set consisting of individual letters in the native orthography. Listeners were also instructed to provide category goodness ratings on a five-point scale, a task that would allow listeners to note the degree of difference between a non-native stimulus and the nearest native category. Listeners identified and rated 2 exemplars each of the 36 stimulus types, for a total of 72 stimuli. ''Stimulus type'' in this study refers to a nasal produced in a particular place of articulation, in a particular syllabic and vocalic context, by a particular talker. The identification test included 2 repetitions of this set, presented in random order, for a total of 144 trials, resulting in 4 judgments by listeners for each stimulus type. The test employed a 6 s intertrial interval and a 6 s interblock interval, with 10 trials per block. Listeners were instructed to ignore ''irrelevant differences'' of duration, tone, or voice quality. Before the test began, ten trials were presented (with no feedback from the investigator) to familiarize listeners with the time allotted for labeling and rating a stimulus.

b. Discrimination test. The discrimination test was a categorial AXB test consisting of 544 trials presented in random order, 16 trials each of the 34 different types of contrasts, where ''contrast'' refers to a particular place distinction in a particular syllabic and vocalic context, produced by a particular talker.[5] A list of the non-native contrasts, ordered by place of articulation, talker, and context (vocalic and syllabic), is given in Table V. As with the identification test, stimuli were presented binaurally over Sony MDR-7506 headphones connected to a Sony TCD-D8 portable DAT recorded. Responses were also made on photocopied answer sheets.

In order to ensure that the results were not dependent on the intelligibility of a single exemplar, 2 exemplars of each member of the 34 contrasts were used. The contrasts appeared in four possible orders, AAB, ABB, BAA, BBA. A and B were always from the same talker, and all stimuli that were paired together were selected to minimize acoustic differences that were not relevant to the identity of the stimulus, such as the overall duration or the fundamental frequency pattern of a stimulus. The interstimulus interval for the discrimination test was 1 s, an interval used in earlier works to elicit cross-language differences in perceptual performance (Best *et al.*, 1988; Best, 1996). The intertrial interval was 3 s and the interblock interval was 6 s, with 20 trials per block. The total time for the discrimination test was 58.5 min.

Subjects were told to indicate whether the nasal consonant in the first or last word was the same as the nasal consonant in the middle word by circling a number on the answer sheet. The term ''nasal consonant'' was defined through the use of simple examples in which nasals appeared in different syllable positions and vocalic contexts. A, X, or

TABLE V. Contrasts presented to listeners in the AXB discrimination test, listed by place of articulation.

Place	Talker	Stimulus pair
m-n̪	Ym	[amːa]–[an̪ːa]
		[imːi]–[in̪ːi]
	Ys	[amːa]–[an̪ːa]
		[imːi]–[in̪ːi]
m-n	Ym	[amːa]–[anːa]
		[imːi]–[inːi]
	Ys	[amːa]–[anːa]
		[imːi]–[inːi]
m-ɳ	Ym	[amːa]–[aɳːa]
		[imːi]–[iɳːi]
	Ys	[amːa]–[aɳːa]
		[imːi]–[iɳːi]
n̪-n	Ym	[an̪ːa]–[anːa]
		[in̪ːi]–[inːi]
	Ys	[an̪ːa]–[anːa]
		[in̪ːi]–[inːi]
n̪-ɳ	Ym	[an̪ːa]–[aɳːa]
		[in̪ːi]–[iɳːi]
	Ys	[an̪ːa]–[aɳːa]
		[in̪ːi]–[iɳːi]
	Ms	[an̪ːa]–[aɳːa]
		[an̪a]–[aɳa]
		[an̪]–[aɳ]
	Mv	[an̪ːa]–[aɳːa]
		[an̪a]–[aɳa]
		[an̪]–[aɳ]
n-ɳ	Ym	[anːa]–[aɳːa]
		[inːi]–[iɳːi]
		[ana]–[aɳa]
	Ys	[anːa]–[aɳːa]
		[inːi]–[iɳːi]
		[ana]–[aɳa]
	Oc	[an̪a]–[aɳa]
	Os	[ana]–[aɳa]

B were not physically identical, so listeners made categorial matches. Listeners were instructed to ignore ''irrelevant differences'' of duration, tone, or voice quality. A familiarization set of 20 trials was presented before the AXB test, with particular trials selected to reinforce the instructions.

B. Predictions

1. Assimilation type predictions

A set of specific predictions for the discriminability of non-native contrasts was generated from the Perceptual Assimilation Model (PAM). These predictions concern only the discrimination scores elicited in the categorial AXB task. Neither PAM, nor any model of cross-language speech perception to date, can predict how non-native sounds will be identified with native categories. For the purposes of this study, each contrast for each listener group was classified as one of five assimilation types of PAM: TC, UC, CG, SC, and UU,[6] based only on the identification test results for the stimuli making up the contrast. A contrast was classified as one of these five assimilation types using the following criteria:

(1) TC: Each stimulus of the contrast had to receive different *top labels*, and both top labels had to represent 90% or more of the listener group's responses. *Top label* refers to the most popular single label selected by subjects in a listener group (i.e., the modal response choice of the listener group).
(2) UC: Each stimulus of the contrast had to receive different top labels. One of those top labels had to represent 90% or more of the listener group's responses; the other had to represent less than 90% of the listener group's responses.
(3) UU: Both stimuli had to have top labels that represented less than 90% of the listener group's responses. The top labels could be the same or different (e.g., both /n/, or one labeled as /m/ and the other labeled as /n/).
(4) SC: Both stimuli had to receive the same top label (e.g., both stimuli labeled as /n/) and each top label had to represent 90% or more of the listener group's responses. In addition, the mean category goodness ratings of the stimuli had to be statistically identical in post hoc *t* tests with an alpha level of 0.05 (see Harnsberger, 2000 for a description of the statistical analyses of the identification test results).
(5) CG: Both stimuli had to receive the same top label and each top label had to represent 90% or more of the listener group's responses. In addition, the mean category-goodness ratings of the two stimulus types had to be significantly different in post hoc *t* tests with an alpha level of 0.05.

Common to the definitions of all five of these assimilation types was the use of a stringent criterion for when a stimulus was to be said to be ''categorized'' as an exemplar of a particular native category: it had to be identified with the same label in 90% of a listener group's responses. The choice of the 90% criterion was unique to this study. The criterion was adopted to insure that category goodness, as a predictor of discriminability, was not confounded with the effect of variation in the identification of stimuli with multiple perceptual categories (a *category membership* effect, well documented in the literature on categorical perception). Lotto *et al.* (1998) and Lotto (2000) cited this confound in their critical assessment of the role of category-goodness in the NLM model.

The confounding of category goodness and category membership in predicting discrimination scores can be illustrated in the following example: two stimuli that are identified as exemplars of the same category and given equivalent ratings are classified by PAM as SC assimilations and are predicted to be very difficult to discriminate. If, however, a less stringent criterion is used for category membership, say 50% or 75%, then it is likely that some SC assimilations will be more discriminable than PAM predicts due to the well-documented effect of categorical perception on discrimination. Compare a contrast in which both stimuli (S_a and S_b) are identified 100% of the time as exemplars of category /X/ to a contrast in which S_a is identified with category /X/ 100% of the time while S_b is identified with category /X/ 50% of the time, category /Y/ 40% of the time, and category /Z/ 10% of the time. For the first contrast, all discrimination test trials involve stimuli that are perceived to belong to the same category. For the second contrast, half of the discrimi-

nation test trials involve stimuli from two different categories. Based on category membership only, the mean discrimination score over all trials would be predicted to be significantly higher for the second contrast than the first, despite the fact that they would both be labeled as SC assimilations. A high criterion such as 90% minimizes the effect of variable category membership on discriminability while hopefully still allowing many contrasts to be classified as SC or CG assimilations for the purpose of testing PAM.

One major drawback to a 90% criterion is the potential allocation of many identification patterns to the UC and UU assimilation types, for which PAM makes limited predictions. In this study, a significant portion of UC and UU assimilation types may involve ''borderline'' cases, such as an S_a-S_b contrast in which S_a and S_b are both identified with the same category in 85% of responses (a borderline SC assimilation), or an S_c-S_d contrast in which S_c and S_d are identified in 85% of responses with category /X/ and /Y/, respectively (a borderline TC assimilation). Of course, the apportioning of any continuous range into a set of categories always involves borderline cases. For this study, such borderline cases were included in the mean scores for the UC and UU assimilation types in order to avoid establishing additional criteria for which no past precedents exist, and to avoid possibly excluding large numbers of UC and UU assimilations from the analysis.

For the assimilation types defined above, the following predictions were made concerning their discriminability.

a. TC assimilations. PAM predicts that these assimilations should elicit the highest scores of any assimilation type. Thus the mean discrimination score for TC assimilations should be significantly higher than that of all other types. In addition, for the experimental paradigm used in this study, Whalen *et al.* (1997) claim that TC discrimination scores should fall between 91% and 100%.

b. UC assimilations. PAM predicts that UC assimilations should be significantly less discriminable than TC assimilations, significantly more discriminable than SC assimilations, and comparable to the best CG assimilations. No attempt was made to predict any significant difference between the mean score for UC assimilations and the mean score of some arbitrarily designated lower subset of the CG assimilations.

c. UU assimilations. The discriminability of UU assimilations is predicted to be a function of the proximity of nonnative sounds to one another (psychophysical or gestural difference) and to nearby native categories (Best, 1995). Since PAM fails to specify how these two metrics are weighted for computing a final discrimination score, it is difficult to predict a range of scores for UU assimilations or to rank order UU assimilations with respect to the other assimilation types. Given that SC assimilations should represent the most difficult ones to discriminate, a tentative prediction was made that UU assimilations would be significantly easier to discriminate than SC assimilations.

d. CG and SC assimilations. Both CG and SC assimilations concern stimuli that assimilate to the same native category. PAM predicts that CG and SC assimilations should be less significantly less discriminable than TC assimilations and that CG assimilations should be significantly more discriminable than SC assimilations. In addition, in this experimental paradigm, SC and CG assimilations were expected to elicit discrimination scores between 50% and 59% and 60% and 90%, respectively. These ranges differ from those used by Whalen *et al.* (1997), who suggested ranges of 50%–79% and 80%–90% for SC and CG assimilations, respectively. A wider range for CG assimilations, and consequently a narrower one for SC assimilations, is more in keeping with the notion that discrimination is a function of different degrees of category-goodness between the sounds that make up the contrast (Best, 1994). A ''strong'' CG assimilation would be a contrast in which one stimulus is identified as a good exemplar of a category and the other stimulus is identified as a poor exemplar of the same category. The difference in their category-goodness allows listeners to discriminate the pair, to some extent. As the difference in category-goodness decreases between two stimuli, they would be classified as weaker examples of the CG assimilation type. As category goodness difference approaches equivalence, Best (1994) states that ''both members of the non-native contrast are equally discrepant from native-category exemplars, in which case we have a SC contrast with poor discriminability'' (pp. 192). If the SC assimilation represents one extreme endpoint of the category-goodness difference continuum, then the range of discrimination scores assigned to it should be narrow relative to the CG assimilations, which have been described in more gradient terms (e.g., ''strong,'' ''weak'').

2. Other predictions

In addition to the predictions of mean discrimination scores and ranges for the five assimilation types, PAM predicts a correlation between category-goodness and discriminability given the definitions of the CG and SC assimilation types. SC and CG assimilations are said to differ in terms of the degree of category-goodness difference between the sounds that make up the non-native contrast. SC assimilations are difficult to discriminate because their constituent sounds are identified with the same category and with the same degree of category-goodness. CG assimilations are easier to discriminate because their constituent sounds differ, in some significant degree, in category-goodness. Thus differences in discriminability should positively correlate with the difference in category-goodness. The latter can be defined simply, as in Eq. (1) below:

$$D_{XY}=|X-Y|, \tag{1}$$

where X and Y are the mean category-goodness of the two stimulus types that make up a contrast, and D_{XY} is their absolute difference score. Both CG and SC assimilations could be incorporated in such an analysis, since both concern stimulus pairs that are members of the same category. A significant correlation between the discrimination scores of SC and CG assimilations and D_{XY} would constitute a measure of the extent to which an important aspect of the PAM

TABLE VI. The mean discrimination test scores for each contrast, for each listener group, with the corresponding assimilation type assigned.

Contrast			Malayalam		Marathi		Punjabi		Tamil		Oriya		Bengali		English	
Place	Talker	Stimulus pair	Type	AXB	Type	AXB	Type	AXB	Type	AXB	Type	AXB	Type	AXB	Type	AXB
m-ɲ̃	Ym	[amːa]–[aɲ̃ːa]	UC	93	UC	81	SC	85	UC	87	UC	76	SC	82	UC	93
		[imːi]–[iɲ̃ːi]	UC	99	TC	97	TC	96	UC	96	TC	96	UC	98	UC	91
	Ys	[amːa]–[aɲ̃ːa]	UC	99	TC	98	UC	87	UC	97	TC	87	TC	92	TC	95
		[imːi]–[iɲ̃ːi]	UC	98	TC	99	UC	91	UC	96	UC	98	TC	99	TC	97
m-n	Ym	[amːa]–[anːa]	UC	100	UC	99	UC	98	UC	98	TC	99	TC	100	TC	100
		[imːi]–[inːi]	UC	99	TC	100	UC	97	UC	97	TC	98	TC	99	TC	99
	Ys	[amːa]–[anːa]	UC	100	TC	100	TC	98	UC	99	TC	97	TC	100	TC	99
		[imːi]–[inːi]	UC	99	TC	100	TC	95	UC	95	TC	95	TC	97	TC	97
m-ɳ	Ym	[amːa]–[aɳːa]	TC	99	UC	97	UC	95	UC	96	UC	96	TC	99	UC	94
		[imːi]–[iɳːi]	TC	99	UC	96	UC	94	UC	96	UC	95	TC	96	UC	90
	Ys	[amːa]–[aɳːa]	TC	98	UC	100	UC	88	TC	94	UC	89	TC	95	TC	95
		[imːi]–[iɳːi]	TC	99	UC	98	UC	86	UC	93	UC	91	TC	93	UC	91
ɲ̃-n	Ym	[aɲ̃ːa]–[anːa]	UU	98	UU	96	UC	94	UU	94	UC	98	TC	99	UC	99
		[iɲ̃ːi]–[inːi]	UU	95	SC	70	UC	66	UU	70	SC	72	UC	82	UC	90
	Ys	[aɲ̃ːa]–[anːa]	UU	89	SC	56	UC	53	UU	55	SC	50	SC	55	SC	59
		[iɲ̃ːi]–[inːi]	UU	88	SC	47	UC	47	UU	53	UC	58	SC	55	SC	59
ɲ̃-ɳ	Ym	[aɲ̃ːa]–[aɳːa]	UC	97	UU	92	UC	89	UU	92	UU	90	TC	95	UU	89
		[iɲ̃ːi]–[iɳːi]	UC	99	UC	81	UC	53	UU	67	UC	63	UC	57	UU	80
	Ys	[aɲ̃ːa]–[aɳːa]	UC	81	UC	81	UU	32	UC	42	UC	42	SC	32	SC	44
		[iɲ̃ːi]–[iɳːi]	UC	96	UC	94	UU	58	UU	58	UU	71	CG	59	UC	68
	Ms	[aɲ̃ːa]–[aɳːa]	UU	87	UC	92	UU	76	UU	77	UC	76	SC	75	SC	67
		[aɲ̃a]–[aɳa]	UU	82	UC	94	UU	68	UU	53	UC	94	UC	69	SC	66
		[aɲ̃]–[aɳ]	UU	81	UC	90	UU	65	UU	61	UC	79	UC	68	SC	68
	Mv	[aɲ̃ːa]–[aɳːa]	UC	82	TC	93	UU	68	UC	65	UC	69	SC	64	SC	69
		[aɲ̃a]–[aɳa]	UC	99	TC	98	UU	93	UC	94	UC	86	UC	93	CG	87
		[aɲ̃]–[aɳ]	UC	96	TC	99	UU	98	UU	93	TC	94	UC	88	CG	97
n-ɳ	Ym	[anːa]–[aɳːa]	UC	97	UU	81	UU	76	UU	84	UC	79	SC	86	UC	80
		[inːi]–[iɳːi]	UC	95	UC	81	UU	69	UU	71	UC	68	SC	71	UC	83
		[ana]–[aɳa]	UC	95	UU	83	UU	70	SC	83	UU	81	UU	87	SC	80
	Ys	[anːa]–[aɳːa]	UC	95	UC	85	UC	54	UC	52	UC	59	SC	54	SC	63
		[inːi]–[iɳːi]	UC	97	UC	89	UC	63	UU	59	UC	72	CG	54	UC	64
		[ana]–[aɳa]	UC	97	TC	97	UU	68	UU	63	TC	88	CG	75	CG	77
	Oc	[ana]–[aɳa]	UC	99	UC	98	UU	97	UU	96	TC	97	UC	98	SC	93
	Os	[ana]–[aɳa]	UC	95	UC	99	UC	94	UU	87	TC	96	UC	94	SC	95

model accounts for the identification-discrimination relationship.

C. Results

The results of the discrimination test are given in Table VI for the seven listener groups. In this table, the mean percent correct discrimination score for each contrast is reported, with ''contrast'' defined not only in traditional feature terms (i.e., ''dental-retroflex'') but also by vocalic context, syllable type, and talker. It was necessary to report results for contrasts defined with such detail due to the significant effects of talker, vocalic context, and syllabic context on the identification test results (see Harnsberger, 2000). The discrimination test results themselves are only interpretable in light of how the listener groups identified the contrasts. The identification test results were first reported and discussed in detail by Harnsberger (2000), who studied the effect of native language phonological inventory on the identification of these non-native sounds. In this study, the identification test results are coded in terms of the Perceptual Assimilation Model in order to evaluate the model's predictions for the discrimination of non-native contrasts. A PAM assimilation type was assigned to each contrast according to the criteria given in Sec. II B. The assigned assimilation types all appear in Table VI with their corresponding contrast.

The identification test results showed that the selection of particular stimulus types and listener groups was for the most part successful in eliciting significant numbers of each assimilation type, with one exception (CG, see below). By far, the most common assimilation type elicited was UC, constituting 111 of the 238 assimilation types, or just under half, followed by UU (47), TC (47), SC (27), and CG (6). Four of the assimilation types were sufficiently represented in the dataset for the purposes of testing PAM, namely, TC, UC, UU, and SC. However, with only six instances of the CG assimilation type, it is difficult to conclude much from a comparison of the SC and CG assimilation types, or to test the correlation between category goodness and discriminability.

The mean percent correct scores on the AXB discrimination test are shown in Fig. 2 for each assimilation type, averaged over all contrasts and listener groups (error bars denote one standard error). The mean scores for the TC and CG assimilations fell within their predicted ranges (91%–100%, 60%–90%, respectively). The mean score for the SC

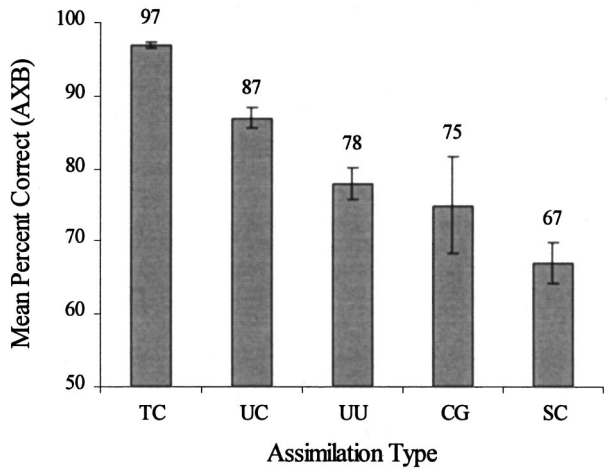


FIG. 2. The mean percent correct discrimination test scores for the five assimilation types.

assimilation type was above the range predicted (50%–59%), but within the broader range (50%–79%) predicted by Whalen *et al.* (1997). However, the range of scores within each assimilation type was greater than that predicted by PAM, particularly for the CG and SC assimilation types. Of the six CG assimilations, two had AXB discrimination scores below the predicted range and one had a mean score above. Of the 27 SC assimilations, 63% had discrimination scores above the predicted range of 50%–59%. Even using the broader range of Whalen *et al.* (1997), 26% of SC assimilations had higher than expected discrimination scores. Apparently, SC assimilations can still be relatively discriminable despite their complete assimilation to a single native category, at least in terms of the method of measuring identification in this experiment (forced-choice with goodness ratings).

The relative discriminability of the five assimilation types fulfilled some, but not all, of PAM's predictions. The discrimination scores for each assimilation type were submitted to a one-way ANOVA. Assimilation type proved to be significant [$F(4,233)=25.8$, $p<0.001$]. The results of post hoc Games–Howell tests comparing the discrimination test scores of the five assimilation types appear in Table VII. First, TC assimilations were, as predicted, significantly more discriminable than UC, UU, and SC assimilation types (but not CG assimilations; see below). UU assimilations were significantly more discriminable than SC assimilations; UC assimilations were significantly more discriminable than SC but not CG assimilations, all as predicted. Most importantly, CG and SC assimilation types were not significantly different. This is a potentially important result, as it bears on the most falsifiable portion of PAM, the relationship between category-goodness and discriminability for non-native contrasts that assimilate to one category. In addition, CG assimilations were not significantly different in discriminability from the TC assimilation type, in contrast with predictions.

TABLE VII. The significance levels of paired comparisons between assimilation types in post hoc Games–Howell tests.

	TC	UC	UU	CG
TC				
UC	$p<0.001$			
UU	$p<0.001$	$p<0.001$		
CG	n.s.	n.s.	n.s.	
SC	$p<0.001$	$p<0.001$	$p<0.05$	n.s.

Unfortunately, the significance of CG assimilation results, particularly the SC-CG comparison, is mitigated by the fact that so few examples of CG assimilations were collected in this study. If there were some way to generate more CG assimilations from the current dataset, a better test of this aspect of PAM would be possible. One method would be to relax the category membership criterion of 90%, which would result in the reclassification of many UC and UU assimilations as CG or SC. However, as argued earlier, as the membership criterion is lowered, there is an increase in the confounding of category membership with category-goodness in their effect on discrimination. One way to minimize this confound while generating more CG assimilations would be to change the criterion for including individual subject scores in the analysis. Currently, all subjects listed in Table IV had their scores included in the group averages reported by Harnsberger (2000). Some of these subjects, however, were relatively inconsistent in their labeling of some non-native contrasts. Such variable labeling contributed to the many examples of UC and UU assimilations in this dataset. Many of the UC and UU would have been classified as CG or SC assimilations if only the top labeling choice of a particular listener group was considered.

To generate more examples of SC and CG assimilations, the discrimination scores for SC, CG, UC, and UU assimilations were recalculated to represent only those subjects who consistently (in 90% of more of responses) used a single label for both stimuli of a contrast. When inconsistent labelers were excluded from the analysis, so few consistent labelers remained that in some cases the corresponding mean discrimination score represented an average over a much smaller group of listeners. In cases in which the remaining pool of labelers dropped below eight, the assimilation was excluded from the analysis. Using only such ''consistent'' subjects, a total of 64 SC and CG assimilations were generated. The reclassified CG and SC assimilations, with recalculated AXB discrimination test scores, are listed in the Appendix (see Tables AI and AII, respectively). Forty-three of the reclassified assimilations were SC assimilations and 21 were classified as CG assimilations. The mean discrimination score for the reclassified CG assimilations was 76%, versus 68% for the reclassified SC assimilations, a difference that was significant [$t(62)=2.1$, $p<0.05$]. Thus with a greater number of examples, this prediction of PAM was supported.

However, looking beyond the mean scores to the range of discrimination scores, a more varied picture emerges that is not in keeping with the predictions of PAM. Figures 3 and 4 are histograms of the discrimination scores for the reclassified SC and CG assimilations, respectively. These figures clearly show that both assimilation types spanned a wider range of possible discrimination scores (32%–95% for SC, 54%–98% for CG) than the predicted ranges (50%–59% for

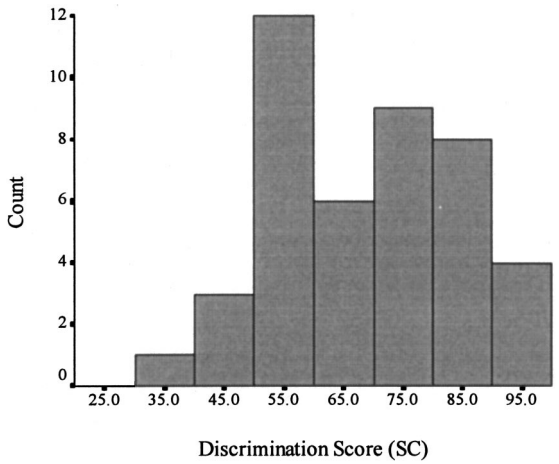


FIG. 3. The range and distribution of the discrimination test scores for the reclassified SC assimilations.

SC, 60%–90% for CG in this study; 50%–79% for SC, 80%–89% for CG proposed by Whalen *et al.*, 1997). This pattern suggests that there was a poor correspondence between the discrimination scores and the ratings data. The greatest discrepancies between the identification patterns and discrimination scores appear with the SC assimilations. Despite the fact that they were pairings of stimuli judged to be excellent exemplars of the same category, 27 out of 43 SCs elicited unexpectedly high mean discrimination scores (>59%). Of those, 12 SC assimilations elicited mean discrimination scores above even the 79% cutoff proposed by Whalen *et al.* (1997). Among the CG assimilations, three received scores below the predicted range (60%–90%), while five received scores above the predicted range.

Another testable prediction of the Perceptual Assimilation Model concerns the relationship between category-goodness and discriminability in assimilations in which both stimuli of the contrast are identified consistently as exemplars of the same native category. This relationship is an inherent part of the definition of the CG and SC assimilations, and is amenable to testing in the form of a correlation comparing the differences in category-goodness ratings of stimuli to their corresponding discrimination test scores. For this analysis, the 64 reclassified CG and SC assimilations were used, to maximize the number of observations in the analysis. Category-goodness difference and discriminability did correlate significantly, though the strength of the correlation was not great ($r=0.31$, $p<0.05$).

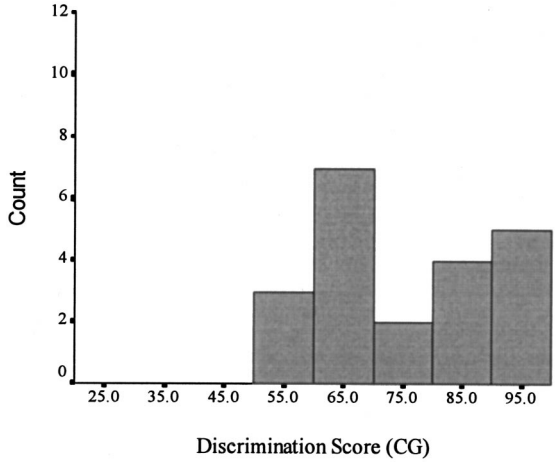


FIG. 4. The range and distribution of the discrimination test scores for the reclassified CG assimilations.

III. DISCUSSION AND CONCLUSIONS

The results obtained in this study revealed a large proportion of uncategorizable assimilations (UC and UU) and an unexpected range of discrimination scores for SC and CG assimilations. Both of these findings potentially raise problems for PAM, as well as the SLM and the NLM models. First, a large proportion of uncategorizable assimilations in this dataset indicate that PAM may not be able to account for a large proportion of the assimilations that occur outside of the laboratory. For example, PAM was unable to account for the discrimination of the UU assimilations (20% of the dataset) and the range in discrimination scores for the UC and UU assimilations (46% and 20% of the dataset, respectively). UC and UU assimilations in PAM, as stated earlier, rely on two metrics, only one of which is easy to test in traditional speech perception experiments: category-goodness and psychophysical (or gestural) difference. The latter is difficult to measure given the complex nature of speech sounds, and PAM offers no weighting of the two metrics in calculating the discriminability of two non-native stimuli. This limitation, of course, applies to the NLM and SLM models as well, and reflects limitations of the field of cross-language speech perception as a whole. To date, we lack a metric of psychophysical (or gestural) similarity that can be used to predict the similarity of a stimulus to a category, or the similarity between two stimuli (Best, 1995). Such a metric would have to integrate a number of spectral and temporal cues (or articulatory gestures) about the non-native stimuli in question. Given the frequency with which uncategorizable assimilations appear in this dataset, the need for such a metric is great.

Second, the range of discrimination scores for SC and CG assimilations did not strongly support claims concerning the relationship between category goodness and discrimination. While the predicted difference between the mean discrimination scores of the (reclassified) CG and SC assimilations was supported, a large proportion of scores fell outside of the predicted range (43% and 63%, respectively; see Figs. 3 and 4). Also, while the category-goodness difference scores did correlate significantly with the discrimination scores as predicted, the strength of the correlation was not particularly compelling in terms of PAM's prediction that category-goodness, and not psychophysical (or gestural) difference, accounts for the discriminability of SC and CG assimilations. Of course, the correlation was based on a small sample of assimilations, which may have mitigated the effect of category-goodness difference.

If category goodness, as measured in this experiment, failed in many respects to account for the discrimination test scores, what are the alternatives? Why, for instance, would two stimuli that are such similar exemplars of a single cat-

egory (as in SC assimilations or stimulus pairs near a prototype) nevertheless be, in many cases, highly discriminable, if native language experience constrains listeners' perceptual abilities? Several explanations are possible. The most obvious has been suggested previously, namely the psychophysical differences between the stimuli. Such differences may remain salient despite the assimilation process, differences that would not be reflected in the goodness ratings. As stated earlier, a complete measure of the psychophysical differences between two speech sounds is not easy to calculate, given the multiple cues that exist for speech sounds, and given the problem of integrating, or weighting, those cues to determine a single measure of psychophysical difference. However, for the purposes of testing this hypothesis, it would be worthwhile to examine SC and CG assimilations involving contrasts that appear to differ along a limited number of acoustic dimensions that are relatively easy to compare or integrate.

For the contrasts used in this study, an acoustic analysis of all of the stimuli was conducted by Harnsberger (1998). A full review of its results is beyond the scope of this paper. However, several general observations can be reported. The stimuli were differentiated by multiple acoustic cues, including the magnitude of difference in the $F2$ and $F3$ transitions,[7] duration differences in the intervocalic nasal murmurs, shifts in the nasal resonances over the course of the murmur, and differences in higher formants ($F4$, $F5$). Most of these cues have been shown in prior work to be important ones in the perception of place of articulation in nasal consonants (Malécot, 1956; Nakata, 1959; Fujimura, 1962; Larkey *et al.*, 1978; Recasens, 1983; Kurowski and Blumstein, 1984, 1987). No one cue correlated significantly with the discrimination scores. For instance, the formant transition differences for contrasts that were CG or SC assimilations yielded only an r of 0.33 ($p=0.27$) due mainly to four outlier contrasts: Malayalam talker Ym's [iɲːi]–[inːi], [amːa]–[aɲːa], and [anːa]–[aɳːa]; and Marathi talker Mv's [aɲa]–[aɳːa]. Ym's [iɲːi]–[inːi] and [anːa]–[aɳːa] were cued primarily by shifts in the nasal resonances; [aɲːa]–[aɳːa] by murmur duration difference; and [amːa]–[aɲːa] by a difference in $F5$, a rather unexpected result given the perceptual salience of high frequency information relative to low frequency (Johnson, 1997). When these 4 were excluded, the formant transition differences of the remaining 11 contrasts correlated significantly with the discrimination scores ($r=0.85$, $p=0.001$). These contrasts represented 77% of the reclassified SC and CG assimilations used to test PAM.

This acoustic analysis of a limited set of data suggests that the phonetic differences (acoustically defined in the post hoc analysis above) between stimuli that assimilate to the same native category can be apparent to listeners, even for contrasts that combined highly similar or prototypical stimuli, as defined by category-goodness ratings. If listeners maintain some sensitivity to the cues that differentiate assimilated speech sounds, then category-goodness ratings are going to be limited in capturing some of these perceived differences because ratings are one-dimensional in character. They capture a speech sound's degree of deviation from the category prototype, as opposed to manner, or direction, of deviation, in multiple acoustic dimensions. Without that additional information, it is not possible to ''locate'' two nonnative stimuli in a multidimensional perceptual space, calculate the ''distance'' between them, and arrive at a number to compare with their discrimination score. If psychophysical differences are important in predicting the discriminability of SC and CG assimilations, then category-goodness ratings alone are insufficient to map out a perceptual category. Categories instead must be modeled as multidimensional. If they are multidimensional, then linguistic communities may differ in their weighting of individual cues, as a number of researchers have claimed in the past (Terbeek, 1977; Gottfried and Beddor, 1988; Rochet, 1991).

A good candidate cross-language speech perception model may be a multidimensional version of PAM. The emphasis on category-goodness difference in the discrimination of SC and CG assimilations would need to be dropped. Instead, the discriminability of *within-category* assimilations (stimuli assimilating to the same category, subsuming PAM's SC and CG assimilation types) would be a function of their perceptual similarity. Perceptual similarity could be calculated as the sum of the two stimulus types' differences along each relevant acoustic (or gestural) dimension [see Juscyzk's (1993) WRAPSA model for an example of feature-weighting in speech perception and spoken word recognition]. Each dimension would be represented by an auditory or gestural scale (such as $F2$ in Bark or glottal aperture) that reflects a degree of abstraction that corresponds to the nature or degree of *phonologicalization* of that dimension by the linguistic community in question. This abstraction would result in cross-language differences in the phonetic detail available to listeners in each dimension. To test discriminability predictions for within-category assimilations, detailed acoustic or gestural analyses of the phonological inventories of language are required (currently, for most languages, only phonemic or allophonic descriptions of their inventories are available). The same perceptual similarity metric could also be applied to TC, UC, and UU assimilations.

A fully elaborated model of this kind would resemble the NLM model, in that both approaches assume that discrimination is a function differences between the stimuli (as opposed to category-goodness difference) and their location in perceptual space relative to category prototypes. However, to data, the NLM model has not been extended to account for non-native contrasts that fall partly or completely outside of perceptual space that is associated with particular perceptual categories, as in the UC and UU assimilations of PAM that were so common in this study. Moreover, NLM assumes a ''perceptual magnet effect'' that operates within single categories that influences the identification and discrimination of speech stimuli above and beyond the influence of categorical perception and category membership. As Lotto *et al.* (1998) and Lotto (2000) point out, it is unclear in prior work on the perceptual magnet effect if nonprototypical stimuli are truly peripheral members of a single category or if they are associated with two or more categories. Thus category-goodness may not play the role in discrimination that is en-

visioned in the NLM. A conservative approach to accounting for the discriminability of non-native contrasts in terms of categories in a multidimensional perceptual space may be to model the language-specific phonologization of dimensions solely in terms of category membership.

While modeling categories in a multidimensional space is one approach suggested by the mismatch between the ratings data and the discrimination scores of this dataset, more prosaic explanations for the mismatch can be proffered, such as the effect of task differences on cross-language speech perception (suggested in studies by Ingram and Park, 1998 and Lotto *et al.*, 1998) and the analysis techniques employed in this study. First, Ingram and Park (1998) noted a mismatch between their identification and discrimination test results. In the identification test, Japanese and Korean listeners showed labeling patterns in accordance with predictions based on their language-specific inventories of allophonic variants. In contrast, their oddball discrimination test failed to show the same language-specific patterns, leading Ingram and Park to conclude that ''identification and discrimination differed in terms of task demands on listeners,'' supporting the hypothesis that ''a phonological level of signal processing was less engaged by the...discrimination task'' (pp. 1172–1173).[8] In essence, the structure of discrimination tests allows listeners to make more precise comparisons of the stimuli, mitigating some of the language-specific constraints on perception. Thus in this study, unexpected variation in the discrimination test scores and the assimilation types (based on the identification test results) could simply reflect differences in the engagement of phonological processing due to task constraints. The fit between identification and discrimination may have improved if the discrimination task was altered to require listeners to judge stimuli at a more categorical level, such as using trials that vary in talker as well as token (e.g., the categorical oddity discrimination test used by Guion *et al.*, 2000).

This explanation seems plausible, given that in a discrimination test, listeners are able to make perceptual judgments about a stimulus relative to other stimuli in the trial, while in most identification tests, a stimulus is presented in isolation. However, what is the source of very poor or intermediate discrimination performance on non-native contrasts if not the native perceptual categories of the listener? Lotto *et al.* (1998) addressed this issue by testing an alternate method for comparing identification and discrimination data in evaluating theories of internal category structure. They cited Fry *et al.* (1962); Eimas (1963); Thompson and Hollien (1970); and Nearey (1989) in arguing that the categorization of vowel stimuli is context-dependent. That is, the identity of a vowel stimulus can be shifted to another category when it was paired with another vowel stimulus, as in a discrimination test. Lotto *et al.* (1998) tested this possibility by running two experiments, one a replication of Iverson and Kuhl (1995) in which synthetic [i] stimuli were presented in isolation for the identification test, and another in which the same set of stimulus pairs was presented in both an identification and an AX discrimination test with only the instructions and response labels differing. As anticipated, they found a robust effect for stimulus context. For instance, the percent [i] identification for the prototypical [i] stimulus varied from 48%–84% depending on the neighboring vowel in a trial. This context effect could account for many of the discrepancies observed in Ingram and Park (1998), as well as in this study, between identification in isolation and discrimination in context. Thus in this study, PAM's predictions for the ranges of scores for CG and SC assimilations may have been upheld if the listeners in this study were asked to rate stimuli in the same context as they heard them in the discrimination test.

Finally, in this study, the comparison of assimilations to discrimination data was made at the level of the seven listener groups. For each group and contrast, the labeling and rating responses of the individual subjects were averaged together before being submitted for correlation analysis with the mean discrimination test scores. A stronger correlation between category-goodness difference and discrimination may have emerged if assimilations were assigned to individual subjects' identification of non-native stimuli. For this study, an analysis of the individual results was not possible given the small number of judgments (four) made by listeners for each stimulus type in the identification test (Harnsberger, 2000). Such an approach was taken by Lively and Pisoni (1997) in their test of the perceptual magnet effect using synthetic tokens of /i/. In their first experiment, Lively and Pisoni (1997) administered a ratings task in which subjects rated on a seven-point scale a set of synthetic stimuli modeled after those used by Kuhl (1991). They found that subjects varied greatly in their rating of synthetic stimuli in a ''prototype'' set versus a ''nonprototype'' set, indicating that subjects can differ significantly in the location and structure of their /i/ category in this task. In a second experiment, Lively and Pisoni (1997) administered both a ratings task and a same–different discrimination task in which the assumed prototypical /i/ stimuli (and their corresponding distribution of less prototypical neighbors) was calibrated to that of the individual subject based on the ratings task. Interestingly, this individual calibration technique did not result in a robust confirmation of the perceptual magnet effect. However, it does serve as a useful example of how best to evaluate some predictions of cross-language speech perception models, such as the identification-discrimination relationship.

To address the issues surrounding the role of category-goodness and discriminability, future studies should take into account the methodological issues discussed above, while directly addressing the issue of multidimensional perceptual categories. Such studies necessitate the use of synthetic stimuli in order to determine the degree of phonetic detail in, or the weighting of, individual perceptual dimensions for a given category for a given listener group or individual subject. In addition, a diverse range of listener groups should be used, spanning multiple types of phonological inventories to ensure the generalizability of the findings to other groups and contrasts. Such future experiments should allow us to successfully model the identification-discrimination relationship in cross-language speech perception, as well as the identification of non-native sounds with native categories.

ACKNOWLEDGMENTS

I would like to thank Patrice Beddor, who advised me throughout the design and execution of this study, and Catherine Best for her many helpful comments at early and late stages of this work. I would also like to thank John Kingston, John Logan, Linda Polk, and one anonymous reviewer for their comments on earlier drafts of this work and on its companion piece, Harnsberger (2000). Questions to the author should be addressed to James Harnsberger, Speech Research Laboratory, Department of Psychology, Indiana University, Bloomington, IN 47405.

APPENDIX

TABLE AI. ''Consistent'' Category-Goodness assimilations. PAM (Old)=assimilation type assigned to the contrast based on the identification test results of all of the subjects; AXB Scores: (Old)=scores averaged over all subjects and (New)=scores averaged over the consistent subjects.

Contrast			Listener group	PAM	AXB scores	
Place	Talker	Stimulus pair		(Old)	New	Old
m-n̪	YM	[amːa]–[an̪ːa]	English	UC	91	93
			Marathi	UC	78	81
n̪-ɳ	YS	[in̪ːi]–[iɳːi]	Bengali	CG	59	59
			English	UC	69	68
			Punjabi	UC	56	58
	MS	[an̪a]–[aɳa]	Bengali	UC	68	69
			English	SC	66	66
		[an̪]–[aɳ]	Bengali	UC	69	68
	MV	[an̪ːa]–[aɳːa]	English	SC	69	69
		[an̪a]–[aɳa]	Bengali	UC	92	93
			English	CG	87	87
		[an̪]–[aɳ]	Bengali	UC	88	97
			English	CG	97	97
n-ɳ	YM	[anːa]–[aɳːa]	English	UC	79	80
			Oriya	UC	80	79
		[ana]–[aɳa]	English	SC	80	80
	YS	[inːi]–[iɳːi]	Bengali	CG	54	54
			English	UC	66	64
			Punjabi	UC	66	63
	OC	[ana]–[aɳa]	Bengali	UC	98	98
	OS	[ana]–[aɳa]	Bengali	UC	94	94

TABLE AII. ''Consistent'' Single-Category assimilations.

Contrast			Listener group	PAM	AXB scores	
Place	Talker	Stimulus pair		(Old)	New	Old
m-n̪	YM	[amːa]–[an̪ːa]	Bengali	SC	82	82
			Oriya	UC	75	76
			Punjabi	SC	84	84
			Tamil	UC	88	87
n̪-n	YM	[in̪ːi]–[inːi]	Bengali	UC	81	82
			English	UC	91	90
			Marathi	SC	70	70
			Oriya	SC	72	72
			Punjabi	UC	69	66
	YS	[an̪ːa]–[anːa]	Bengali	SC	56	56
			English	SC	59	59
			Marathi	SC	56	56
			Oriya	SC	50	50
			Punjabi	UC	54	53
		[in̪ːi]–[inːi]	Bengali	SC	55	55
			English	SC	59	59
			Marathi	SC	47	47
			Oriya	UC	57	58
			Punjabi	UC	47	47
n̪-ɳ	YM	[in̪ːi]–[iɳːi]	Bengali	UC	59	57
			Punjabi	UC	55	53
	YS	[an̪ːa]–[aɳːa]	Bengali	SC	32	32
			English	SC	44	44

TABLE AII. (*Continued*.)

Contrast			Listener group	PAM (Old)	AXB scores	
Place	Talker	Stimulus pair			New	Old
	MS	[an̪ːa]–[anˑa]	Bengali	SC	75	75
			English	SC	67	67
		[an̪]–[anˑ]	English	SC	68	68
	MV	[an̪ːa]–[anˑa]	Bengali	SC	65	65
		[an̪a]–[anˑa]	Punjabi	UU	95	93
n-ɳ	YM	[anːa]–[anˑa]	Bengali	SC	86	86
			Punjabi	UU	78	76
		[inːi]–[iɳˑi]	Bengali	SC	71	71
			English	UC	83	83
			Oriya	UC	60	68
			Punjabi	UU	71	69
		[ana]–[anˑa]	Bengali	UU	82	87
			Tamil	SC	83	83
	YS	[anːa]–[anˑa]	Bengali	SC	54	54
			English	SC	63	63
			Punjabi	UC	51	54
		[ana]–[anˑa]	Bengali	CG	75	75
			English	CG	77	77
	OC	[ana]–[anˑa]	English	SC	93	93
	OS	[ana]–[anˑa]	English	SC	95	95

[1]Complete measures of psychophysical difference between stimuli are difficult to calculate since natural speech sounds are complex signals differing along multiple acoustic or gestural dimensions. Choosing which dimensions to combine in a measure of psychophysical difference for a given pair of speech sounds is not a straightforward matter.

[2]Psychophysical difference is a measure easily obtained for the stimuli Kuhl and colleagues have used, namely synthetic vowels varying along two dimensions, $F1$ and $F2$, scaled in mels or Bark.

[3]Hallé, Best, and Levitt (1999) do consider this factor in their recent study of French perception of English /1/–/r/.

[4]For more information concerning the selection of stimulus materials, please consult Harnsberger (2000).

[5]The results were reported in terms of so many (34) contrasts because, in the identification test results, the labeling and rating results were heavily influenced by talker, vocalic context, and syllabic context (Harnsberger, 2000). Thus averaging together the discrimination scores of, for instance, all of the dental-retroflex contrasts produced by talkers Ym, Ys, Ms, and Mv, in [aNːa], [iNːi], [aNa], and [aN] contexts (N=nasal consonant), would have involved averaging together the discrimination scores of contrasts that potentially differed significantly from one another.

[6]No non-native contrast could have been classified as an NA assimilation type because the identification test was closed set, with no ''Nonspeech'' response choice offered.

[7]''Transitions'' here refer to changes in formants from the midpoint of the vowel to the vowel/nasal boundary. For entry into the correlation analysis, the static formant measures were converted to Bark. Then, the absolute value of the Bark difference scores were calculated and entered in the analysis.

[8]It should be noted that Ingram and Park (1998) did use a relatively short interstimulus interval (0.5 s) that may have allowed listeners to perceive acoustic-phonetic detail that would not be available at a phonological level of processing. Access to such phonetic detail could account for the mismatch between their identification and discrimination results.

Abramson, A. S., and Lisker, L. (**1970**). ''Discriminability along the voicing continuum: Cross-language tests,'' in *Proceedings of the Sixth International Congress of Phonetic Sciences* (Academia, Prague), pp. 569–573.

Best, C. T. (**1994**). ''The emergence of native language phonological influences in infants: A perceptual assimilation model,'' in *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, edited by J. Goodman and H. C. Nusbaum (MIT Press, Cambridge), pp. 167–224.

Best, C. T. (**1995**). ''A direct realist view of cross-language speech perception,'' in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York Press, Baltimore), pp. 171–203.

Best, C. T. (**1996**). ''Assimilation of non-native vowel contrasts to the American English vowel system,'' J. Acoust. Soc. Am. **99**, 2602.

Best, C. T., McRoberts, G. W., and Sithole, N. M. (**1988**). ''Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants,'' J. Exp. Psychol. **14**, 345–360.

Burnham, D. K. (**1986**). ''Developmental loss of speech perception: Exposure to and experience with a first language,'' Appl. Psycholinguist. **7**, 207–240.

Eimas, P. D. (**1963**). ''The relation between identification and discrimination along speech and nonspeech continua,'' Language and Speech **6**, 206–217.

Flege, J. E. (**1995**). ''Second language speech learning: Theory, findings, and problems,'' in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York Press, Baltimore), pp. 233–277.

Fry, D. B., Abramson, A. S., Eimas, P. D., and Liberman, A. M. (**1962**). ''The identification and discrimination of synthetic vowels,'' Lang. Speech **5**, 171–189.

Fujimura, O. (**1962**). ''Analysis of nasal consonants,'' J. Acoust. Soc. Am. **34**, 1865–1875.

Gottfried, T. L., and Beddor, P. S. (**1988**). ''Perception of temporal and spectral information in French vowels,'' Lang. Speech **31**, 57–75.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., and Pruitt, J. C. (**2000**). ''An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants,'' J. Acoust. Soc. Am. **107**, 2711–2724.

Halle, P. A., Best, C. T., and Levitt, A. (**1999**). ''Phonetic vs phonological influences on French listeners' perception of American English approximants,'' J. Phonetics **27**, 281–306.

Harnsberger, J. D. (**1998**). ''The perception of non-native nasal contrasts: A cross-linguistic perspective,'' Doctoral Dissertation, The University of Michigan.

Harnsberger, J. D. (**2000**). ''A cross-language study of the identification of non-native nasal consonants varying in place of articulation,'' J. Acoust. Soc. Am. **108**, 764–783.

Henly, E., and Sheldon, A. (**1986**). ''Duration and context effects on the perception of English /r/ and /l/: A comparison Cantonese and Japanese speakers,'' Lang. Learning **36**, 505–521.

Ingram, J., and Park, S-G. (**1998**). ''Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners,'' J. Acoust. Soc. Am. **103**, 1161–1174.

Iverson, P., and Kuhl, P. K. (**1995**). ''Mapping the perceptual magnet effect

for speech using signal detection theory and multidimensional scaling,'' J. Acoust. Soc. Am. **97**, 553–562.
Iverson, P., and Kuhl, P. K. (**1996**). ''Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/,'' J. Acoust. Soc. Am. **99**, 1130–1140.
Johnson, K. (**1997**). *Acoustic and Auditory Phonetics* (Blackwell, Cambridge).
Jusczyk, P. W. (**1993**). ''From general to language-specific capacities: The WRAPSA Model of how speech perception develops,'' J. Phonetics **21**, 3–28.
Kuhl, P. K. (**1991**). ''Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not,'' Percept. Psychophys. **50**, 93–107.
Kurowski, K. M., and Blumstein, S. (**1984**). ''Perceptual integration of the murmur and formant transitions for place of articulation in nasal consonants,'' J. Acoust. Soc. Am. **76**, 383–390.
Kurowski, K. M., and Blumstein, S. (**1987**). ''Acoustic properties for place of articulation in nasal consonants,'' J. Acoust. Soc. Am. **81**, 1917–1927.
Larkey, L., Wald, J., and Strange, W. (**1978**). ''Perception of synthetic nasal consonants in initial and final position,'' Percept. Psychophys. **23**, 299–312.
Lively, S. E., and Pisoni, D. B. (**1997**). ''On prototypes and perceptual categories: A critical assessment of the perceptual magnet effect in speech perception,'' J. Exp. Psychol. **23**, 1665–1679.
Lotto, A. J. (**2000**). ''Reply to 'An analytical error invalidates the depolarization of the perceptual magnet effect,' '' J. Acoust. Soc. Am. **107**, 3578–3580.
Lotto, A. J., Kluender, K. R., and Holt, L. L. (**1998**). ''Depolarizing the perceptual magnet effect,'' J. Acoust. Soc. Am. **103**, 3648–3655.
Malécot, A. (**1956**). ''Acoustic cues for nasal consonants: An experimental study involving tape-splicing technique,'' Lang. **32**, 274–284.
Miyawaki, K., Strange, W., Verbrugge, R. R., Liberman, A. M., Jenkins, J. J., and Fujimura, O. (**1975**). ''An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English,'' Percept. Psychophys. **18**, 331–365.
Nakata, K. (**1959**). ''Synthesis and perception of nasal consonants,'' J. Acoust. Soc. Am. **31**, 661–666.
Nearey, T. M. (**1989**). ''Static, dynamic, and relational properties in vowel perception,'' J. Acoust. Soc. Am. **85**, 2088–2113.
Polka, L. (**1991**). ''Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions,'' J. Acoust. Soc. Am. **89**, 2961–2977.
Polka, L. (**1992**). ''Characterizing the influence of native language experience on adult speech perception,'' Percept. Psychophys. **52**, 37–52.
Pruitt, J. S. (**1995**). ''The perception of Hindi dental and retroflex stop consonants by native speakers of Japanese and American English,'' Doctoral Dissertation, University of South Florida, Tampa.
Recasens, D. (**1983**). ''Place cues for nasal consonants with special reference to Catalan,'' J. Acoust. Soc. Am. **73**, 1346–1353.
Rochet, B. L. (**1991**). ''Perception of the high vowel continuum: A cross-language study,'' in *Proceedings of the Twelfth International Congress of Phonetic Sciences* (University of Provence, Aix-en-Provence, France), Vol. 4, pp. 94–97.
Sheldon, A., and Strange, W. (**1982**). ''The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception,'' App. Psycholinguist. **3**, 243–261.
Sussman, J. E., and Gekas, B. (**1997**). ''Phonetic category structure of [I]: Extent, best exemplars, and organization,'' J. Speech Hear. Res. **40**, 1406–1424.
Terbeek, D. (**1977**). ''A cross-language multidimensional scaling study of vowel perception,'' Doctoral Dissertation, University of California, Los Angeles.
Thompson, C. L., and Hollien, H. (**1970**). ''Some contextual effects on the perception of synthetic vowels,'' Lang. Speech **13**, 1–13.
Werker, J. F., Gilbert, J. H. V., Humphrey, K., and Tees, R. C. (**1981**). ''Developmental aspects of cross-language speech perception,'' Child Dev. **52**, 349–353.
Whalen, D. H., Best, C. T., and Irwin, J. R. (**1997**). ''Lexical effects in the perception and production of American English /p/ allophones,'' J. Phonetics **25**, 501–528.

Target spectral, dynamic spectral, and duration cues in infant perception of German vowels

Ocke-Schwen Bohn[a)]
Engelsk Institut, Aarhus Universitet, DK-8000 Aarhus C, Denmark

Linda Polka[b)]
School of Communication Sciences and Disorders, McGill University, Montreal, Canada

(Received 21 June 2000; revised 16 April 2001; accepted 30 April 2001)

Previous studies of vowel perception have shown that adult speakers of American English and of North German identify native vowels by exploiting at least three types of acoustic information contained in consonant–vowel–consonant (CVC) syllables: target spectral information reflecting the articulatory target of the vowel, dynamic spectral information reflecting CV- and -VC coarticulation, and duration information. The present study examined the contribution of each of these three types of information to vowel perception in prelingual infants and adults using a discrimination task. Experiment 1 examined German adults' discrimination of four German vowel contrasts (/i/–/e/, /e/–/ɪ/, /ɪ/–/ɛ/, /o/–/ʊ/), originally produced in /dVt/ syllables, in eight experimental conditions in which the type of vowel information was manipulated. Experiment 2 examined German-learning infants' discrimination of the same vowel contrasts using a comparable procedure. The results show that German adults and German-learning infants appear able to use either dynamic spectral information or target spectral information to discriminate contrasting vowels. With respect to duration information, the removal of this cue selectively affected the discriminability of two of the vowel contrasts for adults. However, for infants, removal of contrastive duration information had a larger effect on the discrimination of all contrasts tested. [DOI: 10.1121/1.1380415]

PACS numbers: 43.71.An, 43.71.Ft, 43.71.Hw [KRK]

I. INTRODUCTION

Vowels produced in natural coarticulated consonant–vowel–consonant (CVC) syllables contain at least two potential perceptual cues to vowel identity: (a) formant minima or maxima near the vowel midpoint which reflect the articulatory target of the vowel (i.e., target spectral information), and (b) spectrally dynamic portions of the vowel onsets and offsets which reflect the coarticulation of the surrounding consonants with the vowel, i.e., dynamic spectral information (Jenkins, 1987). In addition, the duration of the vowel portion may provide information on vowel identity in languages with contrastive duration, and patterns of vowel inherent spectral change may be perceptually exploited in languages with diphthongized vowels (Andruski and Nearey, 1992; Nearey, 1989).

Several studies by Strange and her collaborators examined in detail the acoustic sources of information in the perception of native coarticulated vowels by adult American English (AE) listeners (e.g., Strange, 1987; Strange, 1989a; Jenkins *et al.*, 1994) and adult German listeners (Strange and Bohn, 1998). These studies showed that vowel centers (VCs), which consist only of the syllabic nuclei with target information, are not perceived more accurately than silent center (SC) syllables, which consist only of the dynamic portions of the syllable onsets and offsets in their appropriate durational relationship. For adult native listeners of AE and of German, vowel identity is maintained very well in SCs even though the vocalic nucleus with information on formant targets is silenced in SCs. The only major difference between the study examining German listeners (Strange and Bohn, 1998) and the studies examining AE listeners (e.g., Strange, 1989b) was that German listeners were more adversely affected by the removal of duration information than AE listeners. This is not surprising given the fact that duration differences between German vowel contrasts are much larger than between AE vowel contrasts (Strange and Bohn, 1998; Bohn and Flege, 1992).

These findings and others on the importance of dynamic spectral information for vowel perception provide support for Strange's dynamic specification theory (DST), which states that vowels are specified by dynamic information defined over syllable onsets and offsets (Strange, 1989b). According to DST, the dynamic information reflects each vowel's characteristic opening and closing phases in their appropriate durational relationship and style of movement of the vocal tract. Research motivated by DST has pursued the question of whether this information is specific to the vowel, or whether it is speaker-specific or specific to the place and voicing characteristic of the surrounding consonants. For instance, Jenkins *et al.* (1994) examined the perception of ''hybrid'' SCs in which the onsets and offsets were originally produced by two speakers (a female and a male), and Jenkins *et al.* (1999) examined the perception of mixed-consonant SCs in which the onsets and offsets were cross-matched from syllables originally produced in different consonant contexts. The results of these studies strongly

a)Electronic mail: engosb@hum.au.dk
b)Electronic mail: lpolka@po-box.mcgill.ca

suggest that coarticulated vowels are specified by styles of movement that are invariant across consonant contexts and across different speakers. Because this approach attempts to explain perception of vowels as they typically occur in natural speech (i.e., coarticulated in syllables), DST presents a compelling alternative to models of vowel perception built on the assumption that vowel identity is specified primarily by target information.

Much previous research on infant vowel perception has implicitly been based on the assumption that the perception of steady-state vowels is sufficiently representative of vowel perception in general. For instance, research by Trehub (1973, 1976), Swoboda *et al.* (1976), and by Kuhl and her colleagues (Kuhl, 1979; Kuhl *et al.*, 1992) employed synthetic steady-state vowel tokens in which the only acoustic information for vowel identity consisted of target formant frequencies. However, more recent studies by Polka and her colleagues (Polka and Werker, 1994; Polka and Bohn, 1996) have used naturally produced CVC syllables which provide listeners with the full set of acoustic information characteristic of natural coarticulated vowels. An important question that has not been addressed concerns the type of acoustic information that infants exploit in vowel discrimination. Is the implicit assumption of much previous infant vowel perception research valid that vowels are sufficiently specified by their targets? Or, do infants make use of dynamic spectral information to the same extent as adult listeners? One could also ask whether either type of acoustic information takes developmental priority such that, for instance, infants first establish vowel identity on the basis of target spectral information, and only later learn to exploit dynamic spectral information. Finally, what is the role of duration information?

Previous research does not provide much of an indication as to how these questions will be answered. The relative importance of different acoustic cues which signal vowel contrasts has not yet been studied with prelingual subjects. Murphy *et al.* (1989) compared adults and 3-year-old children on their ability to identify synthetic tokens of the English vowels /æ/ and /ʌ/ in unedited and in SC /bVd/ syllables and reported than ''many 3-year-olds can identify vowels correctly in the absence of the full spectral properties of steady-state formants.'' In addition, several studies that examined the contribution of various cues to stop-consonant identification in CV syllables by young children (3–5 year olds) reported that this age group possesses ''skills for decoding information that is provided by a speaker's coarticulatory behavior'' (Parnell and Amerman, 1978, p. 694; see also Nittrouer, 1992; Ohde and Haley, 1997). If one adds to this Nittrouer's (1992) conclusion that 3-year-olds are more sensitive than adults to dynamic aspects of production, and less sensitive to static aspects, one might expect that the ability to exploit dynamic spectral cues for vowel perception either develops before the age of 3 years, or is already present in prelingual infants. However, if the suggestion made by Elliott *et al.* (1979) is correct that children require more acoustic information than adults to identify a speech stimulus, one might expect that prelingual infants would perform less well on the discrimination of edited syllables (like SCs and VCs) than on unedited syllables.

Concerning the importance of contrastive duration information for vowel perception, previous research on adult listeners' perception of non-native vowels suggests that listeners rely on duration cues to differentiate vowel contrasts if they have not had sufficient experience with the spectral cues which contrast spectrally similar vowel categories (Bohn and Flege, 1990; Bohn, 1995; Flege *et al.*, 1997). One could expect that infants likewise lack sufficient experience with the spectral cues that differentiate spectrally similar vowel categories, which might lead them to rely on duration cues to a larger extent than adult native listeners would.

The present study addresses the questions raised above by testing two hypotheses. First, we hypothesize that target spectral information does not have a privileged status in infant vowel perception. Infants are similar to adults in that they can exploit dynamic spectral information to discriminate tokens from different vowel categories. Second, we hypothesize that infants rely on duration differences to a greater extent than adult native listeners do in their attempt to discriminate spectrally similar vowels. The hypotheses were examined in a series of experiments testing German-learning infants' discrimination of naturally produced German /dVt/ syllables which were digitally modified to manipulate the availability of the three types of acoustic information (target spectral, dynamic spectral, and durational).

The present study examined the discrimination of the four German vowel contrasts, /i/–/e/, /e/–/ɪ/, /ɪ/–/ɛ/, and /o/–/ʊ/, produced in a /dVt/ context, as has been used in previous research with German adults. The test contrasts were chosen primarily because they were relatively confusable in the Strange and Bohn (1998) study, and because they differ in the kind of acoustic cues that could be used for their perceptual differentiation. Of the four test contrasts, two consist of a tense (long) and a lax (short) vowel which are spectrally similar but differ greatly in duration (/e/–/ɪ/, /o/–/ʊ/). The other two contrasts consist of spectrally similar vowel pairs which differ less in duration, i.e., (/i/–/e/) with two tense (long) vowels, and (/ɪ/–/ɛ/) with two lax (short) vowels.

Discrimination of the vowel contrasts was examined in two experiments. Experiment 1 examined native adult German listeners' discrimination of unmodified and digitally modified German /dVt/ syllables. This experiment provided a measure of comparison for the infant study (experiment 2). Previous studies of the contribution of the various sources of acoustic information in CVC syllables to adult vowel perception have exclusively employed identification experiments. Experiment 1 used the same stimuli as experiment 2, and the procedure was as similar as possible to the one used in the infant discrimination experiment. Experiment 2, which is the major focus of this study, examined the discrimination of modified and unmodified German /dVt/ syllables by German-learning infants.

II. GENERAL METHODS

A. Stimuli

The speech materials were recorded by the same talker described in Strange and Bohn (1998); his North German (Kiel) dialect matches that of the subjects in the present ex-

TABLE I. Average and range (in parentheses) in durations of onset, center, and offset portions of four tokens each of the six German test vowels. Absolute duration and VOT (in ms) and proportions of total syllable duration are given.

	dit	det	dɪt	dɛt	dot	dʊt
Onset:	36.6	37.7	35.1	34.9	37.7	36.8
Absolute	(35–38)	(37–38)	(29–38)	(33–38)	(35–43)	(36–37)
VOT	12.2	14.1	13.6	11.5	12.8	13.2
	(10–12)	(13–15)	(13–14)	(9–13)	(11–14)	(12–15)
% syllable	24.6	18.6	29.1	26.4	15.9	26.2
	(23–28)	(18–19)	(24–32)	(25–27)	(14–18)	(25–28)
Center:	88.3	141.0	62.3	68.9	158.9	73.5
Absolute	(71–99)	(132–152)	(56–69)	(58–77)	(144–165)	(65–88)
% syllable	59.3	69.6	51.6	52.2	67.0	52.4
	(53–62)	(68–71)	(48–57)	(48–57)	(66–69)	(49–58)
Offset:	24.0	23.7	23.3	28.2	40.6	30.0
Absolute	(23–25)	(23–24)	(23–24)	(23–30)	(39–47)	(26–34)
% syllable	16.1	11.7	19.3	21.3	17.1	21.4
	(15–19)	(11–12)	(18–20)	(17–25)	(16–19)	(18–25)
Syllable duration	148.9	202.4	120.7	132.0	237.1	140.3
	(133–152)	(193–214)	(117–124)	(121–142)	(216–247)	(131–152)

periments. Six tokens each of the German vowels /i/, /ɪ, /e/, /ɛ/, /ʊ, /o/, /ɑ/ were produced in /dVt/ syllables in citation form and recorded on a TEAC DAT recorder in an IAC sound chamber with an AKG microphone. The vowel /ɑ/ was included only to form the /i/–/ɑ/ contrast for use in the conditioning stage of the infant testing. The syllables were digitized on a DEC Vaxstation II/GPX minicomputer at a 20-kHz sampling rate with 12-bit resolution after low-pass filtering at 9.8 kHz and 6-dB/oct pre-emphasis. No modifications of input intensity were made across the utterances within each block. From the digitized waveforms, measurements of syllable duration, voice onset time (VOT), and fundamental frequency ($F0$) were used to select four instances each of the six vowels. Tokens were chosen which had equally short VOT values and stable $F0$.

Seven modified stimulus conditions were generated from the digitized syllables using waveform-editing techniques. First, each /dVt/ syllable was divided into onset, center, and offset portions. The criteria used here are similar to those employed by Strange (1989a) and Jenkins *et al.* (1994). Specifically, the onset included the release burst, VOT, and the first three pitch periods of the vowel portion. Wideband spectrograms (Kay Elemetrics digitial sonograph) of each syllable were visually inspected to ascertain that the formants were still in transition from the initial /d/ at the end of the onset portion, and that the onset and offset portions included the most dynamic part of the formant trajectories.[1] The offset portion contained a minimum of three pitch periods, the closure duration for the final /t/, and the release burst.[2] The center portions were defined as the interval between the end of the onset and the beginning of the offset. Visual inspection of the spectrograms confirmed that the center portions included the formant targets (minima or maxima) and the less dynamic portions of the formant trajectories.

Table I presents the average and range in absolute duration's and proportion of total syllable duration for onset, center, and offset portions of the German vowels /i/, /ɪ/, /e/, /ɛ/, /ʊ/, and /o/. As expected, /dɪt/–/dɛt/ is the only contrast in which the absolute durations of the three portions are very similar across the two vowels. For the three other test contrasts (/dit/–/det/, /det/–/dɪt/, /dot/–/dʊt/), absolute durations of the center and/or offset portions differ considerably.

Table II presents the results of the spectral measurements. Using the Computerized Speech Research Environment, a 25.6-ms Hamming window was centered at the end of the onset portion, at the temporal midpoint of the vowel interval, and at the beginning of the offset portion. The frequencies of the first three formants were established with an LPC algorithm (14 poles). These formant frequency measures are in good agreement with the results of Strange and Bohn (1998) for their /dVt/ tokens in showing that $F1$, $F2$, and $F3$ are changing throughout the center portions of almost all vowels. It is important to note that Strange and Bohn's analyses of citation-form /hVt/ and sentence-context /dVt/ tokens produced by the same speaker as in the present study suggested that formant movement patterns in /dVt/ syllables reflected coarticulatory effects with the surrounding alveolar consonants rather than diphthongization intrinsic to the vowels. Thus, vowel inherent spectral change is not an accessible cue for the perception of vowels produced by this speaker.

In addition to the unmodified original /dVt/ syllables (labeled FULL), the following seven syllable conditions were generated for each contrast: (1) Silent center (SC) tokens were generated by attenuating to silence the center portion of each FULL syllable, leaving onset and offset portions in their original temporal position. The onset and offset portions included the major part of the transitions. (2) Vowel center (VC) tokens were the converse of SC syllables. VC tokens were generated by attenuating to silence the onset and offset portions, leaving the centers with the formant maxima or minima. (3) Initial (INI) tokens were generated by silencing both center and offset portions. (4) Final (FIN) tokens were generated by silencing both onset and center portions. (5) FULL syllables with neutralized duration (FND) were

TABLE II. Mean formant frequencies and ranges of four tokens each of the German vowels /i, e, ɪ, ɛ, o, ʊ/, produced in /dVt/ syllables, measured at the end of the onset portion, the middle of the center portion (Mid), and the beginning of the offset portion.

		Onset			Mid			Offset		
Vowel		*F*1	*F*2	*F*3	*F*1	*F*2	*F*3	*F*1	*F*2	*F*3
/i/	Mean	253	2109	2734	255	2279	2658	244	2299	2660
	Range	252–253	2073–2141	2667–2817	252–258	2249–2315	2602–2717	239–254	2287–2318	2650–2670
/e/	Mean	304	1975	2529	299	2206	2507	294	2183	2538
	Range	290–337	1911–2068	2511–2529	291–309	2168–2223	2438–2533	288–297	2163–2215	2493–2567
/ɪ/	Mean	312	1928	2458	345	2088	2606	344	2010	2427
	Range	281–339	1902–1982	2423–2510	317–367	1934–2162	2386–2843	336–352	1963–2096	2335–2565
/ɛ/	Mean	405	1704	2359	494	1697	2230	493	1669	2266
	Range	386–425	1637–1807	2327–2388	467–517	1547–1741	2139–2360	462–512	1561–1723	2202–2348
/o/	Mean	356	1115	2027	352	682	2194	329	683	2223
	Range	348–361	1042–1250	1997–2047	228–377	634–737	2158–2269	321–341	634–727	2168–2303
/ʊ/	Mean	365	1155	2069	389	924	2060	426	1013	2073
	Range	359–372	1029–1219	2000–2106	370–404	896–940	2012–2140	409–445	980–1035	2030–2101

generated by iterating or deleting complete pitch periods of the eight tokens for each vowel contrast (four per vowel) so that the resulting duration was nearly equal to mean duration of the eight tokens. Pitch periods were iterated or deleted in a way that did not abruptly change the formant movement throughout the center portion. (6) Silent centers with neutralized duration (SND) were generated by adjusting the silent interval to the mean duration of the eight tokens for each contrast. (7) Vowel centers with neutralized duration (CND) were generated by iterating or deleting complete pitch periods so that the resulting duration was nearly equal to the mean of the VC durations for the eight tokens in each contrast.

The 236 stimuli created to fulfill these eight conditions were copied onto DAT tape and redigitized for presentation on an IBM format computer. To ensure that subjects would not discriminate on the basis of loudness differences between contrasting tokens, minor adjustments in amplitude were made to the redigitized tokens. Using the VU meter of an audiometer, which is designed to integrate amplitude over time like a human ear, the amplitude of each token of a set of four tokens (per vowel, per condition) was adjusted by the same amount so that the most intense of the four tokens peaked at the same level on the VU meter as the most intense token of the contrasting set. This approach retained some variability with regard to loudness within each token set, but naive listeners reported no detectable differences in loudness between syllables within each contrast.

B. Procedures and equipment

The equipment was identical and the test procedure was similar for the infant and adult experiment. Subjects were tested in the same IAC sound chamber and seated (on a parent's lap for the infants) across a small table from an experimenter (E1). The loudspeaker and an array of four visual reinforcers, located behind a smoked Plexiglas panel, were placed to the right of the subject. E1 and the parent listened to vocal music over headphones to prevent them from influencing the subject's behavior. A second experimenter (E2) located outside the sound chamber observed the subject through a one-way window and operated the computer with the stimulus presentation program. The stimuli were presented online from an IBM format computer via a DT2801 D/A board and were routed through a Yamaha AX-350 amplifier for delivery via a Cyrus 780 loudspeaker. The peak amplitude of the stimuli varied between 67 and 69 dBA when measured at approximate ear level of the subject. Computer software controlled the stimulus delivery, activation of the reinforcers, and trial selection, and recorded hits, misses, correct rejections, and false alarms.

Subjects were tested in age-appropriate versions of the change/no change procedure, which has been used in previous studies involving a variety of age groups including infants and adults (Werker *et al.*, 1997). The subject was instructed (adults) or conditioned (infants) to provide a response when she/he detects a change in an ongoing series of stimuli, and to suppress a response when no such change occurs. To test infants, a headturn response was used; a hand-signal response was used to test adults. In the present experiments a stream of stimuli was played from a loudspeaker with a 1500-ms ISI. At random intervals the background syllables were changed to foreground syllables for a 4.8-s interval during which three tokens were presented which were either drawn from the same category as the background syllable (control trial) or from a contrasting category (change trial). The maximum number of consecutive control trials was limited to three, and the change from background to foreground was initiated by E2, who could not hear whether she initiated a change or a control trial. When E2 observed a headturn (infants) or a hand signal (adults), she pressed a button connected to the PC which controlled feedback for correct signals, which consisted of the illumination of the display of toys above the loudspeaker. No feedback was provided for false alarms. For infants, correct responses in change trials were also reinforced by verbal praise from E1. For both groups we implemented the change/no change procedure as a category change paradigm in which the background and the target consist of multiple tokens of each syllable type (played in random order).

With the infants, E1 engaged the infant's visual attention with several toys as needed between trials. Adults were engaged in a simple distractor game (*Memory*)[3] with E1; they were instructed to attend to the game and to the stimuli, and to raise their hand when they detected a change. The distrac-

tor game, which required visual attention, was included in an attempt to create comparable conditions for the infant and the adult subjects.

Pilot testing was conducted with both adults and infants. This testing revealed that the SC syllables remained cohesive when presented in the repetition format with a long (1500 ms) ISI used in the change/no change paradigm; the individual portions of the syllable did not stream apart or group in an incoherent fashion.[4] Adults were able to identify the intended vowel in the SC syllables presented in this format and the high levels of discrimination accuracy obtained for adult and infants in the SC conditions are further evidence that the silent gaps in these syllables were not problematic. (For more details see General Discussion, Sec. V.)

III. EXPERIMENT 1

Experiment 1 was an attempt to replicate the general findings reported by Strange and Bohn (1998) using a simple discrimination task. This was necessary because all previous research using the SC paradigm has examined performance using identification rather than discrimination tasks. Thus, there are no adult discrimination data to compare with infant perception assessed via discrimination. The findings can be compared to the results obtained with infants in Experiment 2. Based on Strange and Bohn (1998), we expect that vowel contrasts are equally discriminable on the basis of dynamic spectral information and target spectral information, whereas vowel discrimination will be less accurate when only vowel onset or vowel offset information is available. Specifically, we expect discrimination to be equally accurate in the SC and VC conditions, whereas performance in the INI and FIN conditions will be significantly lower. We would also expect that removal of contrastive duration information will reduce vowel discrimination accuracy for German adults. Specifically, we expect that discrimination will be poorer in the FND than the FULL condition, in SND than the SC condition, and in the CND than the VC condition. The effect of removing vowel duration information may also have a larger effect for the tense–lax contrasts (/e/–/ɪ/, /o/–/ʊ/), given that vowel duration differences are larger for such contrasts (Strange and Bohn, 1998).

A. Method

1. Subjects

Eighty adults (50 females; 30 males) participated as unpaid volunteers. All subjects were native speakers of North German with limited exposure to languages other than German (i.e., less than 1 year in a foreign language environment) and with no history of hearing loss. Participants were randomly assigned to eight experimental groups (ten subjects/group). The mean age of subjects was 25.3 years (s.d.=4.9).

2. Design

Adults were tested individually in a single session. Groups of ten subjects each were assigned to one of eight listening conditions defined by stimulus type: FULL, SC, VC, INI, FIN, FND, SND, and CND. Each subject was tested on each of the four contrasts with testing blocked by contrast. Within each group, every subject was tested in a different order of contrasts. For each contrast the vowel that served as background was counterbalanced within each group.

We decided to test each adult in one condition across all four contrasts. A disadvantage of this design is that it is not optimal for collecting data with infants. As described below, it is clearly necessary to test each infant on the same contrast and to have each baby show the ability to perform the task with unedited syllables before testing them on the edited syllables. We decided to implement a different protocol with the adults because our pilot testing indicated that varying the contrast had the advantage of making this very easy task slightly more challenging and clearly helped the adults to stay engaged in the task.

3. Procedure

In each group, subjects were first briefly familiarized with the task by presenting them with the contrast /i/–/ɑ/ in their stimulus condition (FULL, SC, etc.). For each contrast, the testing stage was initiated after four correct responses to change trials during a training stage. Twenty-five test trials (15 changes and 10 controls) were presented for each contrast. However, testing was terminated after 15 trials (9 changes, 6 controls) if the subject made no errors.

B. Results

Percent correct scores (''hits'' + ''correct rejections''/change+control trials) were analyzed in an analysis of variance (ANOVA) with vowel contrast (/i/–/e/, /e/–/ɪ/, /ɪ/–/ɛ/, /o/–/ʊ/) as a within-subjects factor and condition (FULL, SC, VC, INI, FIN, FND, SND, CND) as a between-subjects factor.[5] The ANOVA revealed significant main effects of condition [$F(7,72)=38.160$; $p<0.01$] and vowel contrast [$F(3,216)=17.753$; $p<0.01$], as well as a significant condition × vowel contrast interaction [$F(21,216)=7.190$; $p<0.01$]. The condition main effects were explored in pairwise *post hoc* tests (Tukey, $\alpha=0.05$), and the interaction was explored in separate one-way ANOVAs for each of the four vowel contrasts.

Condition main effect: Fig. 1 illustrates the results for the eight stimulus conditions collapsed across the four vowel contrast. *Post hoc* tests revealed that, as expected, discrimination did not differ significantly in the conditions FULL, SC, and VC, whereas it was significantly lower in the FIN and INI conditions. Discrimination accuracy was significantly lower in the SND than in the SC condition, as expected, but there were no significant differences between the FULL and FND conditions or between the VC and CND conditions, contrary to our predictions.

Condition × vowel contrast interaction: Four separate one-way ANOVAs examining the effect of condition for each of the vowel contrasts revealed in each case a significant main effect of condition ($p<0.01$). Table III gives the percent correct scores in each of the eight conditions for each vowel contrast separately. Pairwise *post hoc* tests exploring

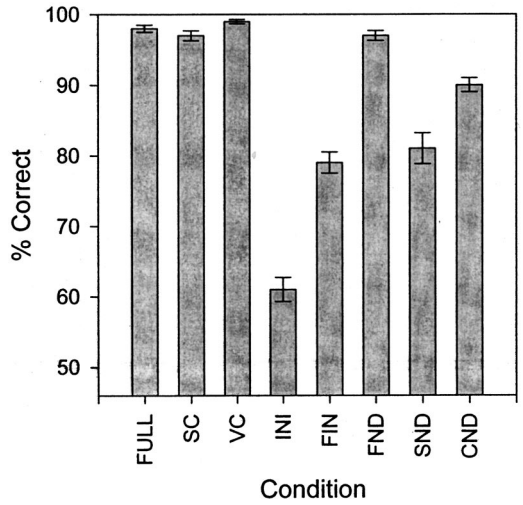


FIG. 1. Overall percent-correct scores in German adults' discrimination of four German vowel contrasts in eight conditions. Error bars represent ±1 standard error.

the main effect of condition for each contrast separately revealed that, for every contrast, the subjects' ability to discriminate stimuli in the FULL, SC, and VC conditions did not differ significantly.

For the /i/–/e/ and /e/–/ɪ/ contrasts, discrimination was significantly better in the FULL, SC, and VC conditions than in the INI and FIN conditions, and thus follows the pattern of the main effect of condition. However, for the /ɪ/–/ɛ/ contrast, discrimination was significantly better in the FULL, SC, VC, and INI conditions than in the FIN condition. For the /o/–/ʊ/ contrast, discrimination was significantly better in the FULL, SC, VC, and FIN conditions than in the INI condition.

With respect to effect of neutralizing vowel duration information, the *post hoc* tests show that there were no differences in discrimination for the FULL vs FND in any of the four contrasts consistent with the main effect. Discrimination accuracy in the CND condition was significantly lower than VC only for the /e/–/ɪ/ contrast. Discrimination accuracy was lower in the SND condition than in the SC condition for two of the four contrasts (/i/–/e/ and /e/–/ɪ/).

TABLE III. Mean percent-correct scores in German adults' discrimination of four German vowel contrasts in eight conditions (standard deviations in parentheses).

	Condition							
Contrast	FULL	SC	VC	INI	FIN	FND	SND	CND
/i/–/e/	98.0	96.5	99.6	72.0	74.8	98.0	80.0	96.0
	(5.1)	(4.7)	(1.3)	(13.7)	(18.2)	(3.4)	(21.2)	(6.5)
/e/–/ɪ/	98.4	95.6	99.6	45.2	69.2	90.8	68.0	79.2
	(3.4)	(9.7)	(1.3)	(10.3)	(21.8)	(10.3)	(20.3)	(10.5)
/ɪ/–/ɛ/	96.8	95.6	97.2	84.4	77.6	94.8	94.4	94.4
	(2.5)	(8.1)	(6.0)	(12.6)	(22.4)	(8.9)	(5.1)	(6.0)
/o/–/ʊ/	96.8	96.8	100	43.2	91.7	98.5	90.4	98.4
	(6.5)	(7.0)	(0)	(5.3)	(15.9)	(3.1)	(11.2)	(2.8)
Mean	97.5	96.1	99.1	61.2	78.3	95.5	83.2	92.0
	(4.5)	(7.3)	(3.2)	(20.7)	(20.7)	(7.6)	(19.7)	(10.2)

C. Discussion

The pattern of results for the discrimination of confusable German vowel contrasts was similar to that reported previously for the identification of native vowels by adult AE listeners (Strange, 1989a) and by adult German listeners (Strange and Bohn, 1998), even though the stimulus materials and the experimental tasks were different. In particular, German adults' discrimination of the four German vowel contrasts was highly accurate in the SC condition, i.e., when only dynamic information specified over onsets and offsets together was available. This replicates the results of previous studies employing the SC paradigm, which have shown that target spectral information is not necessary for an accurate perception of vowels; rather, trajectory information specified over syllable onsets and offsets is a sufficient source of information for vowel identity.

The overall finding that INIs and FINs were discriminated less accurately than SCs is also consistent with Strange's dynamic specification theory (Strange, 1989b), which states that vowel identity is specified by dynamic information defined over syllable onsets and offsets. Even when listeners were engaged in the very simple task of detecting a vowel change, onsets alone (INIs) or offsets alone (FINs) failed to support vowel discrimination as well as either target information or the trajectory information defined over syllable onsets and offsets in their appropriate durational relationships (SCs). However, there were two exceptions to the overall finding that INIs and FINs were discriminated less accurately than SCs: For the /ɪ/–/ɛ/ contrast, INIs were as highly discriminable as SCs, and for the /o/–/ʊ/ contrast, FINs were as highly discriminable as SCs. These exceptions are probably due to the large spectral difference between /ɪ/–/ɛ/ in the INI condition, where /ɪ/ and /ɛ/ differ more in both $F1$ and $F2$ than the other vowel contrasts (see Table II), and to the large duration contrast for the /o/–/ʊ/ contrast in the FIN condition, where /o/–/ʊ/ is the only contrast with nonoverlapping durations.

With respect to the role of duration information, our results differ somewhat from the Strange and Bohn (1998) study, which reported that identification accuracy for SC and VC syllables was reduced when contrastive duration information was removed. This led us to expect that removal of duration information would reduce discrimination for all four contrasts and perhaps have a stronger effect for the two tense/lax pairs. In the present study, however, effects of neutralization were not uniform but instead were evident only for the /e/–/ɪ/ contrast with respect to the VC vs CND conditions and for the /e/–/ɪ/ and /i/–/e/ contrasts with respect to SC vs SND conditions. It is not clear why these specific neutralization effects were observed. However, the duration differences may be more perceptually important for these two contrasts because the $F1$ and $F2$ differences at onset, midpoint, and offset for these two contrasts are smaller compared to the other two vowel contrasts.

The finding that the role of duration information is selective rather than uniform may be due to either differences in the procedure or in the stimuli between the Strange and Bohn study and the present study. Discrimination in the change/no-change procedure is a very simple task compared

to forced-choice identification (with 14 response alternatives), as in the Strange and Bohn study. Also, duration information was removed from the test syllables in different ways in the two studies. For example, CNDs in the present study had minimally eight pitch periods, whereas Strange and Bohn removed duration information from their VCs by deleting all but the middle four pitch periods of the original stimuli. Nevertheless, the fact that we did find some effects of vowel neutralization indicates that the stimuli and task used here were sensitive enough to show that vowel duration affects the perception of some vowel contrasts by adults.

IV. EXPERIMENT 2

Experiment 2 examined the acoustic bases for the perceptual differentiation of native vowels by German infants. As reviewed above, no previous study has examined the relative importance of target spectral, dynamic spectral, and duration information in prelingual infants' vowel perception. To the extent that infants make use of the same acoustic information presented in coarticulated syllables as do adults, we expect that discrimination will be equally accurate when infants are presented with target spectral or dynamic spectral information, whereas vowel discrimination will be less accurate when only vowel onset or vowel offset information is available. Specifically, performance in the SC and VC conditions should not differ, and should be significantly better than performance in either the INI or the FIN condition. In addition, if infants are sensitive to any decrease in cue redundancy, we would expect that discrimination will be best supported when all naturally present cues are available. Specifically, discrimination in the FULL condition should be better than in any modified condition.

If infants assign the same importance to duration as native adult listeners, they will show the same selective response to the removal of this information as observed in experiment 1. However, recall that previous studies indicate that non-native adults exhibit a greater reliance on duration cues to differentiate vowel contrasts when their experience with the spectral cues has been limited. Given their less extensive experience with native vowels, we might expect that infants also rely on duration cues to a larger extent than do native adults. If so, we expect removal of contrastive duration information to reduce discrimination accuracy for all four contrasts regardless of what other information is present. Specifically, discrimination accuracy should be lower in the FND than the FULL condition, in the SND than the SC condition, and in the VC than in the CND condition.

A. Method

1. Subjects

Seventy infants (42 females, 28 males) served as subjects. The infants were between the age of 6 and 12 months (mean age: 8.2 months, s.d.=1.4).[6] All were healthy, full-term infants with no history of ear infections (by parental report). All infants were being raised in monolingual (North) German-speaking families in Kiel, Germany. Two-hundred thirty-one additional infants were tested but failed to provide complete data for the following reasons: fussiness (17), failed to reach criterion on first visit as described below (186),[7] scheduling problems (25), and equipment failure (3). Ten infants were randomly assigned to each of seven experimental groups (as explained below). There were no significant age differences among the seven groups [$F(6,63) = 1.821$; $p = 0.1093$].

2. Procedure

Discrimination of one vowel contrast in one stimulus session (e.g., /i/–/e/ as SC) was tested in a single session. Each infant test session consisted of two conditioning stages and a testing stage. In the first conditioning stage, the infant was trained to response to a change from either a /dit/ or a /dɑt/ to the syllable that will be presented as foreground during the testing stage. Infants tested on the front vowel contrast (/i/–/e/, /e/–/ɪ/, /ɪ/–/ɛ/) heard /dɑt/ as the background in this stage; infants tested on the back vowel contrast /o/–/ʊ/ heard /dit/ as the background syllable. In the second conditioning stage, the /dɑt/ or /dit/ background was replaced with the other test contrast syllable while the foreground syllable remained the same. For example, infants tested on /i/–/e/ in the SC condition were first conditioned to respond to change from an SC syllable with /ɑ/ to an SC syllable with /i/, and then to a change from an SC syllable with /i/ to an SC syllable with /e/. We used two conditioning stages for two reasons. First, starting conditioning with a more distinct vowel pair clearly facilitated infants' learning of the contingency between vowel change and the reinforcer. Second, by switching the background, but not the foreground, we wanted to encourage infants to adopt a ''listen for the foreground syllable'' strategy in the task.[8] Given that all infants began testing with FULL unedited syllables in which vowel quality is clearly the most salient difference between the syllables, we hoped that this approach would serve to focus their attention on the vowel quality of the foreground syllable. Although we cannot know whether this is the one or only strategy that infants used in this task, we wanted to structure the task to encourage infants to make the same sort of judgment across the various test conditions.

Only a single exemplar of each syllable type was used in the conditioning stages. During the conditioning stages, all trials were change trials. Reinforcers were activated immediately or shortly following a vowel change in the initial conditioning trials to establish the contingency and then only after a headturn response to shape an anticipatory headturn. The transition from the first to the second conditioning stage, and from the second conditioning stage to the testing stage, was made after maximally ten change trials or after minimally three consecutive correct anticipatory head turns.

During the testing stage, all four tokens of each syllable type were presented as the background and foreground stimuli. E2 initiated trials when the infant was in a ''state of readiness'' (facing E1, not fussing, etc.). E2 was blind to the trial type (change or control) and pushed a button when she observed a head turn during the trial interval. A retraining protocol implemented during the testing stage provided up to six retraining trials (change trials that are automatically reinforced if no head turn occurs). Performance on retraining trials was excluded from data analysis. Twenty-five trials (15 changes and 10 controls) were presented in the testing stage.

3. Design

Ten infants were randomly assigned to each of seven experimental groups. In each group, infants were tested on one vowel contrast (/i/–/e/, /e/–/ɪ/, /ɪ/–/ɛ/, or /o/–/ʊ/) in four stimulus conditions. The vowel that served as the background (or foreground) during testing for each contrast was counterbalanced within each group. Each condition was tested in a single test session conducted on different days; all testing was completed within a 3-week period. Each subject in every group was first tested in FULL condition on day 1. Only infants who discriminated the unmodified FULL syllables (criterion: 7/8 consecutive correct trials and >60% correct responses[9]) returned for further testing. Any infant who failed to meet this initial criterion was considered either unwilling or unable to perform the headturn task and was excluded from the study; this criterion is reasonable given data showing that infants this age and younger can discriminate native vowel contrasts (Polka and Werker, 1994; Polka and Bohn, 1996).

Infants in groups 1–4 were tested on one of the four contrasts in the FULL, SC, VC, and INI or FIN conditions; the SC and VC conditions were tested on days 2 and 3 (order counterbalanced within each group) and either the INI or the FIN condition were presented on day 4.[10] Infants in groups 5–7 were tested on one of three contrasts in the FULL, FND, SND, and CND conditions;[11] the FND condition was presented on day 2 and the SND and CND conditions on days 3 and 4 (order counterbalanced within the groups). Infants in groups 1–4 and infants in groups 5–7 were tested by two different teams of assistants.

The order of test conditions was not fully counterbalanced within each group. Instead, all infants were tested on the unedited syllables first and the conditions that were likely to be more difficult for infants (determined in pilot testing) were presented last (i.e., INI and FIN in groups 1–4 and SND and CND in groups 5–7). This is not optimal from a pure experimental design perspective. However, such fixed sequencing of conditions is desirable when testing infants in the headturn procedure because this procedure involves operant conditioning and thus if reinforcement is not available (e.g., if discrimination becomes difficult) infants will stop performing in the task. Therefore, when using this procedure in a within-subjects design, researchers often test infants on the easier conditions first in order to elicit optimal performance from babies (e.g., Werker and Lalonde, 1988).

B. Results

Figure 2 shows the percent-correct scores for the eight stimulus conditions, averaged across the four vowel contrasts, and Table IV gives the percent-correct scores for the eight conditions for each of the four contrasts (see footnote 5). As a first step, percent-correct scores in the FULL condition were compared in a one-way ANOVA to determine whether the four contrasts were equally discriminable in the unmodified condition, and to determine whether results for the FULL condition differed depending on which of the two teams of assistants tested infants.[12] This analysis failed to show any significant differences in discrimination accuracy among the seven groups [$F(6,63)=0.595$; $p=0.733$], indicating that across the seven groups, infants attained a comparable level of performance in the FULL condition on day 1 regardless of which contrast they were tested on and which team performed the testing.

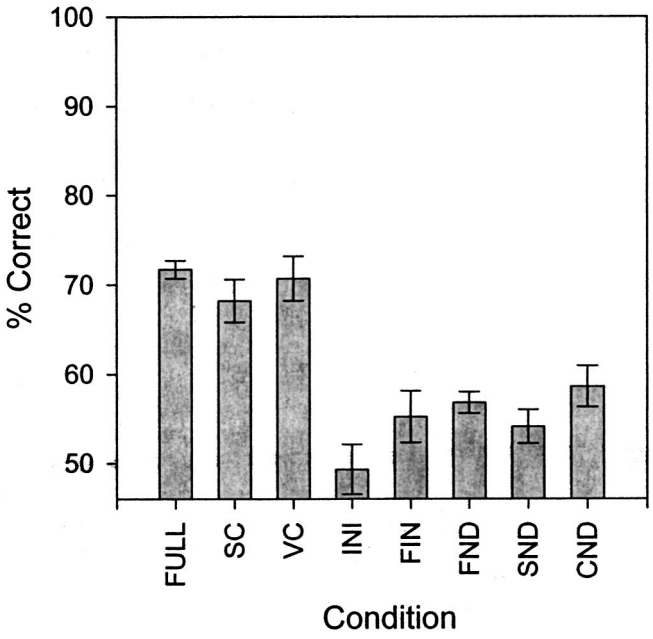


FIG. 2. Mean percent-correct scores in German infants' discrimination of four German vowel contrasts in eight conditions. Error bars represent ±1 standard error.

The discrimination scores for groups 1–4 were first analyzed in a mixed-model ANOVA with vowel contrast (four levels) as a between-subjects factor and condition (FULL, SC, VC) as a within-subjects factor. This analysis failed to reveal a main effect of condition [$F(2,72)=1.375$; $p=0.2595$]; the condition by vowel contrast interaction was also nonsignificant [$F(6,72)=1.857$; $p=0.1001$]. However, the effect of vowel contrast was significant [$F(3,36)=3.418$, $p=0.0274$]. *Post hoc* tests (Tukey, $\alpha=0.05$) revealed that the /o/–/ʊ/ contrast was easier to discriminate than the /i/–/e/ contrast.

Next, two one-way ANOVAs were conducted.[13] Recall that half of the infants tested on FULL, SC, and VC condi-

TABLE IV. Mean percent-correct scores in German infants' discrimination of four German vowel contrasts in eight conditions (standard deviations in parentheses).

	Condition							
Contrast	FULL	SC	VC	INI	FIN	FND	SND	CND
/i/–/e/	70.5	64.9	59.5	58.0	46.7	55.1	55.1	64.8
	(6.9)	(17.9)	(15.6)	(6.9)	(8.6)	(8.8)	(10.6)	(10.7)
/e/–/ɪ/	72.8	66.8	76.9	42.2	57.5	58.1	55.4	52.3
	(13.4)	(13.5)	(19.1)	(7.0)	(12.2)	(4.8)	(9.5)	(10.7)
/ɪ/–/ɛ/	71.4	61.2	70.1	53.2	50.4	…	…	…
	(6.9)	(11.8)	(11.4)	(20.8)	(11.5)	…	…	…
/o/–/ʊ/	75.2	79.8	76.4	47.2	68.0	57.2	51.8	58.7
	(9.6)	(12.6)	(12.6)	(7.2)	(12.3)	(6.3)	(11.2)	(13.2)
Mean	72.5	68.2	70.7	50.2	55.7	56.8	54.1	58.6
	(9.4)	(15.3)	(16.0)	(12.6)	(13.2)	(6.7)	(10.3)	(12.4)

tions were tested in the INI condition and the other half were tested in the FIN condition. The first ANOVA compared performance in the FULL, SC, VC, and INI conditions in those infants who were tested in the INI condition. The second ANOVA compared performance in the FULL, SC, VC, and FIN conditions in those infants who were tested in the FIN condition. Vowel contrast was not included in these analyses because only five infants each were tested on INI (or FIN) for each contrast. Both ANOVAs revealed a significant main effect of condition [ANOVA 1: $F(3,57)=15.567$; $p<0.001$; ANOVA 2: $F(3,60)=9.588$; $p<0.001$]. In both ANOVAs, *post hoc* tests (Tukey, $\alpha=0.05$) showed that performance was comparable across FULL, SC, and VC conditions, but was lower in the INI (ANOVA 1) and the FIN condition (ANOVA 2).

Next, discrimination scores for groups 5–7 were analyzed in a mixed model condition (4)×vowel contrast (3) ANOVA to compare infants tested in the FULL, FND, SND, and CND conditions. The main effect of condition was highly significant [$F(3,81)=18.418$; $p<0.001$]. *Post hoc* tests (Tukey, $\alpha=0.5$) indicated that discrimination was significantly better in the FULL condition than in the FND, SND, and CND conditions, which did not differ significantly from one another. The interaction of vowel contrast and condition was nonsignificant [$F(6,81)=1.692$; $p=0.1335$], indicating that the neutralization of duration information in the conditions FND, SND, and CND reduced discriminability for all three vowel contrasts.

The effect of neutralizing duration information on the discriminability of the stimulus types SC and VC was examined in separate ANOVAs. One analysis examined the effects of neutralization with respect to SC syllables in an ANOVA with neutralization (SC vs SND) and vowel contrast (3) as between-subjects factors. A second analysis examined the effects of neutralization with respect to VC syllables with neutralization (VC vs CND) and vowel contrast as between-subjects factors. Both ANOVAs revealed significant main effects of neutralization, for SC vs SND: $F(1,54)=24.353$; $p<0.001$, and for VC vs CND: $F(1,54)=11.711$; $p=0.0013$. The main effects of vowel contrast were nonsignificant in both ANOVAs. The neutralization ×vowel contrast interaction was marginally significant in the ANOVA with SC syllables [$F(2,54)=3.066$, $p=0.0548$], and reached significance in the ANOVA with VC syllables [$F(2,54)=6.275$; $p=0.0035$]. One-way ANOVAs examining the effect of neutralization separately for each contrast revealed that vowel duration neutralization did not affect the discriminability of /i/–/e/ in either SC [$F(1,19)=2.197$; $p=0.1556$] or VC syllables [$F(1,19)=0.779$; $p=0.3892$]. However, this manipulation produced a significant decrease in discriminability for both syllable types in the other two contrasts. For /e/–/ɪ/ comparing SC vs SND: [$F(1,19)=4.770$; $p=0.0425$] and comparing VC vs CND: [$F(1,19)=12.625$; $p=0.0023$]; for /o/–/ʊ/ comparing SC vs SND: [$F(1,19)=27.425$; $p<0.0001$]; and comparing VC vs CND: [$F(1,19)=9.379$; $p=0.0068$].

Discrimination performance was clearly poorer in those conditions in which contrastive duration information was neutralized (FND, SND, CND). Single mean t-tests were conducted comparing performance in each neutralized condition (collapsed across contrasts) to chance (i.e., 50% correct). These analyses confirmed that performance in these conditions was significantly above chance [for FND $t(29)=5.47$, $p<0.005$; for SND $t(29)=2.14$, $p<0.025$; for CND $t(29)=3.73$, $p<0.005$]. This indicates that neutralizing duration information removed some, but not all of the information that infants can use to discriminate vowel contrasts.

C. Discussion

The present findings show that infants' discrimination of confusable German vowel contrasts is similar in several respects to that reported previously for the perception of native vowels by adult listeners. Like German adults, German infants' discrimination of four German vowel contrasts in the SC condition was comparable to the FULL and the VC conditions, indicating that infants can use dynamic spectral information, the full range of information provided by unmodified (FULL) syllables, or the target spectral information provided by VC syllables. Infants were also similar to adult listeners in that onsets alone (INIs) or offsets alone (FINs) failed to support vowel discrimination as well as either target information or the trajectory information defined over syllable onsets and offsets in their appropriate durational relationships. Thus, infants appear able to make use of either target or dynamic spectral information to discriminate vowel contrasts. Moreover, with respect to cue redundancy, infant perception of coarticulated vowels is not adversely affected by removal of either dynamic or target information, i.e., discrimination in the VC and SC conditions was comparable to the FULL unmodified syllables.

Unlike native German adults, German infants' discrimination of the four vowel contrasts was strongly affected by removal of duration information. For the infants, neutralization of vowel duration produced a uniform reduction in discrimination accuracy across all four contrasts regardless of what other information was retained in the FND, SND, and CND conditions. The only exception to this overall pattern was that removal of vowel duration information failed to affect discrimination of /i/–/e/ in either the VC vs CND or the SC vs SND conditions. With respect to cue redundancy, infant perception of coarticulated vowels was adversely affected by removal of the duration information in that discrimination accuracy was better in the FULL condition compared to every condition in which duration information was removed, i.e., FND, SND, CND, INI, and FIN. It is important to note that when duration information was removed in the FND, SND, and CND conditions, infant performance did not drop to chance levels and that discrimination accuracy was comparable in the SND and CND conditions. Thus, despite the clear importance of duration cues, infants were still able to discriminate the contrasts when only target or dynamic information was retained and neither type of cue was favored.

V. GENERAL DISCUSSION

The primary aim of this study was to examine which of the acoustic cues that are present in naturally produced CVC

syllables infants may use to discriminate coarticulated native vowels. Specifically, we tested the hypotheses that (1) prelingual infants are similar to adults in that either dynamic spectral information or target spectral information is used in vowel perception, and (2) that infants rely on duration cues to a greater extent than adult native listeners do in their attempt to perceptually differentiate spectrally similar vowels. Both hypotheses were supported in two experiments that used the same stimulus materials and comparable test procedures to assess adults' and infants' discrimination of four confusable German vowel contrasts. Experiment 1, which examined native German adults, replicated previous identification studies (Strange, 1989a; Strange and Bohn, 1998) by showing that dynamic and target spectral cues are equally efficient in supporting vowel discrimination in adults, and that syllable onsets alone or offsets alone support vowel discrimination less well than target information or dynamic information defined over syllable onsets and offsets in their appropriate durational relationship. The results of experiment 1 differed somewhat from previous adult studies with respect to the importance of duration cues. The finding that the adults in experiment 1 were selectively (rather than globally) affected by the removal of duration cues is likely due to differences concerning the stimulus materials and procedures between the present and previous studies.

Experiment 2 showed that prelingual infants are like adults in that dynamic or target spectral cues can provide information that is comparable to unmodified syllables in supporting discrimination of coarticulated vowels. Experiment 2 showed that infants are also like adults in that onsets alone or offsets alone support discrimination much less well than information defined over syllable onsets and offsets. However, infants differed from adults in that their performance was more adversely and globally affected by the removal of contrastive duration information.

Concerning the role of target spectral and dynamic spectral information in infant vowel perception, previous studies clearly show that prelingual infants can discriminate vowels on the basis of target information (Trehub, 1976; Kuhl, 1979) or when the full complex of cues is available (Polka and Werker, 1994; Polka and Bohn, 1996). The present findings clearly show that infants discriminate vowel contrasts equally well if presented with unmodified syllables, silent centers, or vowel centers, whereas accuracy is poorer when only the onsets or offset portions are presented. Although we structured the discrimination task to encourage infants to respond on the basis of perceived vowel quality, we cannot be certain that this is what infants were actually doing. For example, one could argue that with SC syllables infants are responding on the basis of differences in the initial /d/ and final /t/ articulations that result from their coarticulation with different vowels and that babies may be unable or simply choose not to respond to the quality of the intended vowel as adults do. Although this is a logical possibility, we suspect that this explanation is unlikely to be correct for several reasons. First, our data and our observations clearly show that infants were able to rapidly adjust to the changes in the information present in the syllables across the FULL, SC, and VC conditions. This was evident when infants were tested on different days but also when we presented very abrupt online changes in stimulus conditions during our pilot testing. In pilot testing we shifted infants from testing on FULL syllables to testing with SC syllables (with the same contrasting vowels) within a single testing stage. Infant performance was not disrupted by this abrupt shift; instead, infants continued to respond appropriately in the task. It would seem difficult for infants to maintain a steady performance in the face of these stimulus changes if they were changing the basis on which they decide whether or not a change has occurred. Moreover, the cues that infants must attend to if they are simply detecting differences in the initial /d/ or final /t/ articulations are very subtle in our stimuli because the contrasting vowels within each pair are spectrally close to one another. Finally, if one assumes that infants are responding to subtle differences in the initial /d/ and final /t/ articulations for the SC stimuli, then one would have to explain why the infants do not do as well on either the INI or the FIN conditions as they do on the SC syllables.

For these reasons we would argue that infants in this study were most likely responding on the basis of perceived vowel quality, at least in the FULL, SC, and VC conditions. As such, it appears reasonable to interpret the present data as evidence that spectral dynamics provide a robust and viable source of information supporting vowel perception in prelingual infants. Accordingly, we propose that dynamic spectral information is exploited for vowel perception in the early stages of language development rather than emerging as a by-product of highly sophisticated language processing skills. Nevertheless, further studies designed to determine whether infants can form vowel equivalence classes for SC syllables produced in different consonantal contexts or by multiple talkers are needed to further substantiate our claims. If infants can do this it would provide stronger support for our conclusions and would also show that the dynamic portions of syllable onsets and offsets contain information that specifies vowel identity that is context- and speaker-independent for infants. If, as we propose, infants can use dynamic spectral information efficiently, then they may also be able to achieve perceptual constancy for vowel categories by extracting context- and speaker-independent information from the dynamic cues that occur in natural coarticulated syllables. The present study also does not address the question of whether the ability to exploit dynamic spectral information for vowel discrimination is present at birth. Because language experience has been shown to affect vowel perception by 6 months of age (Kuhl *et al.*, 1992), studies with even younger infants are needed to address the question of the origins of the ability to use dynamic spectral information for vowel discrimination.

Concerning the role of contrastive duration information, previous research by Flege and Bohn (Bohn, 1995; Flege *et al.*, 1997) suggests that vowel duration is a default cue that adults use to differentiate non-native vowel contrasts cued by small spectral differences that are irrelevant in their native vowel inventory. Infants, like non-native adults, may also lack the extensive experience required to distinguish small spectral differences that code some of their native vowel contrasts. For this reason, we hypothesized that German in-

fants would rely on vowel duration differences to a greater extent than native German adults. The present study further supported this hypothesis. Infants in experiment 2 showed a very robust and uniform reduction in discrimination accuracy when vowel duration was neutralized, regardless of what other information was retained. In contrast, the overall effect of vowel duration neutralization was very weak in the adults and was evident for only two of the four vowel contrasts in only two of the three conditions in which vowel duration was neutralized. Thus, there were clear differences in the extent to which German adults and German infants relied on duration.

Three interpretations of these differences in infant and adult perception appear viable. First, as suggested by Bohn (1995), vowel duration may be the more salient cue for inexperienced listeners when they confront spectrally similar vowels that differ in duration. As such, this may be a language-independent pattern that infants display before they have developed a language-specific interpretative scheme. Alternatively, German infants' sensitivity to vowel duration may reflect an effect of language experience. Adult studies show that vowel duration is a perceptually important cue for distinguishing vowels in North German (Strange and Bohn, 1998). Moreover, the finding that diphthongization is absent in this variety of German may further enhance the salience of vowel duration information (Strange and Bohn, 1998). Thus, the strong bias toward vowel duration in the present study could mean that these German infants have already discovered that vowel duration is a salient feature of vowels in their language. These explanations are not mutually exclusive, i.e., the default bias toward duration may help German infants discover the importance of this feature in their language. Further research is needed to clarify the meaning of the adult and infant response to this duration information. Specifically, data comparing vowel discrimination across infants exposed to languages in which the status of vowel duration differs will be informative.

With respect to duration, a third possible interpretation of our findings must also be considered in light of the stimuli and task employed in the present study. As mentioned above, it was not mandatory for infants to identify or attend to the vowel in the test syllables and thus infants could respond to durational differences in the syllables without identifying the vowel in the syllable. Moreover, durational cues may be particularly enhanced for vowels presented in isolated syllables where other factors like rate, stress, utterance position, etc. are not also affecting vowel duration as they do in natural speech. For these reasons it is possible that access to duration cues for infants, as well as for non-native adults, may be to some extent dependent on the specific testing procedures that have been used with these populations. Further studies in which vowel perception is examined in more natural speech contexts are needed to address this issue. With infants, research on perception of vowel contrasts within the word segmentation paradigm described by Jusczyk and Aslin (1995) would provide relevant data with respect to this issue.

In summary, the present study provides the first evidence to suggest that, like adults, prelingual infants can use dynamic spectral cues to differentiate spectrally similar vowels. Information provided by vowel duration differences also clearly supports vowel discrimination for German infants. However, it is unclear whether the strong bias toward vowel duration cues observed here reflects a language-independent pattern that all infants possess, or whether North German infants develop a special fondness for this cue very early in life, and to what extent the use of durational cues in inexperienced listeners is task specific. One of the interesting questions raised by these data concerns the ontogeny of vowel specification. Could it be that infants start their careers as vowel perceivers with a specific hierarchy of importance attached to the three types of acoustic information? For instance, is there an ontogenetic primacy of dynamic spectral cues over target spectral cues (or vice versa), and how does vowel duration help define vowel contrasts? Further research is needed to address the perceptual weighting of these cues in younger infants and how the relative weighting of acoustic cues changes with age and language experience.

ACKNOWLEDGMENTS

Research supported by grants from the *Deutsche Forschungsgemeinschaft* (DFG Grant No. Bo-1055/3-1 and 3-2) to O. Bohn and the Natural Sciences and Engineering Research Council of Canada (OGP0105397) to L. Polka. We thank Desiderio Saludes, Sonja Trent, and NIDCD Grant No. 00323 for assistance in generating stimulus materials, and Kirsten Schriever, Tatjana Soldat, Anja Steinlen, and Amira Yassine for assistance in testing infants. The comments of Shari Baum, James E. Flege, Anders Højen, James J. Jenkins, Susan Nittrouer, Megha Sundara, an anonymous reviewer, and the editor helped improve this manuscript. Special thanks go to Winifred Strange for her essential support, advice, and encouragement.

[1]For one token of /dɪt/, the onset contained only two pitch periods because the formants were near their maxima during the third pitch period.

[2]Four pitch periods were included in the offset of four tokens of /dot/ and one token of /dʊt/ because intensity and $F0$ were very low in the final three pitch periods of these tokens.

[3]In this game the players are presented an array of picture cards which are placed face down. Players take turns in which they can turn over only two cards; the object of the game is to locate the matched pairs in the set.

[4]Murphy *et al.* (1989) found that some of the 3-year-olds in this study could not accurately identify the intended vowel in an SC syllable unless the silent gap was filled with noise. However, they suggest that this outcome was likely due to the children failing to understand their task rather than an inability to perceive the intended vowel in the SC syllables. In fact, all of the children successfully identified the intended vowel in SC syllables in experiment 4 of this study in which task demands were made clearer.

[5]Results were also analyzed using A' as dependent variable; the same pattern of results was obtained. The A' score is a nonparametric index of sensitivity in which the subject's hit rate is corrected by their false-alarm rate.

[6]Correlations computed to assess the relationship between infant age and discrimination score in each condition were generally insignificant, the only exception being a weak positive correlation observed for the FND condition ($r = 0.3824$, $p = 0.0365$). These findings fail to suggest a clear association between age and task performance.

[7]We discovered that for three of the vowel contrasts, discrimination in the FULL condition on day 1 was clearly more difficult when infants were tested on the vowel change in one direction compared to the other direction (direction was counterbalanced for each contrast). This affected our attrition rate because infants had to reach a preset discrimination criterion to continue in the experiment. It should be noted that these direction effects

cannot be explained on the basis of vowel duration, i.e., a short to long change was harder for one contrast but the reverse was evident for another. Also, the direction effects were evident in the attrition rates for each direction (which was counterbalanced across infants for each contrast), but there was no further evidence of a direction effect among infants who met the preset criterion in the FULL condition on day 1. These direction effects are discussed in detail in Polka and Bohn (submitted).

[8]Observations by the authors as well as anecdotal reports from other researchers suggests that this is what infants generally do in this task. Specifically, when infants have been trained to respond to a stimulus change in one direction they become confused when presented with the same contrast with the assignment of foreground and background reversed.

[9]The chances of attaining 7/8 consecutive correct responses within 25 test trials are between $p<0.05$ and $p<0.001$, depending on the method used to calculate a floating probability (Eilers and Gavin, 1981).

[10]In each group, only five subjects each were tested in the INI and in the FIN condition. This decision was made because adult performance in these conditions in experiment 1 was relatively poor, and because we wanted to avoid testing each infant in two (rather than one) very difficult conditions.

[11]Discrimination of the /ɪ/–/ɛ/ contrast was not tested in these conditions beause the duration contrast between /ɪ/ and /ɛ/ was minimal (see Table I).

[12]Note that every infant was tested on the FULL condition; thus, these averages are based on 20 infants in the FULL condition for /i/–/e/, /e/–/ɪ/, and /o/–/ʊ/, whereas ten infants per contrast were tested in the other seven conditions.

[13]We recognize that we analyze the same data in multiple ANOVAs without applying a correction (e.g., Bonferroni). However, given the *F* values attained in these analyses, such a correction would not affect the results.

Andruski, J. E., and Nearey, T. M. (**1992**). ''On the sufficiency of compound target specification of isolated vowels and vowels in /bVb/ syllables,'' J. Acoust. Soc. Am. **91**, 390–410.

Bohn, O.-S. (**1995**). ''Cross-language speech perception in adults: L1 transfer doesn't tell it all,'' in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York, Timonium, MD), pp. 275–300.

Bohn, O.-S., and Flege, J. E. (**1990**). ''Interlingual identification and the role of foreign language experience in L2 vowel perception,'' Appl. Psycholing. **11**, 303–328.

Bohn, O.-S., and Flege, J. E. (**1992**). ''The production of new and similar vowels by adult German learners of English,'' Stud. Second Lang. Acquis. **14**, 131–158.

Eilers, R. E., and Gavin, W. J. (**1981**). ''The evaluation of infant speech perception skills: Statistical techniques and theory development,'' in *Language Behavior in Infancy and Early Childhood*, edited by R. Stark (North-Holland, Amsterdam), pp. 185–237.

Elliot, L., Connors, S., Kille, E., Levin, S., Ball, K., and Katz, D. (**1979**). ''Children's understanding of monosyllabic nouns in quiet and in noise,'' J. Acoust. Soc. Am. **66**, 12–21.

Flege, J. E., Bohn, O.-S., and Jang, S. (**1997**). ''The production and perception of English vowels by native speakers of German, Korean, Mandarin, and Spanish,'' J. Phonetics **25**, 437–470.

Jenkins, J. J. (**1987**). ''A selective history of issues in vowel perception,'' J. Mem. Lang. **26**, 542–549.

Jenkins, J. J., Strange, W., and Miranda, S. (**1994**). ''Vowel identification in mixed-speaker silent-center syllables,'' J. Acoust. Soc. Am. **95**, 1030–1043.

Jenkins, J. J., Strange, W., and Trent, S. A. (**1999**). ''Context-independent dynamic information for the perception of coarticulated vowels,'' J. Acoust. Soc. Am. **106**, 438–448.

Jusczyk, P. W., and Aslin, R. N. (**1995**). ''Infants' detection of sound patterns from fluent speech,'' Cogn. Psychol. **29**, 1–23.

Kuhl, P. K. (**1979**). ''Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories,'' J. Acoust. Soc. Am. **66**, 1668–1679.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (**1992**). ''Linguistic experience alters phonetic perception in infants by 6 months of age,'' Science **255**, 606–608.

Murphy, W. D., Shea, S. L., and Aslin, R. N. (**1989**). ''Identification of vowels in vowelless syllables by 3-year-olds,'' Percept. Psychophys. **46**, 375–383.

Nearey, T. (**1989**). ''Static, dynamic, and relational properties in vowel perception,'' J. Acoust. Soc. Am. **85**, 2088–2113.

Nittrouer, S. (**1992**). ''Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries,'' J. Phonetics **20**, 351–382.

Ohde, R. N., and Haley, K. L. (**1997**). ''Stop-consonant and vowel perception in 3- and 4-year-old children,'' J. Acoust. Soc. Am. **102**, 3711–3722.

Parnell, M. M., and Amerman, J. D. (**1978**). ''Maturational influences on perception of coarticulatory effects,'' J. Speech Hear. Res. **21**, 682–701.

Polka, L., and Bohn, O.-S. (**1996**). ''A cross-language comparision of vowel perception in English-learning and German-learning infants,'' J. Acoust. Soc. Am. **100**, 577–592.

Polka, L., and Bohn, O.-S. (submitted). ''Asymmetries in vowel perception'', Speech Communication.

Polka, L., and Werker, J. F. (**1994**). ''Developmental changes in the perception of non-native vowel contrasts,'' J. Exp. Psychol. Hum. Percept. Perform. **20**, 421–435.

Swoboda, P. J., Morse, P. A., and Leavitt, L. A. (**1976**). ''Continuous vowel discrimination in normal and at rist infants,'' Child Dev. **47**, 459–465.

Strange, W. (**1987**). ''Information for vowels in formant transitions,'' J. Mem. Lang. **26**, 550–557.

Strange, W. (**1989a**). ''Dynamic specification of coarticulated vowels spoken in sentence context,'' J. Acoust. Soc. Am. **85**, 2135–2153.

Strange, W. (**1989b**). ''Evolving theories of vowel perception,'' J. Acoust. Soc. Am. **85**, 2081–2087.

Strange, W., and Bohn, O.-S. (**1998**). ''Dynamic specification of coarticulated German vowels: Perceptual and acoustical studies,'' J. Acoust. Soc. Am. **104**, 488–504.

Trehub, S. E. (**1973**). ''Infants' sensitivity to vowel and tonal contrasts,'' Dev. Psychol. **9**, 91–96.

Trehub, S. E. (**1976**). ''The discrimination of foreign speech contrasts by infants and adults,'' Child Dev. **47**, 466–472.

Werker, J. F., Polka, L., and Pegg, J. E. (**1997**). ''The conditioned headturn procedure as a method for testing infant speech perception,'' Early Devel. Parenting **6**, 171–178.

Werker, J. F., and Lalonde, C. E. (**1988**). ''Cross-language speech perception: Initial capabilities and developmental change,'' Dev. Psychol. **24**, 672–683.

Category restructuring during second-language speech acquisition

Ian R. A. MacKay
Department of Linguistics, University of Ottawa, Ottawa K1N 6N6, Canada

James Emil Flege[a)]
Department of Rehabilitation Sciences, University of Alabama at Birmingham, Birmingham, Alabama 35294

Thorsten Piske
English Department, Kiel University, Kiel 24098, Germany

Carlo Schirru
Department of Linguistics, University of Padua, Padua 35137, Italy

(Received 27 January 2000; revised 5 April 2001; accepted 10 April 2001)

This study examined the production of English /b/ and the perception of short-lag English /b d g/ tokens by four groups of bilinguals who differed according to their age of arrival (AOA) in Canada from Italy and amount of self-reported native language (L1) use. A clear difference emerged between early bilinguals (mean AOA = 8 years) and late bilinguals (mean AOA = 20 years). The late bilinguals showed a stronger L1 influence than the early bilinguals did on both the production and perception of English stops. In experiment 2, the late bilinguals produced a larger percentage of prevoiced English /b/ tokens than early bilinguals and native English (NE) speakers did. In experiment 3, the late bilinguals misidentified short-lag English /b d g/ tokens as /p t k/ more often than the early bilinguals and NE speakers did. Experiment 4 revealed that the frequencies with which the bilinguals prevoiced /b d g/ in Italian and English were correlated. The observed differences between the early and late bilinguals were attributed to differences in the quantity and quality of English phonetic input they had received, not to a greater likelihood by the early than late bilinguals to establish new phonetic categories for English /b d g/. © *2001 Acoustical Society of America.*
[DOI: 10.1121/1.1377287]

PACS numbers: 43.71.Ft [KRK]

I. INTRODUCTION

Research examining second language (L2) speech acquisition has shown that ''early'' bilinguals who were first exposed to their L2 as children generally produce and perceive L2 phonetic segments more like monolingual speakers of the target L2 than do ''late'' bilinguals who were first exposed to the L2 in adolescence or adulthood. For example, Flege *et al.* (1999a) found that early Italian–English bilinguals produced and perceived English vowels more accurately than late bilinguals did (see also Flege *et al.*, 1995, for consonants). Flege *et al.* (1999b) attributed the observed age effect to the greater likelihood of phonetic category formation for L2 vowels by the early than the late bilinguals. It appears that in some instances, however, L2 vowels and consonants (or ''speech sounds,'' for short) are not identical to native language (L1) speech sounds but nevertheless do not differ sufficiently from the closest L1 sound(s) for category formation to occur, even for early bilinguals. The question addressed by this study was whether phonetic learning takes place for such L2 speech sounds in the absence of category formation.

The Speech Learning Model, or SLM (Flege, 1995), proposes that category formation for an L2 speech sound will be blocked if it is perceptually ''equated'' with an L1 speech sound. Equivalence classification is said to occur when the perceived instances of an L2 speech sound continue to be assimilated by the closest L1 speech sound even after many years of L2 use (e.g., Flege and Hillenbrand, 1984; Flege, 1995). According to the SLM, equivalence classification does not prevent phonetic learning from occurring. By hypothesis, a merged (or ''composite'') category that subsumes the equated L1 and L2 speech sounds will develop over time. It will be used to process the equated L1 and L2 speech sounds, and will reflect the properties of all L1 and L2 speech sounds that have been perceived to be instances of either the L1 category, the L2 category, or the merged category that eventually replaces the original L1 category that has undergone restructuring during L2 acquisition due to its perceptual linkage to an L2 sound.

The development of a merged category during L2 acquisition is predicted to yield two interrelated effects, even after many years of L2 use. Bilinguals' production and perception of the L2 speech sound will continue to differ from L2 monolinguals' production and perception because the bilinguals' production and perception will partially resemble patterns that are typical for the corresponding L1 speech sound. At the same time, bilinguals' production and perception of the L1 speech sound will gradually change so as to partially

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

resemble the corresponding L2 speech sound (see e.g., Flege, 1987; Flege *et al.*, 1995, 1999a).

The SLM's prediction of phonetic learning in the absence of category formation appears to differ from the views offered by other theories. According to the Native Language Magnet model (Kuhl, 2000, p. 106), listeners remain sensitive to subcategorical phonetic differences across languages, although they may no longer ''attend'' to such differences. One might hypothesize that L2 speech learning is governed by a kind of cross-language ''categorical perception.'' Such a view is related to the notion that the L1 phonological system acts as a kind of ''sieve'' that filters out the acoustic properties of L1 sounds that are needed to distinguish sounds in the L1 but not in the L2 (Polivanov, 1931; Trubetzkoy, 1939/1969; Hallé *et al.*, 1999). However, if L2 speech learning was governed by a kind of cross-language categorical perception, one would not expect phonetic learning to take place in the absence of category formation because the sensory input needed to guide learning for an L2 speech sound would be unavailable.

There are several reasons to think that, as proposed by the SLM (Flege, 1995), phonetic learning does take place for an L2 speech sound in the absence of category formation. Subcategorical phonetic differences across languages (or language varieties) appear to remain auditorily accessible to language learners. Flege and Hammond (1982) found that native English (NE) adults reproduced the voice onset time (VOT) values often heard in Spanish-accented English (i.e., values midway between those typical for short-lag and long-lag stops) when asked to mimic a Spanish accent in English sentences. Whalen *et al.* (1997) found that NE adults had difficulty discriminating unaspirated and aspirated allophones of English /p/ ([p] and [p^h]), but they could generally reproduce the VOT difference between the allophones in an imitation task. More globally, Munro *et al.* (2000) observed a measurable shift in the native language (L1) pronunciation of monolingual adults who were exposed to a nonprestige dialect of their L1 that differed from their native L1 dialect primarily in terms of subcategorical phonetic differences.

The results obtained in discrimination studies also suggest that subcategorical phonetic differences across languages remain auditorily accessible to language learners. Werker and Logan (1985) found that adult listeners showed sensitivity to certain cross-language phonetic differences under some task conditions (e.g., short ISIs in an AX task) but not others. An analysis of cortical evoked potentials led Sharma and Dorman (2000, p. 2702) to conclude that phonetic segments are processed at a sensory level that is ''not modified by exposure to the phonetic categories of a language'' and also at a level where ''language specific categories play a role.'' [1]

In this study, we examined Italian–English bilinguals' production of English /b/ and their perception of short-lag /b d g/ tokens in order to test for phonetic learning in the absence of category formation. We had several reasons to think that native Italian learners of English will not establish categories for English /b d g/. Italian /b d g/ and /p t k/ are realized with lead and short-lag VOT values, respectively (Magno-Caldognetto *et al.*, 1971, 1979). English /b d g/ are realized with short-lag VOT values or, less often, with lead VOT values (Lisker and Abramson, 1964; Flege and Eefting, 1986).[2] If English /b d g/ were realized with the lead VOT values that are typical for Italian (see experiment 1), such realizations would be acceptable in English; and so there would be no communicative pressure for Italian–English bilinguals to establish new categories for English /b d g/ (Port and Mitleb, 1983, p. 223).

Another reason to think that Italian–English bilinguals will not establish categories for English /b d g/ derives from universal constraints on phonetic systems. Stop consonants in the world's languages are realized with one of three modal VOT categories: lead (prevoiced), short-lag, and long-lag (Cho and Ladefoged, 1999). Keating (1984, p. 224) proposed that there may only be as many phonetic categories in languages as there are ''contrasting phonetic types.'' The same appears to hold true for individual bilinguals (see Flege and Eefting, 1988). Also, bilinguals usually identify short-lag stops in much the same way in their L1 and L2 (e.g., Elman *et al.*, 1977; Bohn and Flege, 1993). The establishment of short-lag categories for English /b d g/ is therefore likely to be preempted by existing Italian categories (viz., those for /p t k/).

The French–English and English–French bilinguals examined by Flege (1987) appear to have created merged categories for /t/. These bilinguals tended to produce both French /t/ and English /t/ with values that were intermediate to those observed for French and English monolinguals, respectively. No previous study has investigated the predicted effects of L1–L2 category merger for /b d g/, although L1 effects on the production of /b d g/ in an L2 have been observed for native Spanish and French learners of English (Caramazza *et al.*, 1973; Nathan, 1987; Williams, 1977b, 1979; Flege and Eefting, 1987). Based on this and the evidence reviewed earlier, we hypothesized that the Italian–English bilinguals examined here would detect differences between short-lag tokens of English /b d g/ and prevoiced tokens of Italian /b d g/ even if they did not establish categories for short-lag English /b d g/ tokens. If so, then the SLM (Flege, 1995) would predict the development of merged categories embracing the properties of corresponding L1 and L2 stops (e.g., English /b/ and Italian /b/). This led us to expect that the Italian–English bilinguals would differ from Italian monolinguals and also from English monolinguals. Specifically, the bilinguals should rely less on prevoicing as a perceptual cue to the identification of English stops as /b d g/. They should also prevoice English /b/ less often than Italian /b/ is typically prevoiced, but more often than is typical for English /b/.

If phonetic learning occurs in the absence of category formation, both predicted effects might be less evident for early than late Italian–English bilinguals. The early bilinguals examined in this study were probably exposed to more short-lag realizations of English /b d g/ in their lifetimes than the late bilinguals were. As in previous studies (e.g., Yeni-Komshian *et al.*, 2000), the early bilinguals had used their L2 longer than the late bilinguals had, and so were likely to have received more input from native speakers of the L2 (Jia

and Aaronson, 1999; Stevens, 1999). The predicted effects might also be greater for the bilinguals who continued to use their L1 (Italian) often than for the bilinguals who used their L1 relatively seldom. Hazan and Boulakia (1993) found that language dominance, which depends importantly on language use patterns, exerted a strong influence on the frequency of prevoicing in /b/'s spoken by French–English bilinguals (see also MacKay *et al.*, 2001; Meador *et al.*, 2000).

The present study was organized as follows. Experiment 1 provided an acoustic analysis of /b/ tokens that had been produced by English and Italian monolinguals. Experiment 2 examined the production of English /b/ by NE monolinguals and four groups of Italian–English bilinguals. Experiment 3 examined the same participants' identification of naturally produced /b d g/ and /p t k/ tokens. This experiment focused on the identification of word-initial /b d g/ tokens that had been realized with short-lag VOT values. Finally, experiment 4 examined the frequency with which Italian–English bilinguals prevoiced Italian /b d g/. Its aim was to test the prediction that the bilinguals whose productions of English /b d g/ most resembled NE speakers' productions would show the greatest influence of English on their productions of Italian /b d g/.

II. EXPERIMENT 1

Previous research (Magno-Caldognetto *et al.*, 1971, 1979) has shown that, as in other Romance languages, /b d g/ are produced with lead VOT values (i.e., are prevoiced) in Italian. The purpose of this experiment was to provide a direct comparison of /b d g/ production by monolingual native speakers of Italian and English.

A. Method

The participants were monolingual speakers of Italian (ten males, ten females) with a mean age of 25 years (range=19–33 years) and monolingual speakers of English (four males, eight females) with a mean age of 27 years (range=20–44 years). The Italian monolinguals were recorded in Padua, Italy; the English monolinguals were recorded in Birmingham, AL and Columbus, OH. The participants produced /bVdo/ nonwords (where V=/i e ɛ a ɔ o u/ for the Italian monolinguals, /i ɪ eɪ ɛ æ ɑ ʌ ɚ o u ʊ/ for the English monolinguals) after hearing four real words containing each target vowel of interest (e.g., *rido, fido, lido, nido* for Italian /i/). Twenty-one /b/ tokens produced by each Italian monolingual (7 vowel contexts × 3 repetitions) and 11 /b/ tokens produced by each English monolingual (one for each vowel context) were digitized at 22.05 kHz using a waveform editor (Cool Edit '96, Syntrillium Corp.).

Acoustic measurements were made from time domain waveforms displayed on the screen of a PC, supplemented by reference to digital spectrograms as needed. The duration of voicing lead (prevoicing) was measured from the onset of low-frequency periodicity to the onset of the /b/ release burst. In a subset of the tokens produced with lead VOT, the prevoicing ceased prior to the release of /b/. In these tokens, we also measured the duration of the silent gap from the cessation of prevoicing to the onset of the release burst. The remainder of tokens were produced without any prevoicing at all (i.e., were short-lag stops). VOT was measured in these tokens from the beginning of the release burst to the first upward-going zero crossing in the periodic portion (''vowel'') of the signal.

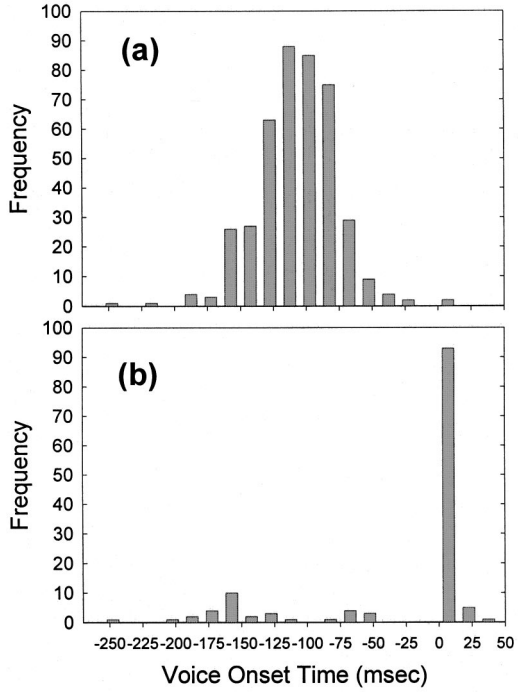


FIG. 1. The mean frequency of voice onset time (VOT) values for productions of /b/ in /bVdo/ nonwords by Italian monolinguals (a) and English monolinguals (b).

B. Results and discussion

Figures 1(a) and (b) show the frequency of VOT values in /b/ tokens produced in the same context by the Italian and English monolinguals, respectively. Of the 420 Italian /b/ tokens, 418 (99.5%) were prevoiced whereas just 32 (24%) of the 131 English /b/ tokens were prevoiced. Prevoicing in the Italian /b/ tokens averaged 106 ms in duration (s.d. =31, range=17–248), as compared to 136 ms (s.d.=9, range=51–251) for the prevoiced English /b/ tokens. The majority (99 or 76%) of the English /b/ tokens were realized as short-lag stops having a mean VOT of 11 ms (s.d.=3, range=1–31). Five of the 12 English monolinguals did not prevoice any /b/ tokens.

One unexpected, but nonetheless interesting, finding involved a qualitative difference in the prevoicing produced by the English and Italian monolinguals. Prevoicing ceased before the /b/ release in 23 (72%) of the prevoiced English /b/ tokens. The duration of the silent gap in these tokens averaged 47 ms (s.d.=24, range=13–95). Prevoicing did not cease before release in any of the prevoiced Italian /b/ tokens, however. The production of voicing through the entire closure of a labial stop requires an active volumetric expansion of the oral cavity, which must be learned (Flege *et al.*, 1987). It is uncertain if the presence of a silent gap before the release of a prevoiced stop will influence perception in Italian. However, the acoustic measurements and perceptual data obtained by Williams (1977a; see also Hazan and Bou-

TABLE I. Characteristics of the five groups of participants in experiment 2.

	AGE	GENDER	AOA	LOR	L1 USE	EDUC	B2FA
Native	50	9m				18	11.6
English	(5)	9f	...	...	...	(2)	(11.7)
Early-	50	8m	7	42	7%	14	12.1
low	(4)	10f	(3)	(4)	(4)	(3)	(8.9)
Early-	49	8m	8	40	43%	11	11.6
high	(6)	10f	(4)	(4)	(15)	(6)	(8.8)
Late-	51	10m	20	31	10%	2	15.0
low	(7)	8f	(3)	(8)	(5)	(2)	(12.0)
Late-	49	8m	20	29	53%	2	12.5
high	(8)	10f	(3)	(9)	(13)	(2)	(9.4)

Note: AGE=chronological age, in years; AOA and LOR=age of arrival and length of residence in Canada, in years; L1 USE=self-reported overall percentage use of Italian; EDUC=years of formal eduction in Canada; B2FA=best two-frequency average obtained in the heating screening, in dB. Standard deviations are in parentheses.

lakia, 1993, for French) suggested to her that such a silent gap may encourage Spanish speakers to hear /b d g/ tokens as /p t k/. Further research will be needed to determine if native Italian speakers learn to avoid producing /b/ with a silent gap in prevoicing in order to prevent perceptual confusions.

The results obtained here confirmed (e.g., Magno-Caldognetto *et al.*, 1971; Lisker and Abramson, 1964) that /b/ is prevoiced more frequently in Italian than English. These results may not necessarily generalize to the native dialect (or variety) of all of the Italian–English bilinguals examined subsequently in this study. However, it seems reasonable to think that when native Italian speakers first begin to speak English, they will prevoice English /b d g/ more frequently than NE speakers do as the result of cross-language differences in the phonetic implementation of /b d g/.

III. EXPERIMENT 2

Research with French–English and Spanish–English bilinguals has shown that they prevoice English /b d g/ more often than NE speakers do because /b d g/ is always, or nearly always, prevoiced in their L1s (Caramazza *et al.*, 1973; Hazan and Boulakia, 1993; Williams, 1977a, b, 1979; Nathan, 1987). The question addressed here was whether the same would hold true for Italian–English bilinguals and, if so, whether the size of the native versus non-native difference in prevoicing would vary as a function of AOA and/or L1 use.

A. Method

1. Participants

Eighteen participants were NE speakers who were not proficient in another language, and 72 were native speakers of Italian who had emigrated from Italy to Canada. As summarized in Table I, the bilinguals were assigned to one of four groups of 18 participants each (roughly half female).[3] Each of the 90 participants passed a pure-tone hearing screening that established thresholds for both ears at 500, 1000, 2000, 4000, and 8000 Hz. Preliminary analyses revealed that the combination of thresholds that was most strongly correlated with the perception data presented in experiment 3 was the best two-frequency average threshold (or ''B2FA,'' for short). The B2FA of the five groups did not differ significantly [$F(4,85)=0.3$, $p>0.10$].

The bilinguals were selected based on their age of arrival (AOA) in Canada and amount of continued L1 (Italian) use. Thirty-six ''early'' bilinguals arrived in Canada between the ages of 2–13 years (mean=8 years, s.d.=4), whereas 36 ''late'' bilinguals arrived between the ages of 15–26 years (mean=20 years, s.d.=3). The early and late bilinguals were then subdivided according to amount of continued L1 use, 1%–15% for the ''low-use'' bilinguals (mean =8%, s.d.=4) vs 25%–80% for the ''high-use'' bilinguals (mean=49%, s.d.=15).

A (2) AOA×(2) L1 use ANOVA revealed that the AOA difference between the low-use and high-use participants (13.2 vs 13.9 years) was nonsignificant [$F(1,68)=0.7$, $p>0.10$]. The lack of an AOA×L1 use interaction [$F(1,68)=0.9$, $p>0.10$] indicated that the AOA differences between the two groups of high-use participants (early-high, late-high), and between the two groups of low-use participants (early-low, late-low), were comparable. The early and late bilinguals differed significantly (25% vs 31%) according to amount of L1 use [$F(1,68)=7.3$, $p<0.01$]. However, the AOA×L1 use interaction in the analysis of self-reported percentage L1 use was nonsignificant [$F(1,68)=2.45$, $p>0.10$].

All but 3 of the 72 bilinguals had lived in Canada for at least 20 years. The early and late bilinguals differed significantly (means=41 vs 30 years) in length of residence (LOR) in Canada [$F(1,68)=47.5$, $p<0.01$], whereas the LOR difference between the low-use and high-use bilinguals (means=37 vs 35 years) was nonsignificant [$F(1,68)=1.4$, $p>0.10$]. Years of education in schools where the L2 is used as the language of instruction is known to affect certain aspects of L2 acquisition (e.g., Flege *et al.*, 1999b). The difference in number of years of education that the early and late bilinguals had obtained in English-speaking schools in Canada (means=13 vs 2 years) differed significantly [$F(1,68)=168.9$, $p<0.01$]. However, the education difference between the high-use and low-use bilinguals (means =7 vs 8 years) was nonsignificant [$F(1,68)=2.2$, $p>0.10$].

2. Speech materials

A delayed repetition procedure was used to elicit the production of word-initial /b/ tokens. A male and a female native speaker of Canadian English produced 15 test words including one token each of *bade*, *booed*, and *bed* and two tokens of *bad*. Their productions were digitized[4] and then presented via loudspeakers at the beginning of carrier phrases in two conditions. In the ''one-word'' condition, each test word to be repeated was followed by ''*...is the next word to say*.'' The participants repeated the target word after hearing the entire carrier phrase. In the ''three-word'' condition, the test words of interest occurred as the second member of three-word series (e.g., *Hid...bad...heed*) followed by ''*...are the next words to say*.'' The participants repeated all three words after hearing the entire carrier phrase. These

TABLE II. The mean percentage of word-initial English /b/ tokens that were prevoiced, and the percentage of stops that were produced with prevoicing that ceased before the stop release.

	Isolated word		Middle word in series	
	Prevoiced	Ceased	Prevoiced	Ceased
Native English	29% (33)	2% (5)	34% (33)	4% (13)
Early-low	61% (32)	17% (21)	53% (32)	18% (19)
Early-high	69% (35)	13% (18)	61% (36)	13% (17)
Late-low	79% (27)	5% (10)	80% (25)	8% (12)
Late-high	86% (17)	4% (9)	73% (31)	3% (6)

Note: Standard deviations are in parentheses.

procedures yielded 1800 /bVd/ words (90 participants $\times 2$ talkers $\times 5$ words $\times 2$ elicitation conditions) for analysis. Of these, seven words in the one-word condition and 115 words in the three-word condition were declared missing because of noise or because they were not repeated. In 21 other instances (of which 18 were in the three-word condition) the participants said a word that resembled the target word to be repeated (e.g., *bid* instead of *bad*). The /b/'s in these ''substitute'' words, along with the other /bVC/ target words, were digitized and measured as described in experiment 1.

B. Results

The total number of words that each participant repeated correctly in each condition (maximum=10) was tabulated. A (5) group$\times$(2) condition ANOVA examining these scores yielded a significant two-way interaction [$F(4,85)=5.3$, $p<0.01$]. Simple effects tests revealed that the effect of condition (isolated=9.9, series=9.8) was nonsignificant for the NE speakers [$F(1,85)=0.1$, $p>0.10$], but significant for all four bilingual groups (early-low: 9.9 vs 8.9; early-high: 9.9 vs 8.3; late-low: 9.9 vs 8.5; late-high: 9.9 vs 7.6) (F-values ranging from 7.9 to 43.3, $p<0.01$). This suggested that the bilinguals experienced greater difficulty than the NE speakers did in retaining three English words in working memory prior to repeating them. However, inasmuch as this finding does not bear directly on how /b/ was produced, it will not be discussed further.

Of the 1042 /b/ tokens that were measured, 646 (62%) were prevoiced. As shown in Table II, all four groups of bilinguals prevoiced /b/ more often than the NE speakers did in both conditions. The ''percent prevoiced'' scores were submitted to a mixed-design (5) group$\times$(2) condition ANOVA to determine if any group of bilinguals prevoiced /b/ more frequently than the NE speakers did. This analysis yielded a significant main effect of group [$F(4,85)=8.98$, $p<0.01$], a nonsignificant effect of condition [$F(4,85)=3.20$, $p>0.05$], and a nonsignificant two-way interaction [$F(4,85)=1.72$, $p>0.10$]. To test for native versus nonnative differences, the average of scores obtained in the two conditions by the four bilingual groups was compared to the average scores obtained for the NE speakers in a series of t-tests. These tests revealed that all four bilingual groups prevoiced /b/ more often (early-low: 57%, early-high: 65%, late-low: 79%, late-high: 79%) than the NE speakers did (mean=31%) (Bonferroni $p<0.05$). A supplementary Tukey's test did not reveal any significant differences between the four groups of bilinguals ($p>0.10$).

Experiment 1 revealed that prevoicing often ceased before the release burst in stops produced by English monolinguals, but never in stops produced by Italian monolinguals. In this experiment, prevoicing ceased before release in 149 of the English /b/ tokens that were examined. Table II shows the mean percentage of prevoiced /b/ tokens produced in which prevoicing ceased before the release. The duration of these silent gaps averaged 35 ms. More early than late bilinguals produced one or more /b/ tokens in which prevoicing ceased before the release (early-low: 16 participants, early-high: 13, late-low: 9, late-high: 5). The number of NE speakers who did so was small ($n=4$), apparently because they prevoiced so few stops.

The percentages of prevoiced stops in which voicing ceased before release were examined in a (5) group$\times$(2) condition ANOVA. It yielded a significant main effect of group [$F(4,85)=5.2$, $p<0.01$], a nonsignificant effect of condition [$F(4,85)=0.5$, $p>0.10$], and a nonsignificant two-way interaction [$F(4,85)=0.1$, $p>0.10$]. The average of scores obtained in the two conditions for the four bilingual groups were compared to the average scores obtained for the NE speakers in a series of t-tests. The two groups of early bilinguals were found to have produced more prevoiced stops in which the prevoicing ceased before the release than the NE speakers did (early-low: 18%, early-high: 13%, NE: 3%), whereas neither group of late bilinguals (late-low: 6%, late-high: 3%) differed from the NE speakers (Bonferroni $p<0.05$). A supplementary Tukey's test that tested for all possible between-group differences revealed that the Early-low group produced more such stops than the NE and the Late-low groups did ($p<0.01$).

The analysis just presented is potentially misleading inasmuch as the NE speakers produced so few prevoiced /b/'s. We therefore computed the percentage of the 20 /b/ tokens produced by each participant that were ''fully'' prevoiced, that is, had prevoicing that continued without interruption until the release burst. (We pooled the data obtained in the two conditions because the earlier analyses indicated that it would be appropriate to do so.) Figure 2 shows that the NE speakers produced fewer fully prevoiced stops than did any of the four bilingual groups. The one-way ANOVA examining these scores was significant [$F(4,85)=9.8$, $p<0.01$]. A series of t-tests revealed that the participants in the late-high, late-low, and early-high groups produced more fully prevoiced /b/'s (means=52%, 73%, 76%) than the NE speakers did (mean=28%; Bonferroni $p<0.05$), whereas the early-low participants (mean=39%) did not differ significantly from the NE speakers (Bonferroni $p>0.10$). A supplementary Tukey's test testing all possible between-group differences revealed that the late-high and late-low groups fully prevoiced /b/ more often than the NE and the early-low groups did ($p<0.01$).

The scores obtained for the four groups of bilinguals were examined in a series of (2) AOA$\times$(2) L1 use

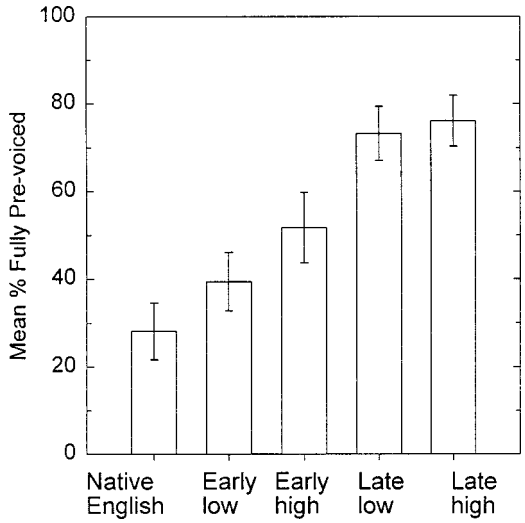


FIG. 2. The mean percentage of fully prevoiced stops produced by the participants in five groups. The brackets enclose ±1.0 SE.

ANOVAs. The variables examined in these analyses were the average percentage of /b/ tokens that were prevoiced in the two conditions (analysis 1); the average percentage of prevoiced /b/ tokens in which the pre-voicing ceased prior to the /b/ release (analysis 2); and the average percentage of /b/ tokens that were produced with full prevoicing (analysis 3). All three analyses revealed a significant main effect of AOA [analysis 1: $F(1,68)=8.1$; analysis 2: $F(1,68)=12.8$; analysis 3: $F(1,68)=18.6$, $p<0.01$]. The early bilinguals produced a smaller percentage of /b/ tokens with prevoicing than the late bilinguals did (61% vs 79%). They produced a larger percentage of prevoiced /b/ tokens in which voicing ceased before the release than the late bilinguals did (15% vs 5%). And they produced a smaller percentage of fully prevoiced /b/ tokens than the late bilinguals did (46% vs 75%). However, in all three analyses, the effect of L1 use was nonsignificant [analysis 1: $F(1,68)=0.3$; analysis 2: $F(1,68)=1.7$; analysis 3: $F(1,68)=1.3$, $p>0.10$], and the two-way interaction was nonsignificant [analysis 1: $F(1,68)=0.4$; analysis 2: $F(1,68)=0.1$; analysis 3: $F(1,68)=0.5$, $p>0.10$]

C. Discussion

The question addressed here was whether bilinguals who were experienced in English would show evidence of phonetic learning for English /b/. It appears that phonetic learning did take place. The percentage of English stops that were prevoiced by the four groups of Italian–English bilinguals ranged from an average of 57% for the early-low group to 79% for the two late bilingual groups. The percentage of stops that were fully prevoiced ranged from 39% for the early-low group to 76% for the early-high group. These percentages are in every case lower than the percentage observed for the production of Italian /b/ by Italian monolinguals in experiment 1 (99.5%).

The present results agree with previous studies examining French–English and Spanish–English bilinguals (Nathan, 1987; Williams, 1977b; Williams, 1979; Hazan and Boulakia, 1993) in showing that the Italian-English bilinguals prevoiced more often than NE speakers did. All four bilingual groups in this study prevoiced English /b/ more frequently than the NE speakers did. When the percentages of /b/ tokens that were fully prevoiced were examined, the participants in all of the bilingual groups except the early-low group differed significantly from the NE speakers. Other analyses revealed that the late bilinguals produced a significantly larger percentage of prevoiced /b/ tokens, and also a larger percentage of fully prevoiced /b/'s than the early bilinguals did. However, amount of continued L1 (Italian) use was not found to influence English /b/ production significantly.

A closer approximation to English phonetic norms by the early than late bilinguals might be attributed to the passing of a critical period (e.g., Scovel, 1988). However, the difference might have arisen from the quantity and quality of L2 input. Recall that the early bilinguals had lived longer in Canada than the late bilinguals had (41 vs 30 years); received more education in English-speaking Canadian schools (13 vs 2 years); and reported using Italian less overall (25% vs 31%). The early bilinguals may, therefore, have used English more often with NE speakers than the late bilinguals had, and may have been exposed to Italian-accented English less often than the late bilinguals had been.

IV. EXPERIMENT 3

Experiment 1 revealed that phonologically voiced stops are prevoiced more often by Italian than English monolinguals. Here we sought to determine if Italian–English bilinguals would misidentify short-lag tokens of /b d g/ as /p t k/ more often than NE speakers due to the lack of prevoicing.

A. Method

The participants from experiment 2 were tested in a quiet room in a single session using a notebook computer after having produced the speech materials examined earlier. The perceptual stimuli used here were non-words of the form /ˈCɑmɑ/, /ˈmɑCɑ/, and /ɑˈmɑC/ (where ''C'' indicates a token of /b d g p t k/). The stimuli were spoken by two NE males, then digitized at 22.05 kHz.[5] The /b d g/ tokens in the /ˈCɑmɑ/ stimuli were realized as short-lag stops having an average VOT of 15 ms. The /b d g/ tokens in the /ˈmɑCɑ/ stimuli, on the other hand, were produced with voicing through most (67 ms or 93%) of the closure interval. The same held true for the /b d g/ tokens in the /ɑˈmɑC/ stimuli (voicing in 80 ms or 94% of the closure intervals).

The stimuli just described were mixed with varying levels of noise to provide stimuli in which ceiling effects would not be evident. After the 36 stimuli (2 talkers×18) were normalized to 50% of full scale, three copies were made of each. The copies were digitally added to three 1000-ms pink noise segments. This yielded 36 stimuli each having S/N ratios of 16, 10, and 4 dB[6] in addition to the original 36 no-noise stimuli.

The 144 stimuli were presented via headphones (Sennheiser Model HD535) at a comfortable level that was determined individually for each participant before the experiment began. Test stimuli similar to the experimental stimuli

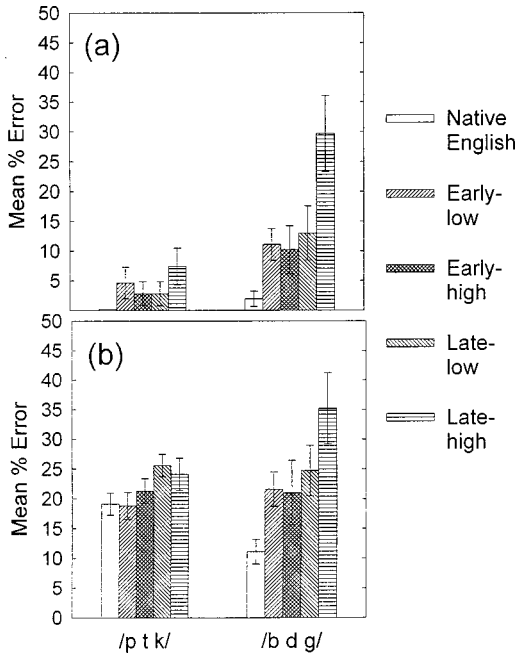


FIG. 3. The mean percentage of errors that the participants in five groups made in the identification of word-initial /p t k/ and /b d g/ tokens that were presented in the quiet (a) or in noise (b). The brackets enclose ±1.0 SE.

were played out. The volume was adjusted repeatedly until each participant confirmed hearing the test stimuli ''clearly'' in both ears.

The participants were told to identify the stop consonant in each stimulus as one of six stop consonants (''p, t, k, b, d, g'') by clicking one of six buttons shown on the computer screen.[7] Stimuli produced by the two talkers were presented in separate, counterbalanced blocks. The 36 no-noise stimuli were randomly presented a single time. The three sets of with-noise stimuli were then presented in a fixed order such that the stimuli became progressively noisier over the final three blocks (i.e., S/N levels of 16, 10, and 4 dB). A fixed order was used to help avoid ceiling effects that might arise as the participants gained familiarity with the small set of stimuli. The participants were given practice with feedback using stimuli produced by another talker before the experiment began. However, they received no feedback during the experiment. The participants were required to label the stop in each stimulus, and were told to guess if uncertain. The interval between each response and the next stimulus was 1.0 s.

B. Results

The percentages of errors that each participant made identifying /b d g/ and /p t k/ in the no-noise condition were calculated. The identification errors arose from confusions of the voicing feature (e.g., /p/ tokens heard as /b/), place of articulation (e.g., /d/ tokens heard as /g/), or both (e.g., /t/s heard as /g/). As shown in Fig. 3(a), the four groups of bilinguals erred more in identifying the /b d g/ than the NE speakers did, but did not differ much from the NE speakers for /p t k/.

The percent error scores were examined in a (5) group ×(2) phonological voicing ANOVA with repeated measures on the voicing factor. It yielded significant main effects of group [$F(4,85)=6.22$, $p<0.01$] and voicing [$F(1,85)=24.9$; $p<0.01$], as well as a significant two-way interaction [$F(4,85)=3.2$, $p<0.05$]. The interaction arose, in part, because the simple effect of voicing was nonsignificant for the NE group [$F(1,17)=2.12$, $p>0.05$], the early-low group [$F(1,17)=3.7$, $p>0.05$] and the early-high group [$F(1,17)=4.2$, $p>0.05$], whereas participants in the late-low [$F(1,17)=5.6$, $p<0.05$] and the late-high group [$F(1,17)=10.1$, $p<0.05$] made significantly more errors identifying /b d g/ than /p t k/. Also, a significant effect of group was obtained for /b d g/ [$F(4,85)=5.9$; $p<0.01$] but not /p t k/ [$F(4,85)=1.5$; $p>0.10$]. A series of *t*-tests was carried out to determine which bilingual group(s) differed from the NE speakers for /b d g/ in the no-noise condition. The late-high participants were found to have made more errors for /b d g/ than the NE speakers did (30% vs 2%; Bonferroni $p<0.05$) whereas those in the remaining three bilingual groups (early-low: 11%, early-high: 10%, late-low: 13%) did not differ from the NE speakers (Bonferroni $p>0.10$.)

As expected, a preliminary analysis revealed that the frequency of errors increased systematically as the stimuli became progressively more noisy (means=10% at 16 dB, 17% at 10 dB level, and 47% at the 4 dB S/N level). However, as in a study by MacKay *et al.* (2001), adding noise appeared to exert a comparable effect on the responses given by all five groups,[8] so we calculated an average percent error score for /b d g/ and /p t k/ in the three with-noise conditions. Each of these scores was based on 18 judgments (2 talkers ×3 stops×3 S/N levels). Figure 3(b) shows the average percent error scores obtained for /b d g/ and /p t k/ in the with-noise conditions.

A (5) group×(2) phonological voicing ANOVA examining the average with-noise scores yielded a significant main effect of group [$F(4,85)=4.2$, $p<0.01$]. The main effect of phonological voicing was nonsignificant [$F(1,85)=0.2$, $p>0.10$] but entered into a significant interaction with group [$F(4,85)=2.5$, $p=0.05$]. The interaction arose, in part, because the NE speakers made fewer errors for /b d g/ than /p t k/ [$F(1,17)=13.7$, $p<0.01$] whereas the simple effect of voicing was nonsignificant for all four bilingual groups [*F*-values ranging from 0.0 to 3.6; $p>0.05$]. Also, the simple effect of group was significant for /b d g/ [$F(4,85)=3.9$, $p<0.01$] but not /p t k/ [$F(4,85)=1.9$, $p>0.10$]. A series of four *t*-tests revealed that the late-high participants made more errors for /b d g/ than the NE speakers did (35% vs 11%; Bonferroni $p<0.05$), whereas the other three bilingual groups (early-low: 22%, early-high: 21%, late-low: 25%) did not differ from the NE speakers (Bonferroni $p>0.10$).

To summarize so far, the same results were obtained for stops presented in the no-noise and with-noise conditions. Only the late-high participants made more errors identifying short-lag tokens of English /b d g/ than the NE speakers did. The late-high participants may have misidentified the English /b d g/ tokens often because they were produced without the prevoicing that is typical for Italian /b d g/ (see experiment 1). We cannot be certain of this, however, because the scores we examined included place of place of articula-

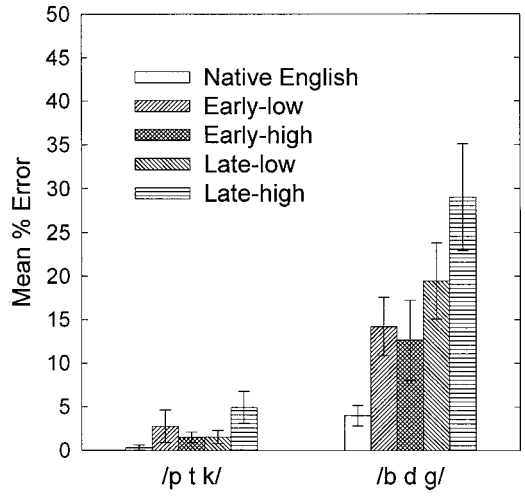


FIG. 4. The mean percentage of errors made by the participants in five groups that were due to misidentification of the voicing feature in word-initial tokens of /p t k/ and /b d g/ presented in noise. The brackets enclose ±1.0 SE.

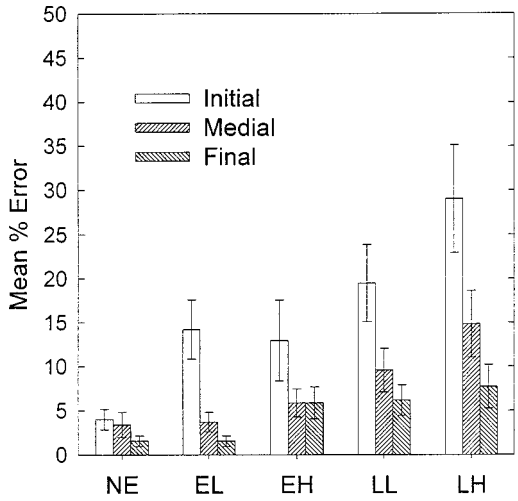


FIG. 5. The mean percentage of times that word-initial, word-medial, and word-final tokens of /b d g/ presented in noise were identified as /p t k/ by five groups of participants. The brackets enclose ±1.0 SE.

tion errors (e.g., the labeling of /t/ as /k/) as well as voicing errors.

We next calculated the percentage of errors involving just the voicing feature (e.g., the labeling of /b/ tokens as /p/) in the with-noise conditions.[9] As shown in Fig. 4, the four bilingual groups made more errors identifying the voicing feature in /b d g/ than the NE speakers did, but did not differ much from the NE speakers for /p t k/. A (5) group×(2) phonological voicing ANOVA examining the voicing error scores yielded significant main effects of group [$F(4,85) = 6.1$, $p < 0.01$] and voicing [$F(1,85) = 46.3$, $p < 0.01$] and a significant interaction [$F(4,85) = 3.0$, $p < 0.05$]. The interaction arose because the simple effect of group was significant for /b d g/ [$F(4,85) = 4.7$, $p < 0.01$] but not /p t k/ [$F(4,85) = 1,9$, $p > 0.10$]. A series of *t*-tests revealed that the late-low and late-high participants misidentified the /b d g/ tokens as /p t k/ more often (means = 19% and 29%) than the NE speakers did (mean = 4%) (Bonferroni $p < 0.05$), whereas neither group of early bilinguals (early-low: 14%, early-high: 13%) differed significantly from the NE speakers (Bonferroni $p > 0.10$).

The percentages of voicing errors made for /b d g/ by the four bilingual groups were examined separately in a (2) AOA×(2) L1 use ANOVA. It revealed that the late bilinguals made more voicing errors than the early bilinguals did (means = 24% vs 13%) [$F(1,68) = 5.3$, $p < 0.05$]. However, the difference between the high-use and low-use bilinguals (means = 21% vs 17%) was nonsignificant [$F(1,68) = 0.7$, $p > 0.10$], as was the two-way interaction [$F(1,68) = 1.4$, $p > 0.10$]. One possible explanation for a difference between the early and late bilinguals is a difference in the amount of English-language input. In support of this, the early–late difference for /b d g/ became nonsignificant [$F(1,67) = 3.0$, $p = 0.09$] when LOR was used as a covariate in a (2) AOA ×(2) L1 use ANOVA.

Unlike the word-initial /b d g/ tokens, those occuring in the word-medial and word-final /b d g/ positions were produced with closure voicing. We compared the frequency of voicing errors in word-initial, word-medial, and word-final /b d g/ tokens to evaluate how the absence of prevoicing in word-initial stops affected the bilinguals' identification judgments. As shown in Fig. 5, the participants in all five groups made more voicing errors for word-initial stops than for word-medial or word-final stops. However, the effect of position was greater for the four bilingual groups than for the NE speakers.

The voicing error scores shown in Fig. 5 were submitted to a (5) group×(3) word position ANOVA. It yielded significant main effects of group [$F(4,85) = 6.5$, $p < 0.01$] and position [$F(2,170) = 28.1$, $p < 0.01$] and also a significant two-way interaction [$F(8,170) = 2.3$, $p < 0.05$]. The interaction arose because word position affected the NE speakers and Italian-English bilinguals differently. The simple effect of position was nonsignificant for the NE speakers [$F(2,34) = 1.3$, $p > 0.10$] but it was significant (or marginally significant) for all four bilingual groups [late-low: $F(2,34) = 6.7$, $p < 0.01$; late-high: $F(2,34) = 9.6$, $p < 0.01$; early-low: $F(2,34) = 10.2$, $p < 0.01$; early-high: $F(2,34) = 2.8$; $0.05 < p < 0.10$]. Tukey's tests revealed that the participants in the early-low, late-low, and late-high groups made more errors for the word-initial /b d g/ tokens than for the word-medial or the word-final /b d g/ tokens ($p < 0.05$).

C. Discussion

This experiment showed that the late but not the early Italian-English bilinguals misidentified word-initial English /b d g/ tokens as /p t k/ more frequently than the NE speakers did. It appears that most of the late bilinguals' errors were due to the absence of prevoicing, for the word-initial /b d g/ stimuli were realized as short-lag stops rather than as lead (prevoiced) stops, as in Italian (see experiment 1). In a follow-up analysis, we compared the frequency of voicing errors for the word-initial, word-medial, and word-final /b d g/ tokens. The effect of position was nonsignificant for the NE speakers, but it was significant (or marginally significant) for all four bilingual groups, whose error rates were highest for the word-initial stops. We attributed the relatively

high rate of errors for the initial stops to the fact that they, but not the medial or final stops, lacked closure voicing. We acknowledge, however, that the medial and final tokens of /b d g/ differed from the word-initial /b d g/ in more than just closure voicing (see, e.g., Flege *et al.*, 1992).

Important individual differences existed in the frequency of voicing identification errors. The following number of 18 participants in the five groups made no errors identifying the voicing feature in the short-lag /b d g/ tokens: Native English: 9, early-low: 4, early-high: 6, late-low: 4, late-high:2. The following number of participants made more than 30% voicing errors for the word-initial /b d g/ tokens: Native English: 0, early-low: 2, early-high: 2, late-low: 6, late-high: 7. The basis of intersubject differences among the Italian–English bilinguals is uncertain. They might have arisen from differences in speech-learning ability, from differences in phonological short-term memory (MacKay *et al.*, 2001), from differences in the quantity and quality of phonetic input that had been received from native speakers of English, from degree of motivation to sound like a native speaker, or from some combination of factors.

V. EXPERIMENT 4

The final experiment examined Italian–English bilinguals' production of Italian /b d g/. Its purpose was to determine if learning English would cause Italian–English bilinguals to prevoice Italian /b d g/ less than Italian monolinguals. We examined the percentage of times that Italian /b d g/ were fully prevoiced because, as shown in experiment 1, /b/ is fully prevoiced less often by English than Italian monolinguals.

Four groups of bilingual participants differing in AOA and L1 use (early-low, early-high, late-low, late-high) participated. Experiment 2 showed that participants in an early-low group produced a smaller percentage of English /b/ tokens with full prevoicing than the participants in two late bilingual groups (late-low, late-high) did. A ''phonetic category merger'' hypothesis (see the Introduction and Flege, 1987, 1991, 1995) would therefore lead one to expect a greater influence of English on the production of Italian /b d g/ by participants in the early-low group than by late Italian–English bilinguals.

Extrapolating from the experiment 2 results for English, one might predict that only the participants in an early-low group would produce fewer fully prevoiced /b d g/ tokens than Italian monolinguals. It would have been ideal to obtain data from Italian monolinguals for this experiment, but we were unable to do so. We therefore evaluated the merger hypothesis by testing the prediction that participants in an early-low group would produce fewer fully prevoiced Italian /b d g/ tokens than participants in a late-low group would, whereas early and late bilinguals who often spoke Italian (early-high, late-high) would not differ significantly.

A. Method

Fifty bilinguals from experiments 2 and 3 (14 early-low, 13 early-high, 9 late-low, and 13 late-high participants) returned for this experiment a year later. Fourteen new participants were recruited in Ottawa to provide four groups of 16 participants each who differed in AOA and L1 use (see Table III). The 64 bilinguals had a mean age of 50 years (s.d.=6, range=30–63), and had been living in Canada for an average of 36 years (s.d.=8, range: 9–51 years). None reported an auditory disorder; and all passed a pure-tone hearing screening at octave frequencies between 500 and 4000 Hz (*re*: 35 dB HL). The 32 early bilinguals had an average AOA of 8 years (s.d.=3, range: 3–13) whereas the 32 late bilinguals had an average AOA of 19 years (s.d.=3, range: 15–28). The 32 low-use bilinguals reported using Italian 8% of the time on the average (s.d.=4, range: 2%–15%) whereas the 32 high-use bilinguals reported using Italian 46% of the time (s.d.=15, range: 29%–75%).

TABLE III. Characteristics of the four groups of native Italian participants in experiment 4.

	AGE	GENDER	AOA	LOR	L1 USE	EDUC	AGE
Early-low	49 (4)	7m 9f	7 (3)	42 (4)	6% (3)	14 (3)	49 (4)
Early-high	49 (6)	8m 8f	8 (4)	41 (6)	40% (13)	11 (5)	49 (6)
Late-low	51 (6)	7m 9f	18 (3)	33 (5)	10% (5)	2 (2)	51 (6)
Late-high	49 (8)	7m 9f	20 (4)	29 (9)	52% (15)	1 (2)	49 (8)
M	50 (6)	···	13 (7)	36 (8)	27% (22)	7 (6)	50 (6)

Note: AGE=chronological age, in years. Standard deviations are in parentheses; AOA and LOR=age of arrival and length of residence in Canada, in years; L1 USE=self-reported percentage use of Italian; EDUC=years of eduction in Canada, in years.

An adult female native speaker of Italian produced a list of 24 Italian words including nine words that began with prevoiced /b d g/ tokens (*babbo, bada, batto, dado, danno, data, gamma, gatta, gatto*). Tokens of these words that had been digitized at 11.025 kHz (with 16-bit resolution) were randomly presented to the participants via headphones at a comfortable level. The participants were told to repeat each word a single time after hearing it presented twice in a row, then to choose the correct English definition for the word from among the three definitions shown on the computer screen (e.g., ''a female cat'' for *gatta*). This last procedure ensured that the bilinguals' L2 (English) system was activated as they repeated the Italian words (see Grosjean, 1999).

The participants' productions of the 24 Italian words were recorded using a portable DAT tape recorder (Sony TCD8). The words beginning in /b d g/ were digitized (11.025 kHz) and then measured as in experiments 1 and 2. A total of 26 (4.5%) of the 576 test words were declared missing because of noise or failure to repeat.

B. Results

Of the 550 nonmissing /b d g/ tokens, 456 (83%) were prevoiced and 94 (17%) were realized as short-lag stops. Voicing ceased before the stop release in 59 (13%) of the prevoiced tokens. The duration of the silent gap in these tokens averaged 26 ms in duration. The percentage of /b d g/ tokens in which voicing continued without interruption until

release was calculated for each participant (maximum=9). All four bilingual groups produced a smaller percentage of fully prevoiced stops (early-low: 55%, early-high: 76%, late-low: 85%, late-high: 73%) than was observed for /b/'s spoken by the Italian monolinguals in experiment 1 (99.5%). The same held true if we consider only the percentage of fully prevoiced Italian /b/ tokens (early-low: 65%, early-high: 81%, late-low: 90%, late-high: 85%).

A two-way ANOVA examining the percentage of fully prevoiced Italian /b d g/ tokens indicated that the effect of AOA (early=65%, late=79%) was significant [$F(1,60)=2.9$, $p<0.01$] whereas the effect of L1 use (low use =70%, high use=74%) was not [$F(1,60)=0.5$, $p>0.10$]. Simple effects tests indicated that a significant two-way interaction was obtained [$F(1,60)=7.2$, $p>0.01$] because the effect of AOA was significant only for the low-use bilinguals. As predicted, the participants in the early-low group produced fewer fully prevoiced /b d g/ tokens than did those in the late-low group [$F(1,30)=13.0$, $p<0.01$], whereas the difference between late-high and early-high groups was nonsignificant [$F(1,30)=0.1$, $p>0.10$].

The low-use bilinguals tended to produce fewer fully prevoiced /b d g/ tokens than the high-use bilinguals did. However, the differences between the early-low and the early-high groups [$F(1,30)=4.0$], and those between the late-low and late-high groups [$F(1,30)=3.4$], were only marginally significant ($0.05<p<0.10$).

C. Discussion

The Italian–English bilinguals fully prevoiced Italian /b/ tokens less often (early-low: 65%, early-high: 81%, late-low: 90%, late-high: 85%) than Italian monolinguals did in experiment 1 (viz., 99.5%). There were, of course, important methodological differences between this experiment and experiment 1. Here we examined the immediate repetition of real Italian words beginning with fully prevoiced /b d g/ tokens whereas the participants in experiment 1 produced nonwords without a direct model. Still, these results suggest that learning English affected how all four groups of Italian–English bilinguals produced /b d g/ in their L1, Italian.

The SLM's (Flege, 1995) prediction that the participants in the early-low group would fully prevoice Italian /b d g/ less often than those in late-low group was confirmed. That is, the bilinguals whose productions of English /b/ was most English-like showed the greatest influence of English on their production of Italian /b d g/. In fact, for the 50 Italian–English bilinguals who participated in this experiment as well as in experiment 2, a significant correlation existed between the percentage of fully prevoiced English and Italian stops that were produced [$r(48)=0.47$, $p<0.01$]. That is, the less often the Italian–English bilinguals fully prevoiced English /b/, the less often they fully prevoiced Italian /b d g/. Significant partial correlations were also obtained when variations in self-rated ability to speak and understand Italian were partialled out [$r(46)=0.47$, $p<0.01$], when variations in self-rated ability to speak and understand English were partialled out [$r(46)=0.42$, $p<0.01$], and when four additional variables (age, LOR, percent use of Italian, and AOA) were partialled out [$r(44)=0.43$, $p<0.01$].

These findings supported our working assumption (see the Introduction) that even those bilinguals who learned English as children and seldom use Italian (early-low) had not developed new phonetic categories for English /b d g/. We suspect that most if not all of the bilinguals continued to identify English /b d g/ tokens as instances of their Italian /b d g/ categories. As the result of using a single category to process the many instances of /b d g/ they encountered in Italian and English words, the bilinguals' representations for /b d g/ may have gradually evolved to reflect all of the phonetic input they received (Flege, 1991, 1995). That is, their learning of English may have resulted in merged categories for /b d g/ that reflected a two-language source of phonetic input.

VI. GENERAL DISCUSSION

This study examined Italian–English bilinguals' production of English /b/ and their perception of short-lag tokens of English /b d g/ to determine if phonetic learning takes place in the absence of category formation. Experiment 1 showed that /b/ is fully prevoiced far more often in Italian than in English. Experiment 2 examined the production of /b/ by Italian–English bilinguals. The early bilinguals were found to prevoice English /b/ significantly less often than the late bilinguals did, and so resembled the NE speakers to a greater extent than the late bilinguals. However, the early bilinguals nevertheless prevoiced /b/ more often than the NE speakers did. Importantly, both the early and the late bilinguals fully prevoiced English /b/ less often than the Italian monolinguals in experiment 1 fully prevoiced Italian /b/. In experiment 3, the late but not the early Italian bilinguals misidentified /b d g/ tokens as /p t k/ significantly more often than NE speakers did, probably because the short-lag English /b d g/ tokens lacked the pre-voicing that is typically found in Italian /b d g/ (see experiment 1).

The difference between the early and late bilinguals might be attributed to the passing of a critical period (e.g., Bever, 1981; Scovel, 1988). However, in our opinion, the early and the late bilinguals differed primarily as the result of differences in the phonetic input they had received. As in previous studies examining immigrants to North America (e.g., Flege *et al.*, 1999b), the early bilinguals had lived for a longer time in an English-speaking environment, had received more education in English-speaking schools, and were likely to have used English more than the late bilinguals did. As a result, the early bilinguals may well have received more phonetic input from NE speakers over the course of their lives than the late bilinguals had (see Jia and Aaronson, 1999; Stevens, 1999). The early Italian–English bilinguals may have resembled English monolinguals more than the late bilinguals did because they had heard /b/ realized as a short-lag stop (or without full pre-voicing) more often than the late bilinguals had. In support of this, the effect of AOA on the percentage of voicing identification errors for the short-lag English /b d g/ tokens became nonsignificant when length of residence in Canada was used as a covariate.

As discussed in the Introduction, our working assumption was that the Italian–English bilinguals generally did not establish new phonetic categories for English /b d g/ in prestressed, word-initial position. According to the SLM (Flege, 1995), segmental production of an L2 speech sound may change in the absence of category formation. By hypothesis, it will do so through the merger of the phonetic properties of corresponding L1 and L2 sounds.

An analysis of L2 production from the perspective of dynamic systems theory provides another potential account of phonetic change in an L2 in the absence of category formation. Sancier and Fowler (1997) measured a Portuguese–English bilingual's production of Portuguese /p t/ and English /p t/ at several times. This bilingual produced longer VOT values in both English and Portuguese stops while living in the United States than in Brazil. As a result of a ''gestural drift'' towards ambient-language VOT norms, the bilingual's L1 (Portuguese) stops became somewhat less authentic in the United States, whereas her L2 (English) stops became less authentic in Brazil (see also Major, 1992). Sancier and Fowler (1997, p. 433) interpreted this to mean that concomitant changes in L1 and L2 stops arose through an ongoing change in a potential function as phonetic input (especially recent input) was received rather than through the establishment of new attractors.[10]

Experiment 4 provided an indirect test of our working assumption that the English stops /b d g/ are too similar to Italian /b d g/ for category formation to occur. It tested the SLM's prediction that, in the absence of category formation for L2 stops, L1 stops will begin to resemble L2 stops. Experiment 4 revealed that the bilinguals whose productions of English /b/ were most English-like (viz., the early-low participants) also showed the greatest influence of English on their production of Italian /b d g/. A positive correlation was found to exist between the production of stops in English and Italian. The less frequently the participants produced /b d g/ with full prevoicing in English, the less frequently they did so in Italian. This finding, which is analogous to the results of Flege (1987) for voiceless stops,[11] is consistent with the view that the Italian–English bilinguals had not established separate phonetic categories for English /b d g/. If they had done so, there would be no reason to expect their production of L1 stops to change so as to resemble L2 stops. We acknowledge, however, that additional research is needed to further probe for category formation for /b d g/ in both early and late bilinguals.

Additional work will also be needed to provide a better understanding of how L1 categories evolve to accommodate the properties of L2 sounds when category formation does not occur (Flege, 1995). We propose that the internal category structure of the bilingual participants' existing (Italian) /b/ category evolved to encompass the phonetic properties of both Italian and English /b/ tokens in proportion to the input they received (see related discussions by Kluender *et al.*, 1998 and Sancier and Fowler, 1997). More specifically, we propose that progressively less weight (or prominence) was accorded prevoicing in the Italian–English bilinguals' perceptual representations for word-initial tokens of /b d g/ as they gained experience with English. This is because prevoicing does not provide a reliable cue to the identity of phonologically voiced stops in English as it apparently does in Italian (see also Williams, 1977a, for Spanish). If so, one would expect the perceptual weight accorded to low-frequency periodicity just prior to the release burst to decrease as English input was received (see Williams, 1977a). This proposal is consistent with the conclusion drawn by Hazan and Boulakia (1993) regarding the perception of stops by French–English and English–French bilinguals. The bilinguals tended to weight spectral and temporal cues to the voicing feature in stops in a way that was not ''language specific'' (i.e., not just like those of English monolinguals or French monolinguals).

A final comment is in order regarding the role of category formation in L2 speech acquisition. Some investigators (e.g., Kluender *et al.*, 1998) have questioned the need for the construct ''phonetic category'' in speech acquisition research. Researchers in Barcelona have suggested that distinct long-term memory representations may not be established for the sounds encountered in an L2, even under seemingly ideal learning conditions (Sebastián-Gallés and Soto-Faraco, 1999; Bosch *et al.*, 2000). According to the SLM (Flege, 1995), on the other hand, the capacity to form new long-term memory representations (phonetic categories) for L2 speech sounds remains intact across the life span. However, the SLM proposes that the likelihood of category formation will vary according to the state of development of L1 categories at the time of first exposure to the L2, and the degree of perceived dissimilarity of an L2 speech sound from the closest L1 speech sound(s). By hypothesis, whether or not a new category is established for an L2 speech sound will affect how accurately the L2 sound will ultimately be produced and perceived.

The present study focused on L2 speech sounds for which category formation was unlikely, even by early bilinguals (see the Introduction). The findings of this study suggested that phonetic learning did take place for these speech sounds. In our view, the limits on learning observed in this study arose from the influence of previous phonetic learning and the distribution of L1 and L2 phonetic input that was received, not maturational constraints due to normal neurological development (e.g., Scovel, 1988). According to the SLM (Flege, 1995), L1 sounds will exert less influence on the perception and production of an L2 speech sound for which an independent category has been established. As a result, L2 sounds for which a category has been formed will be perceived and produced in a more nativelike fashion than L2 sounds that are processed using a merged category (see, e.g., Flege *et al.*, 1996a, b, 1999a). It is important to note, however, that the present study did not provide direct evidence that the Italian–English bilinguals did not establish new categories for English /b d g/. Additional work will be needed to better define the conditions under which categories are or are not established for L2 speech sounds, as well as the effects of category formation on L2 segmental production and perception.

In summary, this study suggested that phonetic learning for L2 stops takes place in the absence of category formation. Early bilinguals perceived English /b d g/ and produced

English /b/ more accurately than late bilinguals did. Some Italian bilinguals (mostly late bilinguals) continued to misidentify short-lag English /b d g/ tokens as /p t k/ and to fully prevoice /b/ more often than NE speakers did despite having spoken English for several decades. The bilinguals' divergences from English phonetic norms can be attributed to the fact that /b d g/ are fully prevoiced far more often in Italian than English. An analysis of Italian stop production suggested that both the early and late bilinguals' /b d g/ categories reflected experience with corresponding English and Italian stops. We suggest that the early bilinguals approximated English phonetic norms for /b d g/ more closely than the late bilinguals did because they had received more phonetic input from NE speakers, not because they were more likely to have established new phonetic categories for English /b d g/.

ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute for Deafness and Other Communicative Disorders (DC00257). The authors thank L. Gunnin for making VOT measurements, U. Lockridge for editorial assistance, R. Lanni for providing the Italian stimuli used in experiment 4, and K. Aoyama, A. Højen, and three anonymous reviewers for comments on earlier versions of the article.

[1]Sharma and Dorman (2000) found that native Hindi speakers identified and discriminated Hindi syllables with lead and short-lag VOT values (/ba/, /pa/) more accurately than NE speakers did because such differences are phonemic in Hindi but not English. Their analyses revealed that, for both native Hindi and English participants, N1 latencies increased as a function of the duration of the voicing lead in the /ba/ stimuli whereas a robust mismatch negativity (MMN) was observed for the native Hindi but not the NE participants.

[2]English /p t k/ are usually produced with long-lag VOT values, but may be realized with short-lag VOT values in certain contexts (Whalen *et al.*, 1997).

[3]The participants' place of origin in Italy (Abruzzo—24, Calabria—12, Sicilia—8, Veneto—7, Campania—6, Basilicata—4, Lazio—3, Friuli—2, Puglia—2, Lombardia—1, Marche—1, Piemonte—1, Toscana—1) did not vary systematically across the four groups.

[4]The male talker's /b/ tokens were all prevoiced (mean = 107 ms, s.d. = 38), whereas the female talker's /b/ tokens were all produced with short-lag VOT values.

[5]We used nonwords to minimize effects of lexical frequency on the participants' identification responses. Stimuli produced by three NE males were recorded and presented during the experiment. However, we found in a preliminary analysis that, unlike the case for the phonetically trained listeners who had taken part in a pilot experiment, native English-speaking listeners sometimes misidentified /b d g/ tokens in the no-noise condition. Their errors were due almost exclusively to the /d/ token produced by just one of the three talkers, so we decided to examine the responses for stops produced by just two speakers in subsequent analyses.

[6]The S/N estimates were based on the peak intensities of the disyllables and noise segments rather than on rms values. This is because the medial and final consonants contained silent intervals (i.e., the period of supraglottal constriction for /p t k/) whereas the stimuli with initial consonants did not.

[7]The participants were told that the six response alternatives corresponded to pronunciations, not spellings. They were told that /k/ is often spelled with ''c'' at the beginning of words such as cow and with ''ck'' at the end of words such as *tack*.

[8]A mixed design ANOVA examining the percentage of errors that each participant made identifying /p t k b d g/ on the four successive presentations (no-noise followed by the 16, 10, and 4 dB S/N levels) yielded significant main effects of group [$F(4,85)=5.2$; $p<0.01$] and presentation [$F(3,255)=285.2$; $p<0.01$] but a nonsignificant two-way interaction [$F(12,255)=1.4$; $p>0.10$]. Another ANOVA was carried out to determine if greater native versus non-native differences existed for stops presented in the with-noise conditions than in the no-noise condition. It too yielded significant main effects of group [$F(4,85)=6.08$; $p<0.01$] and condition [$F(1,85)=192.3$; $p<0.01$] but a nonsignificant interaction [$F(4,85)=1.11$; $p>0.10$].

[9]There were too few responses for the stops in the no-noise condition to support an analysis. The voicing error scores computed for both /p t k/ and /b d g/ were again based on 3 stops×2 talkers×3 S/N levels=18 judgments.

[10]On this view, the bilingual's sensory experience with a class of English phones such as [p^h] caused her to establish a new potential function. This new function was ''incorporated into...the original potential function'' for Portuguese [p] because it was based on ''far less experience'' than for Portuguese [p] phones. As a result of exposure to ''corresponding'' classes of phones in the L1 and L2 (see also Flege, 1987), the bilingual developed a new intrinsic coordinative dynamic, /p/, that was comprised of two attractors in ''close proximity'' to one another (viz., [p^h] and [p]).

[11]As mentioned in the Introduction, Flege (1987) found that French–English and English–French bilinguals tended to produce /t/ in their L2 with VOT values that were intermediate to the VOT values that are typical for French and English. The bilinguals' production of L1 stops also changed so as to partially resemble those of corresponding L2 stops.

Bever, T. (**1981**). ''Normal acquisition processes explain the critical period for language learning,'' in *Individual Differences in Language Learning Aptitude*, edited by K. Diller (Newbury House, Rowley, MA), pp. 176–198.

Bohn, O.-S., and Flege, J. E. (**1993**). ''Perceptual switching in Spanish/English bilinguals: Evidence for universal factors in stop voicing judgments,'' J. Phonetics **21**, 267–290.

Bosch, L., Costa, A., and Sebastian-Gallés, N. (**2000**). ''First and second language vowel perception in early bilinguals,'' European J. Cognitive Psychology **12**, 189–221.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., and Carbone, E. (**1973**). ''The acquisition of a new phonological contrast: The case of stop consonants in French–English bilinguals,'' J. Acoust. Soc. Am. **54**, 421–428.

Cho, T., and Ladefoged, P. (**1999**). ''Variation and universals in VOT: evidence from 18 languages,'' J. Phonetics **27**, 207–229.

Elman, J., Diehl, R., and Buchwald, S. (**1977**). ''Perceptual switching in bilinguals,'' J. Acoust. Soc. Am. **62**, 971–974.

Flege, J. E. (**1987**). ''The production of ''new'' and ''similar'' phones in a foreign language: Evidence for the effect of equivalence classification,'' J. Phonetics **15**, 47–65.

Flege, J. E. (**1991**). ''Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language,'' J. Acoust. Soc. Am. **89**, 395–411.

Flege, J. E. (**1995**). ''Second-language speech learning: Theory, findings, and problems,'' in *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, edited by W. Strange (York, Timonium, MD), pp. 233–272.

Flege, J. E., and Eefting, W. (**1986**). ''Linguistic and developmental effects on the production and perception of stop consonants,'' Phonetica **43**, 155–171.

Flege, J. E., and Eefting, W. (**1987**). ''The production and perception of English stops by Spanish speakers of English,'' J. Phonetics **15**, 67–83.

Flege, J. E., and Eefting, W. (**1988**). ''Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation,'' J. Acoust. Soc. Am. **83**, 729–740.

Flege, J. E., and Hammond, R. (**1982**). ''Mimicry of non-distinctive phonetic differences between language varieties,'' Studies in Second Lang. Acquis. **5**, 1–18.

Flege, J. E., and Hillenbrand, J. (**1984**). ''Limits on pronunciation accuracy in adult foreign language speech production,'' J. Acoust. Soc. Am. **76**, 708–721.

Flege, J. E., MacKay, I. R. A., and Meador, D. (**1999a**). ''Native Italian speakers' production and perception of English vowels,'' J. Acoust. Soc. Am. **106**, 2973–2987.

Flege, J. E., McCutcheon, M., and Smith, S. (**1987**). ''The development of skill in producing word-final English stops,'' J. Acoust. Soc. Am. **82**, 433–447.

Flege, J. E., Munro, M. J., and MacKay, I. R. A. (**1995**). ''The effect of age

of second language learning on the production of English consonants,'' Speech Commun. **16**, 1–26.
Flege, J. E., Munro, M., and Skelton, L. (**1992**). ''Production of the word-final English /t/-/d/ contrast by native speakers of English, Mandarin and Spanish,'' J. Acoust. Soc. Am. **92**, 128–143.
Flege, J. E., Schmidt, A. M., and Wharton, G. (**1996a**). ''Age of learning affects rate-dependent processing of stops in a second language,'' Phonetica **53**, 143–161.
Flege, J. E., Takagi, N., and Mann, V. (**1996b**). ''Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/,'' J. Acoust. Soc. Am. **99**, 1161–1173.
Flege, J. E., Yeni-Komshian, G. H., and Liu, S. (**1999b**). ''Age constraints on second language learning,'' J. Mem. Lang. **41**, 78–104.
Grosjean, F. (**1999**). ''Studying bilinguals: Methodological and conceptual issues,'' Bilingualism: Lang. Cogn. **1**, 117–130.
Hallé, P. A., Best, C. T., and Levitt, A. (**1999**). ''Phonetic vs phonological influences on French listeners' perception of American English approximants,'' J. Phonetics **27**, 281–306.
Hazan, V., and Boulakia, G. (**1993**). ''Perception and production of a voicing contrast by French-English bilinguals,'' Lang. Speech **36**, 17–38.
Jia, G., and Aaronson, D. (**1999**). ''Age differences in second language acquisition: The dominant language switch and maintainance hypothesis,'' in *Proceedings of the 23rd Annual Boston University Conference on Language Development*, edited by A. Greenhill, H. Littlefield, and C. Tano editors (Cascadilla, Somerville, MA), pp. 301–312.
Keating, P. (**1984**). ''Phonetic and phonological representations of stop consonant voicing,'' Lang. **60**, 286–319.
Kluender, K., Lotto, A., Holt, L., and Bloedel, S. (**1998**). ''Role of experience for language-specific functional mappings of vowel sounds,'' J. Acoust. Soc. Am. **104**, 3568–3582.
Kuhl, P. (**2000**). ''Language, mind, and brain: Experience alters perception,'' in *The New Cognitive Neuroscience*, 2nd ed., edited by M. S. Gazzaniga (MIT, Cambridge, MA), pp. 99–115.
Lisker, L., and Abramson, A. (**1964**). ''A cross-language study of voicing in initial stops: Acoustical measurements,'' Word **20**, 384–422.
MacKay, I. R. A., Meador, D., and Flege, J. E. (**2001**). ''The identification of English consonants by native speakers of Italian,'' Phonetica **58**, 103–125.
Magno-Caldognetto E., Abati, A., and Dossi, L. (**1971**). *Consonanti occlusive sorde e sonore della lingua italiana* (Riccardo Pàtron, Bologna).
Magno-Caldognetto, E., Ferrero, F., Vagges, K., and Bagno, M. (**1979**). ''Indici acustici e indici percettivi nel riconoscimento dei suoni linguistici,'' Acta Phoniatrica Latina **2**, 219–249.
Major, R. (**1992**). ''Losing English as a first language,'' Mod. Lang. J. **76**, 190–208.
Meador, D., Flege, J. E., and MacKay, I. A. R. (**2000**). ''Factors affecting the recognition of words in second language,'' Bilingualism: Lang. and Cogn. **3**, 55–67.
Munro, M. J., Derwing, T. M., and Flege, J. E. (**2000**). ''Canadians in Alabama: A perceptual study of dialect acquisition in adults,'' J. Phonetics **27**, 385–403.
Nathan, G. S. (**1987**). ''On second-language acquisition of voiced stops,'' J. Phonetics **15**, 313–322.
Polivanov, E. (**1931**). ''La perception des sons d'une langue étrangère,'' Travaux du Cercle Linguistique de Prague **4**, 79–96.
Port, R. F., and Mitleb, F. M. (**1983**). ''Segmental features and implementation in acquisition of English by Arabic speakers,'' J. Phonetics **11**, 219–229.
Sancier, M. and Fowler, C. (**1997**). ''Gestural drift in a bilingual speaker of Brazilian Portuguese and English,'' J. Phonetics **25**, 421–437.
Scovel, T. (**1988**). *A Time to Speak. A Psycholinguistic Inquiry into the Critical Period for Human Speech* (Newbury House, Cambridge, MA).
Sebastián-Gallés, N., and Soto-Faraco, S. (**1999**). ''On-line processing of native and non-native phonemic contrasts in early bilinguals,'' Cognition **72**, 111–123.
Sharma, A., and Dorman, M. (**2000**). ''Neurophysiological correlates of cross-language phonetic perception,'' J. Acoust. Soc. Am. **107**, 2697–2703.
Stevens, G. (**1999**). ''Age at immigration and second language proficiency among foreign-born adults,'' Lang. in Soc. **28**, 555–578.
Trubetzkoy, N. S. (**1939/1969**). *Principles of Phonology*, translated by C. A. Baltaxe (Univ. of California, Berkeley, CA).
Werker, J., and Logan, J. (**1985**). ''Cross-language evidence for three factors in speech perception,'' Percept. Psychophys. **37**, 35–44.
Whalen, D. H., Best, C. T., and Irwin, J. R. (**1997**). ''Lexical effects in the perception and production of American English /p/ allophones,'' J. Phonetics **25**, 421–436.
Williams, L. (**1977a**). ''The voicing contrast in Spanish,'' J. Phonetics **5**, 169–184.
Williams, L. (**1977b**). ''The perception of stop consonant voicing by Spanish-English monolinguals,'' Percept. Psychophys. **21**, 289–297.
Williams, L. (**1979**). ''The modification of speech perception and production in second-language learning,'' Percept. Psychophys. **26**, 95–104.
Yeni-Komshian, G. H., Flege, J. E., and Liu, S. (**2000**). ''Pronunciation proficiency in the first and second languages of Korean-English bilinguals,'' Bilingualism: Lang. Cognition **3**, 131–150.

Effects of degradation of intensity, time, or frequency content on speech intelligibility for normal-hearing and hearing-impaired listeners[a)]

Nicolle H. van Schijndel,[b)] Tammo Houtgast, and Joost M. Festen[c)]
Department of Otolaryngology, University Hospital VU, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands

(Received 8 November 1999; accepted for publication 16 April 2001)

Many hearing-impaired listeners suffer from distorted auditory processing capabilities. This study examines which aspects of auditory coding (i.e., intensity, time, or frequency) are distorted and how this affects speech perception. The distortion-sensitivity model is used: The effect of distorted auditory coding of a speech signal is simulated by an artificial distortion, and the sensitivity of speech intelligibility to this artificial distortion is compared for normal-hearing and hearing-impaired listeners. Stimuli (speech plus noise) are wavelet coded using a complex sinusoidal carrier with a Gaussian envelope ($\frac{1}{4}$ octave bandwidth). Intensity information is distorted by multiplying the modulus of each wavelet coefficient by a random factor. Temporal and spectral information are distorted by randomly shifting the wavelet positions along the temporal or spectral axis, respectively. Measured were (1) detection thresholds for each type of distortion, and (2) speech-reception thresholds for various degrees of distortion. For spectral distortion, hearing-impaired listeners showed increased detection thresholds and were also less sensitive to the distortion with respect to speech perception. For intensity and temporal distortion, this was not observed. Results indicate that a distorted coding of spectral information may be an important factor underlying reduced speech intelligibility for the hearing impaired. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1378345]

PACS numbers: 43.71.Ky, 43.71.An, 43.66.Dc, 43.66.Fe, 43.66.Mk [RVS]

I. INTRODUCTION

The difficulty hearing-impaired listeners have in perceiving speech in noise has been the subject of many investigations, but is still not entirely understood. Although audibility plays an important role, several studies have shown that this cannot explain the whole problem [see, e.g., Moore (1996) or Noordhoek *et al.* (2000)]. These studies have demonstrated that factors apart from reduced audibility, called suprathreshold deficits, degrade speech processing. Suprathreshold deficits can distort the auditory processing of either intensity, time, or frequency information, or a combination of these types of information. For example, excessive forward and backward masking are consequences of suprathreshold deficits that may be reduced to a single factor of distorted temporal coding; excessive upward and downward spread of masking may be related to distorted spectral coding. Impaired loudness perception probably relates to a distorted representation of intensity information. This study evaluates these three types of information. The aim is to investigate how reduced speech intelligibility relates to distorted coding of intensity, time, or frequency.

Auditory coding cannot be manipulated directly. However, one can investigate the differences in auditory functions among hearing-impaired subjects on specific auditory tests related to accuracy of intensity, time, or frequency coding, and correlate these with their speech-perception performance. In several studies this correlation approach was applied, concentrating on the role of reduced temporal or spectral resolution. The role of reduced temporal resolution in reduced speech intelligibility in noise is not yet clear. In some studies a significant correlation between speech intelligibility and temporal resolution was found (Tyler *et al.*, 1982; Dreschler and Plomp, 1985; Moore and Glasberg, 1987); in other studies this was not so (Festen and Plomp, 1983; van Rooij and Plomp, 1990). With respect to reduced spectral resolution, in most studies a significant correlation with speech intelligibility was found (Patterson *et al.*, 1982; Festen and Plomp, 1983; Dreschler and Plomp, 1985; Horst, 1987). On the other hand, this was not the case in a few other studies (van Rooij and Plomp, 1990; Smoorenburg, 1992).

The correlation approach results in statistical relations between reduced speech perception and suprathreshold deficits. A drawback of this approach is that one cannot exclude that an underlying common factor causes the observed correlation. For example, if a correlation between speech intelligibility and spectral resolution is observed, an underlying common factor can be the hearing threshold. Then, higher hearing thresholds instead of reduced frequency selectivity may cause reduced speech perception. In different studies, underlying factors probably had different effects, which may explain the different results. Relations between distorted auditory coding and speech perception can be investigated in a

[a)]Part of the results was presented at the Joint ASA/EAA/DEGA Meeting ''Berlin 99'' [Collected papers from the joint meeting ''Berlin 99,'' Deutsche Gesellschaft für Akustik, Oldenburg].

[b)]Current address: Philips Research Laboratories, Digital Signal Processing (WY 82), Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands; electronic mail: Nicolle.van.Schijndel@philips.com
[c)]Electronic mail: JM.Festen@azvu.nl

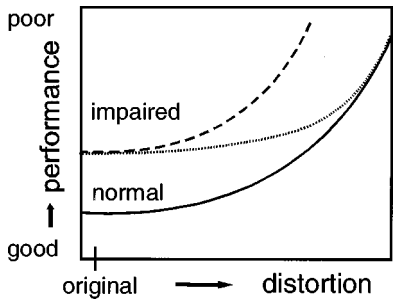


FIG. 1. Illustration of the distortion-sensitivity model. Performance for hearing-impaired listeners as a function of distortion is compared with that of normal-hearing listeners (solid line). The possible outcome of such an experiment is ''convergence'' (dotted and solid lines) or ''no convergence'' (dashed and solid lines).

more direct way using the distortion-sensitivity model (Houtgast, 1995; van Schijndel *et al.*, 2001).

Under the distortion-sensitivity model (Fig. 1), the relation between speech intelligibility and distorted *auditory* coding is studied by simulating the effect of the auditory deficit by *artificial* distortion of the speech signal. The idea is that removing cues that are not perceived by the hearing impaired will not affect their performance. Performance is measured as a function of distortion, and compared for normal-hearing and hearing-impaired listeners. Two trends may be observed: *convergence* (dotted and solid lines) or *no convergence* (dashed and solid lines). In the convergence case, hearing-impaired listeners are less sensitive to the distortion than normal-hearing listeners. Then, it may be concluded that the artificial distortion relates to distorted auditory coding that impedes performance. The artificial distortion affects the sound characteristics in the same way as the auditory deficits. In the no-convergence case, hearing-impaired listeners are as sensitive to the distortion as normal-hearing listeners, indicating that the artificial distortion has no relation to the hearing deficits causing difficulties in speech perception. In the no-convergence case, performances of the normal-hearing and hearing-impaired listeners will run parallel, or may even diverge. Divergence will be discussed in more detail in Sec. III C 2. A few studies (Duquesnoy and Plomp, 1980; ter Keurs *et al.*, 1993; Turner *et al.*, 1995; van Schijndel *et al.*, 2001) used the principles of the distortion-sensitivity model so far, but they did not explicitly explain their results in terms of the model, except the last study.

In van Schijndel *et al.* (2001) the distortion-sensitivity model was used with respect to the coding of intensity information. It was concluded that reduced intensity coding accuracy may partly explain impaired speech perception.

With respect to the coding of temporal information, Duquesnoy and Plomp (1980) measured speech reception of normal-hearing and hearing-impaired listeners as a function of reverberation time. Their results show that hearing-impaired listeners are as sensitive to reverberation as normal-hearing listeners. In terms of the distortion-sensitivity model, this leads to the conclusion that speech-perception problems are not caused by a deficit that introduces a delay to parts of the speech energy, as distorted temporal coding may do.

With respect to coding of spectral information, ter Keurs *et al.* (1993) compared the effect of reduced spectral contrast on speech perception in normal-hearing and hearing-impaired listeners. They concluded that ''limited resolution of spectral contrast is only loosely associated with hearing loss for speech in noise.'' Turner *et al.* (1995) compared speech reception of hearing-impaired and normal-hearing listeners for unprocessed speech and for speech of which spectral cues were removed. For the original speech, hearing-impaired listeners had lower speech-intelligibility scores than the normal-hearing listeners. However, for speech without spectral cues, hearing-impaired listeners understood as well as normal-hearing listeners. In terms of the distortion-sensitivity model, this convergence indicates that the reduced speech intelligibility by hearing-impaired listeners is related to a degraded processing of spectral cues. It should be mentioned that this is our interpretation of the data. Turner *et al.* were interested in the ability of hearing-impaired listeners to use temporal cues. Their conclusion, not in conflict with ours, is that the temporal accuracy of speech coding of hearing-impaired listeners is not impaired in terms of speech recognition.

The studies mentioned previously obtained data that can be analyzed in terms of the distortion-sensitivity model. The effects of distortion of intensity, time, and frequency information on speech perception were studied in isolation, although these three domains are not completely independent. Manipulation in one domain will affect the other domains. For example, spectral smearing introduces temporal smearing and vice versa. In Sec. II A 4, this will be illustrated. Awareness of these unwanted by-products of the speech-processing algorithm is important. Therefore, in the present study, the interdependency of the intensity, time, and frequency domains was taken into account.

In short, this study addresses which domains in auditory coding (i.e., intensity, time, or frequency) cause speech-perception problems for hearing-impaired listeners. First, it is investigated which sound domains are less clearly perceived by hearing-impaired listeners. For this, detection thresholds for artificially applied distortions of intensity, time, or frequency are measured. If a particular type of information is less clearly perceived by hearing-impaired listeners, the detection thresholds for the distortion of this information will probably be higher. The influence of distorted coding on speech perception was investigated by means of the distortion-sensitivity model. Speech intelligibility is measured as a function of the degree of artificial distortion of intensity, time, or frequency information. Comparison of the performance for normal-hearing and hearing-impaired listeners may provide insight into the role of reduced accuracy in auditory coding as a possible explanation for the degraded performance of the hearing impaired.

II. METHOD

A. Degradation of intensity, time, and frequency information

In this study, a sound-processing algorithm is used to degrade artificially the intensity, time, and frequency content of speech. The degradation is intended to simulate the effects of distorted auditory coding. By means of the speech-

reception threshold test (SRT, Sec. II D 3 a), speech intelligibility of sentences was measured as a function of applied artificial distortion. In order to simulate auditory coding, a perceptually relevant spectro-temporal decomposition and recomposition method was developed. This method was also used in van Schijndel *et al.* (2001), and is described in the following.

1. Spectro-temporal decomposition and recomposition

To model auditory spectro-temporal coding, sounds were described in the time–frequency domain by means of a wavelet transform. Compared with the short-time Fourier transform, the wavelet transform matches auditory system coding more closely because it uses a logarithmic frequency scale (e.g., Rioul and Vetterli, 1991). An important criterion in the choice of the mother wavelet is its spectral and temporal width. Results of van Schijndel *et al.* (1999) suggest that a Gaussian-windowed sinusoid with a shape factor between 0.15 and 0.3 roughly matches the auditory time-frequency window. Therefore, as the prototype analysis function, a Gaussian wavelet was chosen. The Gaussian wavelet is a complex sinusoidal carrier with a Gaussian envelope:

$$s(t)=\sqrt{\alpha f_0}\exp(i2\pi f_0 t)\exp(-\pi(\alpha f_0 t)^2), \tag{1}$$

in which f_0 is the carrier frequency, α is the shape factor, and $\sqrt{\alpha f_0}$ normalizes the energy of the analysis function. This time–frequency window has an effective bandwidth of $\Delta_f=\alpha f_0$ and an effective duration of $\Delta_t=1/(\alpha f_0)$ (Gabor, 1947). The effective bandwidth of the analysis function was set to $\frac{1}{4}$ octave [roughly equal to the auditory critical band (Scharf, 1970)]. This corresponds to a shape factor $\alpha=0.1735$. As a result, the effective duration of the time–frequency window is 5.76 ms at 1 kHz (1.44 ms at 4 kHz). The effective number of periods contained within the Gaussian envelope equals 5.8 (i.e., $1/\alpha$).

This Gaussian wavelet was used to construct a wavelet decomposition that covers the time–frequency plane. Shifts of this prototype analysis function cover the temporal range; scales of the prototype function cover the spectral range. The scaling is controlled by varying the carrier frequency f_0. The decomposition results in complex wavelet coefficients, which can be characterized by a modulus, a phase, and a position in the time–frequency plane.

For simultaneous sampling in time and frequency, the Nyquist sampling theorem was applied twice (Allen, 1977; Allen and Rabiner, 1977). The sampling interval was based on the temporal and spectral range over which the Gaussian wavelet is essentially different from zero. Since the Gaussian wavelet does not have compact support[1] in time, or in frequency, the range between the points that were 25 dB down from the peak was taken as the range over which the window is significant, i.e., essentially different from zero. Thus, outside these 25-dB down points, the window is considered to be negligible. This definition corresponds to a duration of about twice the effective duration and a bandwidth of about twice the effective bandwidth. Application of this criterion leads to a sampling of one wavelet per three periods of the wavelet carrier frequency along the time axis, and eight wavelets per octave along the frequency axis. Theoretically, the number of complex coefficients needed to describe the signal using the 25-dB criterion for sampling is about two coefficients per input sample (Allen, 1977). In this study, the frequency of the signals was limited to the range from 250 to 4000 Hz. As a result, 1 s of speech (sampling frequency: 44.1 kHz; no information below 250 Hz or above 4 kHz preserved) was described by 16×10^3 complex wavelet coefficients.

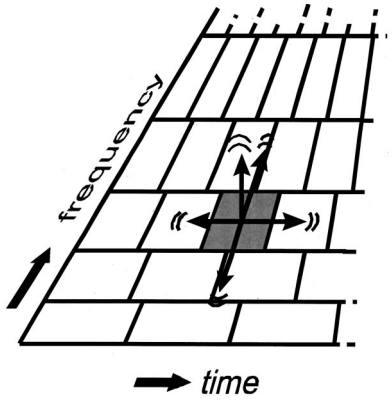


FIG. 2. Schematic illustration of the perturbation of the intensity, time, or spectral information. The Gaussian wavelets are symbolized by rectangles. Each wavelet is given a random perturbation with respect to its intensity, temporal position, or spectral position.

Using these wavelet coefficients, sounds can be reconstructed by an overlap-add procedure. Theoretically, the reconstruction is not perfect. However, using the 25-dB criterion for sampling in time and frequency, little or no aliasing occurs in either the time or the frequency domain. Adequate sampling is important for two reasons (Allen and Rabiner, 1977). First, the difference between the recomposed signal and the original signal must not be noticeable to a listener. Second, in this study modifications to the spectro-temporal decomposition of sound are performed. When modifying undersampled spectro-temporal representations of sound, interactions between modification and window shape may occur. Such interactions will lead to unwanted by-products. As a result of the careful sampling in our decomposition and recomposition scheme, (1) the difference between an original and a recomposed signal was very small and not noticeable to the listener, and (2) the scheme is robust for interactions between window shape and modifications in the decomposed signal.

Between decomposition and recomposition, the integrity of the intensity, time, or frequency information was reduced to simulate poor auditory coding. Intensity degradation was obtained by introducing uncertainty in the modulus of each wavelet coefficient. Temporal and spectral degradations were obtained by introducing uncertainty in the temporal and spectral position of each wavelet, respectively. In Fig. 2, this is illustrated schematically. In the following paragraphs, these different types of degradation will be explained in more detail. After the perturbation, the energy contained in each frequency band over the whole test sentence was scaled to equal the original energy in that band. Since this study

aims at investigating speech-perception performance in noise, speech and noise were summed before processing.

2. Degradation of the intensity accuracy

To degrade the accuracy of the intensity information, the modulus of the wavelet coefficients was perturbed (intensity perturbation). This was achieved by multiplying each wavelet coefficient by a random factor. As a result, silence will remain silence after perturbation. The random perturbation factor ε (in dB) was chosen from a uniform distribution with zero mean and boundaries[2] $-L_D/2$ and $+L_D/2$. Thus the modulus of each individual coefficient was multiplied by a different random factor $10^{\varepsilon/20}$. The intensity perturbation levels used in this study were moderate and the uncomfortable loudness levels of the subjects were never approached.

3. Degradation of the temporal accuracy

To degrade the accuracy of the temporal information, the positions of the wavelets were shifted randomly along the temporal axis (temporal perturbation). To avoid a degradation of the accuracy of spectral information as much as possible, only the temporal envelope of the wavelets was displaced, not the underlying fine structure. The new fine structure was calculated by extrapolation of the original fine structure to the new position of the envelope. As a result, the information contained within the original fine structure was left unaffected. The position of the envelope of each wavelet was shifted independently by a random value chosen from a uniform distribution ranging from $-T_D/2$ to $+T_D/2$. The degree of temporal distortion T_D is expressed in terms of the duration of the wavelets (inversely proportional to the bandwidth). If T_D equals two wavelets, the maximal displacement along the time axis is one effective duration of the wavelet from its original position. At 1 kHz, this is 5.76 ms; at 4 kHz, this is 1.55 ms.

4. Degradation of the spectral accuracy

To degrade the accuracy of the spectral information, the position of each wavelet was shifted randomly along the spectral axis (spectral perturbation). The positions of all wavelet coefficients were shifted independently by a random value chosen from a uniform distribution ranging from $-F_D/2$ to $+F_D/2$. The degree of spectral distortion F_D is expressed in octaves. If $F_D=0.5$ octaves, the maximal displacement along the frequency axis is 0.25 octaves (equals the effective bandwidth of the analysis window).

After wavelet decomposition, the spectral information of the signal is not only encoded in the position of the wavelets along the spectral axis, but also in the phase of the coefficients. The relative phases of the coefficients in each frequency band contain information about the spectral structure within this band. The random shifts of the wavelet positions along the spectral axis result in a smeared spectrum over bands. However, if the phase is kept intact, part of the spectral information within a band is reintroduced in the overlap-add procedure by interactions between neighboring wavelets.[3] By distorting the phase information we tried to bypass this problem. The phase was distorted by a desynchronization of the regular pattern of the wavelet coefficients along the temporal axis. This desynchronization was obtained by shifting the position of each wavelet (envelope plus fine structure) along the temporal axis by a random value chosen from a uniform distribution ranging from -0.0375 to $+0.0375$ of the wavelet bandwidth. In all conditions with spectral distortion including the spectral reference condition (0-octaves spectral perturbation), the phase was distorted in this way.

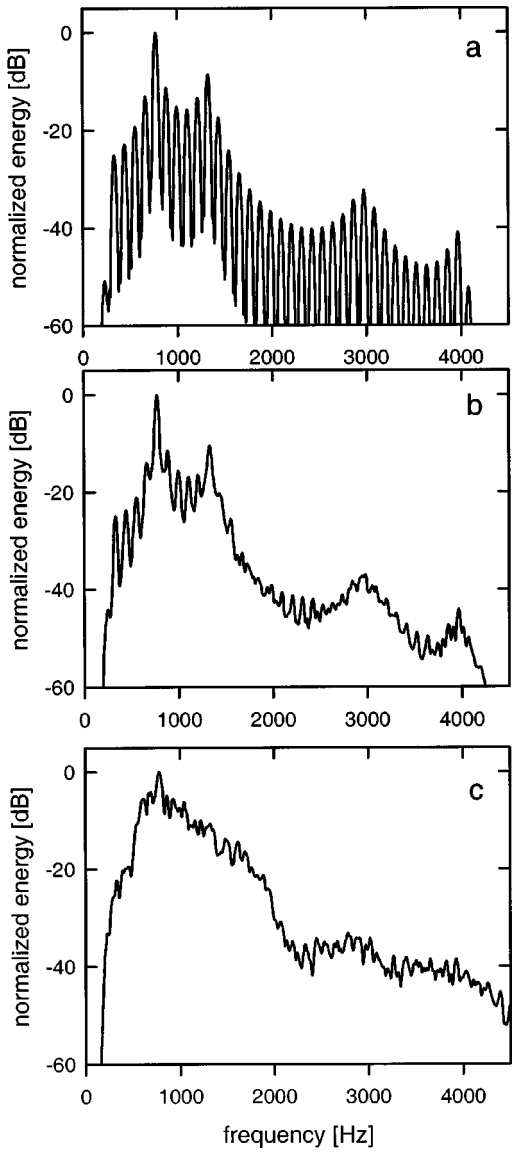


FIG. 3. The effect of the artificial distortion of the spectral information on an artificial vowel/a/: (a) undistorted vowel, (b) spectral reference condition (phase distorted), (c) spectral perturbation of 0.75 octaves (phase distorted and spectrally perturbed).

In Fig. 3, the effect of distorting the spectral information of an artificial vowel /a/ is illustrated. Panel (a) shows the undistorted vowel. In panel (b), the vowel is plotted in the spectral reference condition. In this condition, the phase of the complex coefficients is distorted, but the positions of the wavelets along the spectral axis are retained. As a result, most of the spectral fine structure is lost, but the spectral envelope is intact. In panel (c), the vowel is plotted in the most severe spectral distortion condition used in this study,

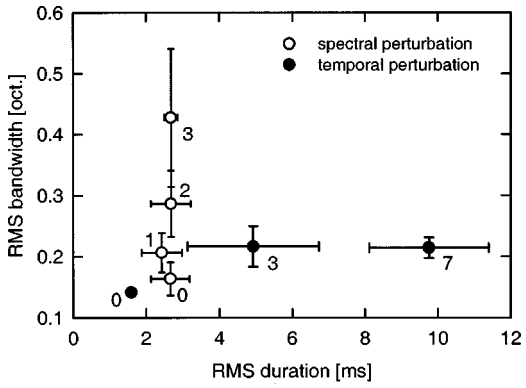


FIG. 4. The effect of the nondeterministic perturbation process on the rms duration (footnote 4) and rms bandwidth of a Gaussian-windowed tone with a center frequency of 1 kHz and a shape factor of 0.1735, i.e., an effective bandwidth of $\frac{1}{4}$ octave. Closed and open symbols represent the values corresponding to temporal perturbation and spectral perturbation, respectively. The numbers represent the degree of perturbation (expressed in the number of wavelets). The error bars represent the standard deviation.

i.e., when F_D equals 0.75 octaves. The phase is distorted as in the reference condition, and in addition the wavelets were shifted randomly over a maximum of $F_D/2$ along the spectral axis. As a result, the spectral envelope is smeared almost fully. Thus the overall spectral effect of the applied spectral uncertainty is a broadening of the spectral peaks.

As mentioned in Sec. I, degradation of the accuracy of the information of one domain is not possible without collateral degradation of the information of other domains. For example, the degradation of the accuracy of the intensity information also affects the spectral and temporal content of a signal. The effects of distortion of temporal information on spectral information and vice versa are illustrated in Fig. 4 for a Gaussian-windowed sinusoid (center frequency $=1$ kHz; $\alpha=0.1735$). This Gaussian tone pulse was wavelet decomposed, followed by spectral or temporal degradation of the wavelet coefficients, and recomposed. For recomposition of the unperturbed signal, the original wavelet coefficients were used. Then, reconstruction is perfect (see Sec. II A 1). The rms duration[4] and rms bandwidth of this signal are indicated by the closed circle with index ''0.'' If the wavelet coefficients are perturbed, the reconstructed signal is different from the original signal. Temporal or spectral perturbation of the coefficients increase the duration and bandwidth of the reconstructed signal. The effects of temporal perturbation on the duration and bandwidth of the signal are represented by the other closed circles; the effects of the spectral perturbation are indicated by open circles. For each degree of distortion, the perturbation procedure was applied to the input signal six times. Since the perturbation values are random values chosen from a uniform distribution, for equal degrees of distortion the increase will not be exactly the same. The error bars represent the standard deviations of the resulting duration and bandwidth of the output signals.

Looking at the effect of temporal perturbation, it can be observed that, when a temporal perturbation of three wavelets is applied, both the rms duration and rms bandwidth of the Gaussian tone pulse increase. For the seven-wavelet condition, the rms duration is longer than in the three-wavelet condition, but the rms bandwidth is the same. Thus for temporal perturbation up to three-wavelets, both the spectral and the temporal contrasts of sound are reduced. At that point, the spectral smearing reaches a maximum of about 0.25 octaves. Beyond that, temporal perturbation only reduces the temporal contrasts while the spectral contrasts stay unaltered.

With respect to spectral perturbation, it should be noted that in all spectral conditions the phase was distorted. As a result, the duration and bandwidth of the Gaussian-windowed sinusoid in the spectral reference condition (open circle ''0'') are larger than the duration and bandwidth of the original signal (closed circle ''0''); the spectral reference condition is slightly spectro-temporally smeared. The effect of additional spectral perturbation is just a reduction of the spectral contrasts, while the resulting (after phase distortion) temporal contrasts are maintained.

B. Subjects

Twelve normal-hearing listeners, aged 20–63 years with a mean age of 26 years, participated in the experiment. Pure-tone air-conduction thresholds of the normal-hearing listeners did not exceed 15 dB HL at any octave frequency from 250 to 4000 Hz. In addition, twenty-six sensorineurally hearing-impaired listeners took part in the experiment, aged 24–67 years with a mean age of 48 years.[5] Their intelligibility scores for monosyllabic words in quite were at least 75% correct. The pure-tone, air-conduction threshold in the hearing-impaired listener's better-hearing ear was at least 30 dB HL at one or more frequencies between 250 and 4000 Hz. Thresholds of the better-hearing ear averaged over 0.5, 1, and 2 kHz (the pure-tone average, or PTA) ranged from 17 to 70 dB HL, with a mean PTA of 50 dB HL. All listeners were native Dutch speakers.

C. Stimuli and apparatus

The speech stimuli consisted of sentences and words. The sentence sets contained lists of 13 everyday Dutch sentences of eight to nine syllables read by a female and male speaker (Versfeld *et al.*, 2000). The word sets consisted of lists of balanced meaningful CVC words (Bosman and Smoorenburg, 1995).

Signals were played out over TDT (Tucker Davis Technologies) System II hardware. Stimuli were presented in the middle of the dynamic range of each listener by frequency shaping them using a programmable filter (TDT PF1). The stimuli were presented monaurally through Sony MDR-V900 headphones. To avoid the risk of cross hearing, the listener's better-hearing ear was tested. For calibration, sound pressure levels of the stimuli were measured on a Brüel & Kjær type 4152 artificial ear with a flat-plate adapter. The entire experiment was controlled by a personal computer. Subjects were tested individually in a soundproof room.

D. Procedures

First, the hearing threshold and the uncomfortable loudness level (UCL) of each listener were determined. In the detection and intelligibility tests, sounds were adapted to fit

the dynamic range of each listener. To familiarize the subjects with the procedure, a training session preceded data collection. All conditions were measured twice in order to determine measurement reliability. Speech intelligibility tests were performed once using sentences spoken by the female talker and once using those by the male talker. In the distortion-sensitivity model, the performance for individual hearing-impaired listeners is compared with that for normal-hearing listeners. Therefore, for all listeners, the same order of conditions and sentence lists was used.

1. Threshold and UCL

The dynamic range of each listener was estimated by measuring the hearing threshold and the uncomfortable loudness level (UCL) for narrow bands of noise. The UCL was corrected for broadband stimulation, as described in the following.

Thresholds and UCLs were measured using 1/3-octave noise bands at center frequencies of 250, 500, 1000, 2000, and 4000 Hz. Hearing thresholds were measured using a Békésy tracking (Yantis, 1994) procedure (300-ms noise bursts; repetition rate 2.5 Hz; step size 1 dB). The measurement was ended after 11 level reversals. The average of all but the first reversal level was taken as the hearing threshold. Narrow-band UCLs were measured with 1/3-octave noise bursts that were presented with a 3-dB increase in level for each presentation (300-ms noise burst; repetition rate 1.4 Hz). Listeners were asked to press a button when the noise bursts became uncomfortably loud. Then, the level of the noise burst was immediately diminished by a random amount between 21 and 30 dB, and the ascending procedure was repeated until six responses were obtained. The average of the levels at which the button was pushed was taken as the narrow-band UCL.

To correct the UCL for broadband stimulation, a 4-s broadband noise burst was presented, spectrally shaped according to the narrow-band UCLs and starting 40 dB below the narrow-band UCLs. The level of the broadband noise burst was gradually increased in steps of 5 dB. After each presentation the listener was asked whether the signal was experienced as uncomfortably loud. If this was the case, the corresponding level was taken as the broadband UCL.

2. Detection threshold for distortion

The detection thresholds for the distortion of intensity, temporal, or spectral information were estimated using words. A 3I-3AFC two-down one-up adaptive procedure was used, leading to a 70.7% correct score. In each trial, the subject was presented with three signals, twice the reference word and once the distorted word. The listener had to point out the distorted signal. For each trial, a random choice out of 90 bandpass filtered (250–4000 Hz) preprocessed (at different degrees of distortion) words was loaded from disk. The difficulty of the task was increased by dividing the distortion factor by $\sqrt{2}$ following two consecutive correct responses; the difficulty of the task was decreased by multiplying the distortion factor by $\sqrt{2}$ following one incorrect response. A transition from increasing to decreasing difficulty or vice versa defined a reversal. A run was ended after 20 reversals. The geometric mean of the last 16 reversals was used as an estimate of the detection threshold for distortion. To define the experiment with respect to presentation level, all words were presented in the middle of the dynamic range of the listener, in noise with a speechlike spectrum (Wandel and Goltermann RG-1) at a signal-to-noise ratio of 15 dB.

3. Speech intelligibility

a. Speech-reception threshold in noise for an adapted spectrum (SRTa). The speech-reception threshold (SRT) is an estimate of the ability to perceive speech in daily life (Plomp and Mimpen, 1979). The SRT in noise is defined as the signal-to-noise ratio (SNR) at which 50% of the sentences are reproduced correctly. The speech level is varied in an adaptive, up–down procedure with a step size of 2 dB. A continuous stationary noise is presented from 500 ms before to 500 ms after each sentence. In our experiments, speech and noise are adapted to fit in the dynamic range of individual listeners. This adapted speech-reception threshold is called SRTa. In the SRTa tests in this study, all stimuli were bandpass filtered from 250 to 4000 Hz.

After a SRT test using undistorted speech, the SRTa was measured as a function of the degree of distortion (distortion-sensitivity model). The intensity-distortion conditions were 0 (undistorted), 10, and 20 dB. The temporal-distortion conditions were 0 (undistorted), 3, and 7 wavelets. The spectral-distortion conditions were 0, $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{3}{4}$ octave (recall that wavelet phases were distorted in all spectral-distortion conditions).

b. Speech-reception bandwidth threshold (SRBT). In addition to the SRTa, the speech-reception *bandwidth* threshold (SR*B*T) was measured to estimate suprathreshold speech processing. The SR*B*T measure of speech intelligibility was introduced by Noordhoek *et al.* (1999). The SR*B*T is highly sensitive for suprathreshold deficits, as shown in a recent study of Noordhoek *et al.* (2000).

The SR*B*T procedure is similar to the SRT procedure, except that the bandwidth (center frequency: 1 kHz) of the undisturbed speech is varied instead of the level when estimating the 50% intelligibility threshold. Complementary shaped bandstop noise is added to the bandpass-filtered speech. Speech and noise are presented in the middle of the listener's dynamic range.

E. Speech intelligibility index

To estimate the quality of speech processing of listeners, the SRTa and SR*B*T data were converted to a speech intelligibility index. The speech intelligibility index (SII) (ANSI, 1997) is a physical measure of how much information of speech is available to the listener. The SII correlates highly with speech intelligibility. To perceive speech, normal-hearing listeners need a certain amount of information which can be converted to a SII value. If hearing-impaired listeners need more information, this suggests that their speech processing is degraded. Thus elevated SII values are an indication for a low speech processing quality. The SII model ac-

counts for hearing threshold, self-masking in speech, normal upward spread of masking, and level distortion at high presentation levels. To calculate the SII, speech spectra, noise spectra, and hearing thresholds must be known. As mentioned in Sec. II D 1, hearing thresholds were measured with 1/3-octave noise bands, using Békésy tracking (Yantis, 1994). This procedure probably results in hearing thresholds that are systematically about 4 dB higher than the methods on which the ISO (1961) threshold is based (Noordhoek *et al.*, 2000; Noordhoek *et al.*, 2001). Therefore, in the SII calculations the internal noise level was lowered by 4 dB. The band-importance function for speech material of average redundancy (Pavlovic, 1987) was used.

III. RESULTS AND DISCUSSION

A. Detection thresholds

To obtain insight into which attributes of sound processing are distorted for hearing-impaired listeners, detection thresholds for the distortion of intensity, time, and frequency information were measured. If the auditory coding of a particular type of information is degraded, the detection thresholds for the distortion of this type of information will probably be higher.

1. Degradation of the intensity accuracy

For the normal-hearing listeners, the detection thresholds for the intensity perturbation, described in Sec. II A 2, ranged from 13 to 23 dB, with a median of 17 dB. For the hearing-impaired listeners, the detection thresholds ranged from 9 to 53 dB, with a median of 18 dB. The overall (normal-hearing plus hearing-impaired listeners: 38 subjects) mean standard error of an individual detection threshold (two measurements) was 3 dB. Individual detection thresholds are shown in Fig. 5(a). This figure will be explained in more detail in Sec. III B. Some hearing-impaired listeners had detection thresholds that were much larger than those for the normal-hearing listeners, but a Mann–Whitney *U* test showed that the difference in detection threshold between the group of normal-hearing and the group of hearing-impaired listeners was not significant.

2. Degradation of the temporal accuracy

For the normal-hearing listeners, the detection thresholds for temporal perturbation ranged from 0.9 to 1.5 wavelets, with a median of 1.1 wavelets; for the hearing-impaired listeners, the thresholds ranged from 0.6 to 7.4 wavelets, again with a median of 1.1 wavelets. The mean standard error of an individual detection threshold was 0.4 wavelets. Individual detection thresholds are shown in Fig. 5(b). This figure will be explained in more detail in Sec. III B. A few hearing-impaired listeners had detection thresholds that were much larger than those for the normal-hearing listeners, but a Mann–Whitney *U* test showed that the detection thresholds for the group of hearing-impaired listeners were not significantly higher than those for the group of normal-hearing listeners.

3. Degradation of the spectral accuracy

For the normal-hearing listeners, the detection thresholds for spectral perturbation ranged from 0.22 to 0.39 octave, with a median of 0.26 octave. For the hearing-impaired listeners, the detection thresholds ranged from 0.17 to 1.4 octave, with a median of 0.36 octave. The mean standard error of the individual detection threshold was 0.06 octave. Individual detection thresholds are shown in Fig. 5(c). This figure will be explained in more detail in Sec. III B. A Mann–Whitney *U* test showed that the detection thresholds for the group of the hearing-impaired listeners were significantly ($p<0.05$) higher than those for the normal-hearing listeners.

In summary, with respect to the detection of distortion of intensity and temporal information, no significant difference was observed between the group of normal-hearing and the group of hearing-impaired listeners. With respect to the detection of spectral distortion, a significant difference between normal-hearing and hearing-impaired listeners was observed. Thus spectral cues were probably less clearly perceived by the hearing-impaired listeners.

B. Suprathreshold speech intelligibility

The aim of this study is to gain insight into the suprathreshold speech processing problems of hearing-impaired listeners. Therefore, speech processing performance was measured by means of the SRTa and SR*B*T tests. For the normal-hearing listeners, the SRTa ranged from -1.8 to 0.3 dB, with a median of -0.8 dB. For the hearing-impaired listeners, the SRTa ranged from -1.1 to 8.5 dB, with a median of 2.0 dB. The mean standard error of an individual SRTa (six measurements) was 0.7 dB. The hearing-impaired listeners had significantly higher SRTa's than the normal-hearing listeners (Mann–Whitney *U* test: $p<0.05$). The SR*B*T for the normal-hearing listeners ranged from 1.1 to 1.7 octave, with a median of 1.6 octave. The SR*B*T for the hearing-impaired listeners ranged from 1.5 to 3.4 octave, with a median of 2.1 octave. The standard error of an individual SR*B*T (two measurements) was 0.3 octave. The hearing-impaired listeners had significantly higher SR*B*T values than the normal-hearing listeners (Mann–Whitney *U* test: $p<0.05$).

For both the SRTa and the SR*B*T tests, hearing-impaired listeners performed worse than normal-hearing listeners, which confirms the problems hearing-impaired listeners have in perceiving speech. To quantify the degree of deterioration of suprathreshold speech processing, the individual SRTa and SR*B*T data were converted to SII units. For the normal-hearing listeners, the SII for the SRTa ranged from 0.36 to 0.42, with a median of 0.39; the SII for the SR*B*T ranged from 0.26 to 0.39, with a median of 0.35. For the hearing-impaired listeners, the SII for the SRTa ranged from 0.37 to 0.54, with a median of 0.43; the SII for the SR*B*T ranged from 0.32 to 0.52, with a median of 0.43. The individual standard error of the SII_{SRTa} (six measurements) was 0.02. The individual standard error of the SII_{SRBT} (two measurements) was 0.05. Both the SII_{SRTa} and the SII_{SRBT} for the hearing-impaired listeners were significantly higher than

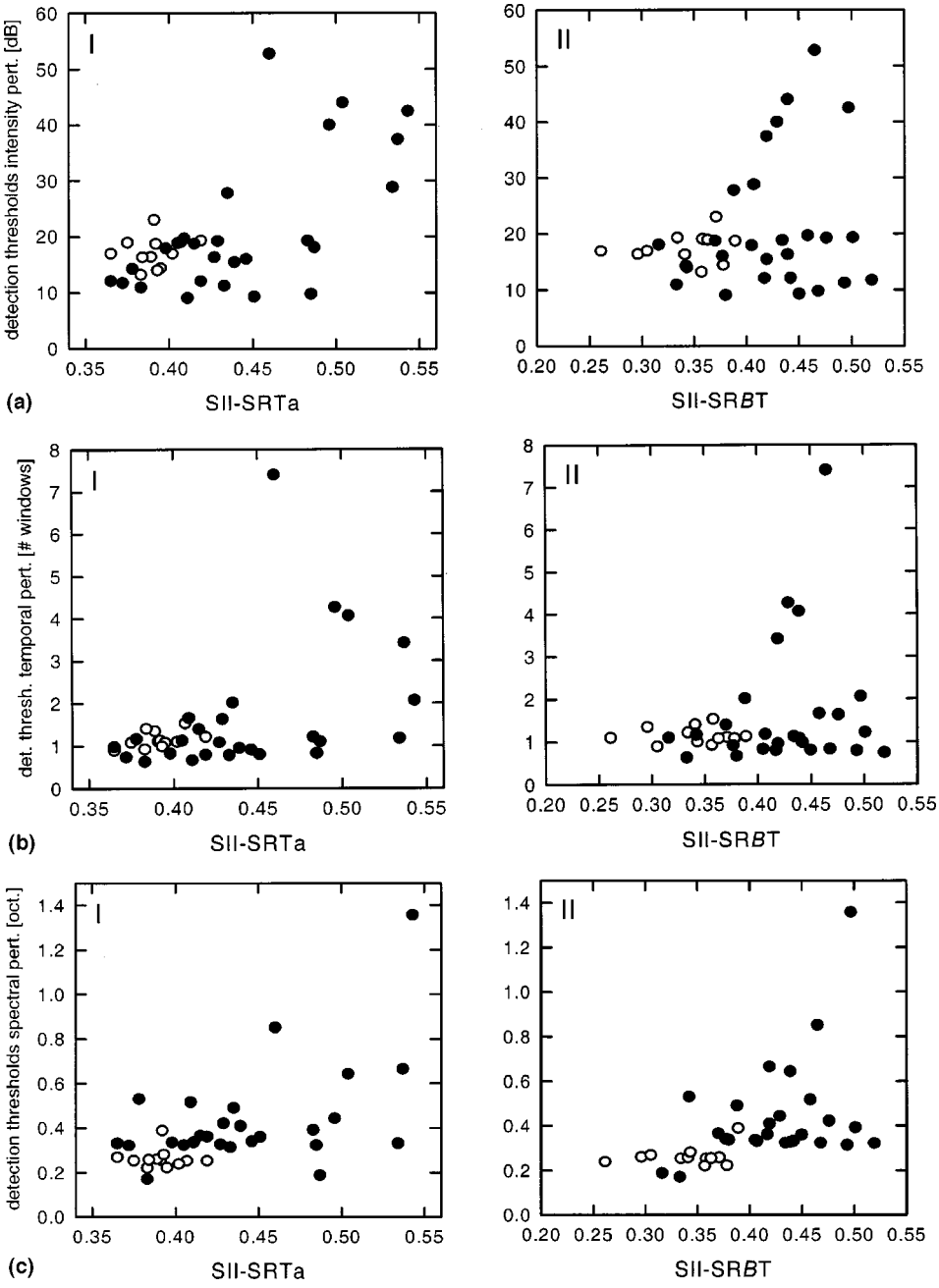


FIG. 5. Individual detection thresholds for intensity (a), temporal (b), and spectral (c) perturbation vs the speech intelligibility index (SII) corresponding with the mean of SRTa scores (panel I) and SR*B*T scores (panel II) for normal-hearing listeners (open circles) and hearing-impaired listeners (closed circles).

those for the normal-hearing listeners (Mann–Whitney *U* test; $p<0.05$).

The SII values of the hearing-impaired listeners indicate that their suprathreshold speech processing is clearly distorted. The next step is to explore what aspects of auditory coding are distorted. The detection threshold experiments suggest that hearing-impaired listeners perceive spectral information less clearly than normal-hearing listeners. In Fig. 5(c) the individual detection thresholds for spectral perturbation are plotted as a function of the SII_{SRTa} (panel I) and as a function of the SII_{SRBT} (panel II). Open symbols represent the detection thresholds for the normal-hearing listeners, closed symbols those for the hearing-impaired listeners. The figure shows a correlation between the SIIs and the detection threshold for spectral perturbation. A statistical analysis (Spearman rank correlation) on the combined data for the normal-hearing and hearing-impaired listeners confirmed this: There is a significant ($p<0.05$) correlation of 0.5 between the detection threshold for spectral perturbation and SII_{SRTa}, and a significant ($p<0.05$) correlation of 0.6 between the detection threshold and SII_{SRBT}.

Also correlations between the SIIs and the detection threshold for intensity and temporal perturbation were considered. The individual detection thresholds for intensity and temporal perturbation are plotted in Figs. 5(a) and (b), respectively. Details are the same as in Fig. 5(c). The Spearman rank correlation of the combined data for the normal-hearing and hearing-impaired listeners between the SII_{SRTa} and the detection thresholds for intensity and temporal perturbation were significant. Both were 0.4 ($p<0.05$). The Spearman rank correlations between the SII_{SRBT} and the detection thresholds for intensity and temporal perturbation were not significant.

Summarizing, a correlation between the detection

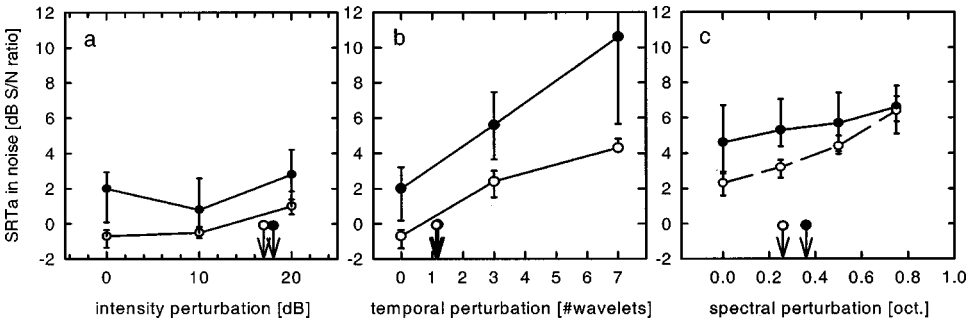


FIG. 6. The median of SRTa values for normal-hearing (open symbols) and hearing-impaired listeners (closed symbols) as a function of distortion. The error bars represent the interquartile ranges. Arrows indicate the median of the detection threshold for each distortion for the normal-hearing listeners (open circle) and hearing-impaired listeners (closed circle): (a) distortion of intensity information; (b) distortion of temporal information; (c) distortion of spectral information.

threshold for the distortion of spectral information and the SII was observed. With respect to the detection thresholds of intensity and temporal perturbation, only the correlation with the SII_{SRTa} was significant. Thus distorted processing of spectral information by hearing-impaired listeners relates statistically to their speech-processing deficits. With respect to the processing of intensity and temporal information, this is less clear. In Sec. III C, the relation between the auditory coding accuracy and reduced speech intelligibility is analyzed in a more direct way by means of the distortion-sensitivity model.

C. Distortion-sensitivity model: Group results

Applying the distortion-sensitivity model, the SRTa was measured as a function of the artificial degradation of the spectro-temporal coding of sound, for normal-hearing and hearing-impaired listeners. The results are plotted in Fig. 6. The SRTa is plotted as a function of the degree of distortion of intensity information [panel (a)], temporal information [panel (b)], and spectral information [panel (c)]. Open and closed circles represent the medians of the data for the normal-hearing and hearing-impaired listeners, respectively. The bars represent the interquartile ranges. The arrows represent the medians of the detection thresholds for normal-hearing (open circle) and hearing-impaired listeners (closed circle).

1. Degradation of the intensity accuracy

For all levels of intensity degradation, the hearing-impaired listeners perform poorer than the normal-hearing listeners on the speech intelligibility tests [Fig. 6(a)]. The difference in performance between normal-hearing and hearing-impaired listeners appears to decrease somewhat as a function of the intensity distortion. However, a Mann–Whitney *U* test showed that this effect was not significant. This is in agreement with the lack of a significant difference in detection thresholds for intensity distortion between normal-hearing and hearing-impaired listeners (Sec. III A 1; medians of the groups represented by arrows). The absence of a difference in sensitivity between normal-hearing and hearing-impaired listeners could be the result from the low perturbation levels used in this study. Higher intensity distortion levels were not measured, because this leads to unwanted spectro-temporal by-products of the signal processing (see van Schijndel *et al.*, 2001). In conclusion, the results do not show a relation between reduced speech intelligibility in noise and a distorted representation of intensity information.

2. Degradation of the temporal accuracy

For all levels of temporal degradation, the medians of the SRTa's for the hearing-impaired listeners are higher than those for the normal-hearing listeners [Fig. 6(b)]. The performances of normal-hearing and hearing-impaired listeners certainly do not converge as a function of temporal perturbation. Instead the performances seem to diverge. This divergence may be related to the fact that the hearing-impaired listeners have less information available from the other, nonperturbed cues. Let's explain this by an example.

Assume as an extreme example that a hearing-impaired listener cannot use the spectral information in speech, but his/her processing of temporal information is as good as that of normal-hearing listeners. When all temporal information is removed from the speech, this hearing-impaired listener cannot understand the speech. The reason is that he/she is deprived of both spectral and temporal cues. Normal-hearing listeners will also be bothered by the removal of temporal information. However, they can still use the spectral information. Thus, looking at the effect of distortion of temporal information on speech intelligibility, performance of this hearing-impaired listener will diverge compared to normal-hearing listeners. However, this divergence does not indicate that this hearing-impaired listener has problems to perceive temporal information. It simply indicates that this hearing-impaired listener has less information available from the other, nonperturbed, cues.

To summarize, the difference in performance between normal-hearing and hearing-impaired listeners does not decrease as a function of temporal perturbation. Actually, the divergence suggests that other cues, not in the temporal domain, are processed less efficiently. In addition, the group of hearing-impaired listeners performed as well as the normal-hearing listeners on the temporal perturbation detection task (Sec. III A 2). In conclusion, the results do not suggest a relation between reduced intelligibility in noise and a distorted representation of temporal information.

3. Degradation of the spectral accuracy

For the most extreme spectral perturbation condition, only the results using the male talker are used, because the male talker was just intelligible in this condition while the female talker was not [see Fig. 6(c)]. The SRTa for the normal-hearing listeners in the spectral reference condition is about 3 dB (median value: 3.1 dB) higher than in the intensity and temporal reference conditions, because the fine structure was perturbed in all spectral conditions (Sec.

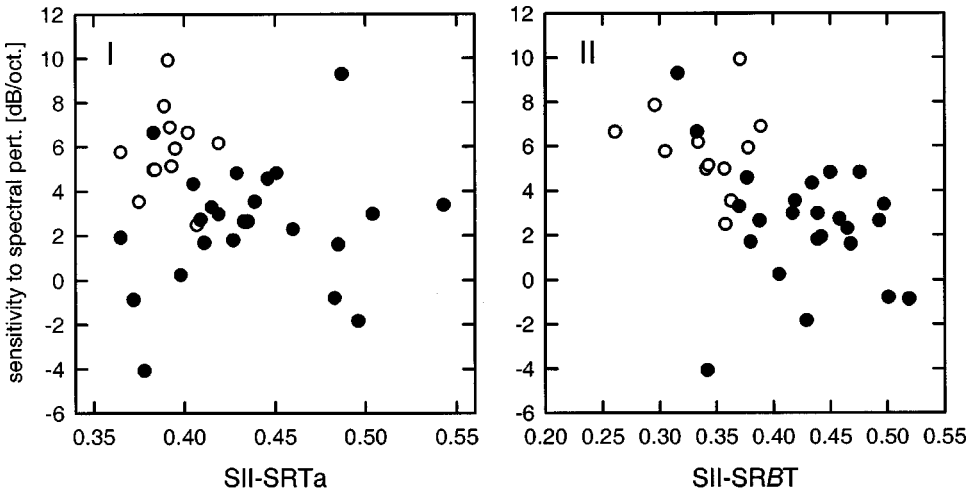


FIG. 7. The individual sensitivities to spectral perturbation for normal-hearing (open symbols) and hearing-impaired listeners (closed symbols) vs the SII_{SRTa} (panel I) and vs the SII_{SRBT} (panel II).

II A 4). The SRTa for the hearing-impaired listeners in the spectral reference condition is about 2.5 dB (median value: 2.4 dB) higher than in the intensity and temporal reference conditions. The difference in SRTa for original speech and speech without fine structure tells something about the role of fine structure. For hearing-impaired listeners this elevation is slightly less than for normal-hearing listeners, but this is not statistically significant. This suggests that perception of fine structure does not relate to speech-perception problems of hearing-impaired listeners.

In the reference condition the median SRTa is higher for the hearing-impaired listeners than for the normal-hearing listeners. When spectral perturbation is applied, the performance for the hearing-impaired listeners converges toward that for the normal-hearing listeners. At $\frac{3}{4}$ octave of spectral perturbation, the performance for the hearing-impaired listeners equals that for the normal-hearing listeners. Mann–Whitney *U* Tests confirm the observed trends: at 0- and $\frac{1}{4}$-octave perturbation the performance for the hearing-impaired listeners is significantly worse than that for the normal-hearing listeners ($p<0.05$), whereas at $\frac{1}{2}$ and at $\frac{3}{4}$ octave no significant difference exists.

As already shown in Sec. II A, intensity, temporal, and spectral information cannot be manipulated completely independently. Perturbation in one domain will also affect the other domains. However, as shown in Figs. 3 and 4, the spectral perturbation as applied in this study only had a negligible effect on the other domains. Therefore, it seems reasonable to assume that effects present in Fig. 6(c) can be accounted for by a distortion of only the spectral information in speech.

In this study, interdependency of intensity, time, and frequency was only considered from a signal-processing point of view. Also, another interdependency may exist, i.e., within the auditory system. Indeed, a recent study of Loizou *et al.* (1999) demonstrated an interaction between intensity accuracy (expressed in terms of number of quantization levels) and spectral accuracy (expressed in terms of number of spectral channels). This interdependency was not investigated in this study.

In summary, the detection threshold for spectral perturbation is significantly higher for hearing-impaired listeners than for normal-hearing listeners; moreover, convergence of the speech-processing performance of normal-hearing and hearing-impaired listeners is observed. This strongly points to a relation between a reduced intelligibility in noise and a distorted representation of spectral information.

D. Distortion-sensitivity model: Individual results

In the preceding section, the group results of the distortion-sensitivity model for normal-hearing and hearing-impaired listeners were compared. Now, the individual results will be used to further examine the relation between distorted coding of information and reduced speech intelligibility. As an estimate of individual performance, the sensitivity to the distortion is taken. The sensitivity to the distortion of individual listeners is defined as the slope of the linear regression line fitted through the individual SRTa values for different degrees of distortion. It quantifies how sensitive a listener is to the distortion of specific cues in speech. The underlying idea is that if a hearing-impaired listener is less sensitive to a particular artificial distortion than normal-hearing listeners, this artificially applied distortion probably relates to the internal deficit causing his/her speech-perception problems. In this study two measures for suprathreshold speech-perception quality are used: SII_{SRBT} and SII_{SRTa}. The relation between speech-perception quality and the sensitivity to distortion of information will be evaluated.

For intensity information, no correlation between the sensitivity to the distortion and SII_{SRTa} or SII_{SRBT} was observed in the individual data: the Spearman rank correlation between sensitivity to intensity distortion and SII was -0.3 ($p=0.09$) for the SII_{SRTa}, and -0.3 ($p=0.1$) for the SII_{SRBT}.[6] For temporal information, sensitivities were not considered, because of divergence (see Sec. III C 2).

In Fig. 7, the sensitivity to distortion of spectral information is plotted against the individual SII_{SRTa} (panel I) and SII_{SRBT} (panel II). Open symbols represent the data for the normal-hearing listeners; closed symbols those for the hearing-impaired listeners.[7] As is already clear from Fig. 6(c), the median sensitivity of the hearing-impaired listeners is less than that of the normal-hearing listeners. No clear trend between SII_{SRTa} and sensitivity is shown [Spearman

rank correlation: -0.2 $(p=0.2)$]; however, there is a trend between SII_{SRBT} and sensitivity: The higher the SII_{SRBT}, the lower the sensitivity to spectral distortion [Spearman rank correlation: -0.6 $(p<0.05)$].[8]

SII_{SRTa} and SII_{SRBT} show a different picture: The sensitivity to spectral distortion is significantly correlated with the SII_{SRBT}, but not with the SII_{SRTa}. This difference may be explained by the different experimental setup: The speech-reception *bandwidth* threshold is measured using bandpass filtered speech signals embedded in complementary bandstop noise, whereas the speech-reception threshold test uses a noise spectrum equal to the average speech spectrum. Therefore, the SR*B*T is probably more sensitive to excessive spread of masking than the SRTa. As a result, the sensitivity to spectral distortion is likely to relate more directly to the SII_{SRBT} than to the SII_{SRTa}.

In summary, the individual results show a relation between suprathreshold speech processing as quantified by the SII_{SRBT} and the sensitivity to spectral distortion. This is in agreement with the observed relation between speech processing quality and the detection threshold for spectral perturbation (Sec. III B), and the observed convergence of the performance for normal-hearing and hearing-impaired listeners for increasing degrees of spectral distortion (Sec. III C 3). These results suggest that the auditory processing of spectral information of hearing-impaired listeners is distorted and that this affects speech perception. The poorer the spectral coding, the more problems hearing-impaired listeners have in perceiving speech.

The question remains whether distorted spectral auditory coding is the only cause of suprathreshold speech-processing deficits. A considerable variance is present in the data of Fig. 7. This may be the result of measurement error, but this may also be variance due to factors other than distorted coding of spectral information. By calculating the reliability (Nunnally, 1967) of the variables in the correlation, an estimate of the influence of measurement error can be made. The square root of the product of the reliabilities of two tests gives an estimate of the unsigned maximum correlation possible, given the measurement accuracy.

The reliability of the SII_{SRTa} (six measurements) is 0.9. The reliability of the sensitivity to the distortion is much smaller: about 0.3. This is because the measurement errors add up when the slope is estimated. Between SII_{SRTa} and sensitivity, the maximum unsigned correlation possible is about 0.5. The correlation observed was -0.2. Thus in the speech processing problems of hearing-impaired listeners as quantified by the SII_{SRTa}, spectral cues are probably not the only ones.

The reliability of the SII_{SRBT} (two measurements) is 0.7. As a result, the estimate of the unsigned maximum correlation possible between SII_{SRBT} and sensitivity is 0.5. The correlation observed was -0.6. It may surprise that the absolute value of the observed correlation is larger than the predicted maximum correlation. However, the predicted maximum correlation is only a rough estimate. Therefore, all variance seems explained.

In summary, the distorted speech processing of hearing-impaired listeners measured by the SR*B*T test can be fully explained by distorted processing of spectral information, but with respect to the SRTa test other factors seem to affect intelligibility as well. This may be explained by the fact that upward spread of masking plays a dominant role in the SR*B*T test, but not in the SRTa test.

E. Comparison to literature

1. Degradation of the intensity accuracy

The median detection threshold for intensity distortion of hearing-impaired listeners is not significantly higher than that of normal-hearing listeners. However, some hearing-impaired listeners showed abnormally high distortion thresholds. This is consistent with the literature about intensity discrimination (for a review, see Florentine *et al.*, 1993). Overall, hearing-impaired listeners discriminate as well as normal-hearing listeners at equal sound pressure levels, and intensity discrimination may even be better at equal sensation levels. However, for some hearing-impaired listeners markedly higher discrimination thresholds are observed (Schroder *et al.*, 1994; Buus *et al.*, 1995).

With respect to speech intelligibility as a function of intensity distortion, no significant convergence of the performances for normal-hearing and hearing-impaired listeners was observed. In addition, no significant correlation between the sensitivity to intensity distortion and the SII was found. In contrast, in van Schijndel *et al.* (2001) a significant correlation between sensitivity to intensity distortion and SII_{SRBT} was observed. Several factors may account for this. Different listener groups were used in the previous and the present study. Since among hearing-impaired listeners a diversity of auditory deficits is observed (see, e.g., Noordhoek *et al.*, 2001), this may lead to a different result. Moreover, although both groups of hearing-impaired listeners had comparable hearing loss, the presentation levels for the second group of listeners was 7 dB lower than for the first group due to lower uncomfortable loudness levels. Due to this difference in dynamic range, the same intensity perturbations may have introduced different loudness perturbations (see van Schijndel *et al.*, 2001). These factors may explain why the correlation in the present study is not significant while in the previous study it was.

2. Degradation of the temporal accuracy

The median detection threshold for temporal distortion of hearing-impaired listeners was not significantly higher than that of normal-hearing listeners. However, some hearing-impaired listeners showed abnormally high detection thresholds. This is in agreement with the literature about temporal resolution. Temporal-resolution deficits occur in some hearing-impaired listeners and not in others (see, e.g., Noordhoek *et al.*, 2001). Whether or not hearing-impaired listeners show temporal-processing deficits also depends on the temporal-resolution test that is used. On some tests of temporal resolution, most hearing-impaired listeners perform as well as normal-hearing listeners (Moore, 1995). Other tests clearly show that hearing-impaired listeners suffer from, for example, excessive forward masking (Festen and Plomp, 1983; Oxenham and Moore, 1995).

The performances of normal-hearing and hearing-impaired listeners did not converge as a function of the distortion of temporal information. This agrees with the study of Duquesnoy and Plomp (1980). They measured how sensitive normal-hearing and hearing-impaired listeners were to reverberation. Reverberation can be considered a very systematic type of distortion of temporal information. The sensitivity of the listeners to reverberation was compared to the Speech Transmission Index (Houtgast and Steeneken, 1973). Their results showed that hearing-impaired listeners were as sensitive to reverberation as normal-hearing listeners.

Based on the previous text, one may conclude that distorted temporal processing is not a factor underlying poor speech intelligibility for the present group of hearing-impaired listeners. This conclusion seems to be in contrast with a recent study of Noordhoek *et al.* (2001). In this extensive study, relations between speech intelligibility and auditory functions in the 1-kHz frequency region were investigated. Results show that a factor related to ''reduced temporal resolution and reduced frequency discrimination seemed to relate to speech-processing deficits.'' As already mentioned (Sec. III D), in the present study not all variance can be explained by distorted spectral processing. Some variance remains unexplained. The underlying factor of this unexplained variance may be distorted temporal (or intensity) processing.

The question remains why in the present study reduced temporal resolution did not show up clearly, while in Noordhoek *et al.*'s study a factor related to temporal resolution and frequency discrimination did. Probably, two differences between Noordhoek's study and the present study may account for this: First, different listener groups were used in Noordhoek's and the present study. As already mentioned in Sec. III E 1, since among hearing-impaired listeners a diversity of auditory deficits is observed, different listener groups may lead to different results. Second, Noordhoek's study concentrated on the 1-kHz frequency region, while the present study looked at the total region from 250 Hz to 4 kHz. Problems related to reduced temporal resolution and/or frequency discrimination may be so frequency specific that looking at a broad frequency range obscures the problem.

3. Degradation of the spectral accuracy

The detection thresholds for spectral distortion were significantly higher for the group of hearing-impaired listeners than for the group of normal-hearing listeners. In addition, convergence of speech-perception performance for normal-hearing and hearing-impaired listeners as a function of spectral distortion was observed. This agrees with the results of Turner *et al.* (1995) that also showed convergence (see Sec. I). It is also in agreement with conclusions of the recent study of Noordhoek *et al.* (2001) that concludes that spectral resolution ''seemed to be related to suprathreshold speech deficits.''

The results of this study suggest that hearing-impaired listeners suffer from reduced frequency selectivity and that this causes reduced speech intelligibility. This agrees with the literature, in which it has been reported frequently that hearing-impaired listeners suffer from reduced spectral resolution. [For review see Tyler (1986).] Reduced frequency selectivity affects speech intelligibility in two ways. First, because of reduced frequency selectivity the spectral contrasts in speech itself are less clear. Second, when frequency selectivity is reduced, hearing-impaired listeners will suffer from excessive upward and downward spread of masking.

Ter Keurs *et al.* (1992), (1993) investigated the first effect. Speech and noise, having the same long-term average spectrum, were added *after* the smearing of the spectral envelope. As a result, the effect of excessive masking was not simulated. Ter Keurs *et al.* (1993) observed that hearing-impaired listeners were as sensitive to reduced spectral contrasts in speech as normal-hearing listeners. They did find a small but significant correlation between the SRT for unsmeared speech and auditory filter bandwidth, but they could not explain this by a reduction of the spectral contrasts in speech.

In our study, the first and second effects were evaluated in combination, because first the noise was added to the speech and then the spectral information was distorted. Our results strongly suggest that reduced frequency selectivity influences speech intelligibility in noise. Since the results of ter Keurs *et al.* (1993) suggest that the first effect is not responsible for reduced speech perception, the reduced speech intelligibility in noise observed in hearing-impaired listeners is probably mainly due to the second effect, i.e., excessive spread of masking. Thus for hearing-impaired listeners, it is more difficult to separate speech from competing background noise.

IV. SUMMARY AND CONCLUSIONS

In this study, the central question was how degraded speech perception of hearing-impaired listeners relates to distorted auditory coding. To investigate this, the intensity, time, and frequency information of sound were artificially distorted after wavelet coding. The detection thresholds for the different types of distortion were measured to obtain insight into how clearly hearing-impaired listeners could perceive a particular type of information. To investigate the relation between distorted auditory coding and speech perception, the distortion-sensitivity model was used. If hearing-impaired listeners are less sensitive with respect to speech perception than normal-hearing listeners to a particular type of distortion (intensity, time, or frequency), this indicates that this artificial distortion relates to the distorted auditory coding causing speech-perception problems.

The group results showed that the detection thresholds for hearing-impaired listeners with respect to the distortion of intensity and temporal information were not significantly higher than those for normal-hearing listeners. For the distortion of spectral information, the detection thresholds for the hearing-impaired listeners were significantly higher than those for the normal-hearing listeners. Thus hearing-impaired listeners may perceive spectral information less clearly than normal-hearing listeners. With respect to the distortion-sensitivity model, the results (Fig. 6) did not show that the group of hearing-impaired listeners was less sensitive than the group of normal-hearing listeners to intensity and temporal distortion. The group of hearing-impaired lis-

teners was less sensitive than normal-hearing listeners to the distortion of spectral information. Thus the group results suggest that distorted coding of spectral information is an important factor underlying the reduced speech intelligibility observed in hearing-impaired listeners.

Also, the individual results were considered to investigate the relation between reduced speech intelligibility and distorted coding of spectral information in more detail. A significant correlation between the SII, both SII_{SRTa} and SII_{SRBT}, and the detection threshold for spectral distortion was observed (Fig. 5). Thus the data reveal a statistical relation between the quality of speech processing, quantified by the SII, and the spectral coding accuracy, quantified by the detection threshold for spectral distortion. In addition, the correlation between the SII_{SRBT} and the sensitivity to spectral distortion with respect to speech perception was significant (Fig. 7). Thus there is a statistical relation between the quality of speech processing and the effect of distortion of the spectral cues on speech perception. The more pronounced the speech-perception problems of hearing-impaired listeners (in terms of the SII), the less accurate the spectral auditory coding (higher detection thresholds) and the less influence the distortion of spectral information has on speech intelligibility (lower sensitivity to spectral distortion). The individual results support the group result, strongly suggesting that distorted coding of spectral information is the factor underlying the suprathreshold problems encountered by many hearing-impaired listeners when trying to perceive speech.

The sensitivity to spectral distortion could explain all ''true'' variance in the SII_{SRBT}, i.e., all variance not due to measurement error. Thus distorted auditory coding of spectral information may be the only factor underlying speech-processing deficits measured by means of the SR*B*T test. However, sensitivity to spectral distortion could not explain all ''true'' variance in the SII_{SRTa}. This suggests that, besides distorted coding of spectral information, other factors play a role in the suprathreshold speech processing problems of hearing-impaired listeners as reflected in the SRTa test.

From the data of the present study the following general conclusions can be drawn.

(1) The distortion-sensitivity model may be a valuable tool to investigate the underlying causes of reduced speech perception.
(2) Distorted auditory spectral coding may be an important factor underlying the speech-perception problems of hearing-impaired listeners.
(3) Besides distorted coding of spectral information, other factors may play a role in reduced speech intelligibility as well.

ACKNOWLEDGMENTS

This work was supported by The Netherlands Organization for Scientific Research (NWO). The authors wish to thank Ingrid Noordhoek for many stimulating discussions. The authors thank Mary Florentine for useful comments on an earlier version of this paper.

[1]A function $f(t)$ has compact support if it is zero outside the interval $T_0 < t < T_0 + \Delta T$.

[2]In van Schijndel *et al.* (2001), the random perturbation factor with which the modulus of each wavelet coefficient was multiplied was chosen from a uniform distribution with boundaries $-L_D$ and $+L_D$.

[3]This inherent characteristic of overlap-add procedures was described in more detail by Baer and Moore (1993). Without phase distortion, even for large random shifts along the spectral axis, basic periodicity in the spectrum is preserved due to the preserved coherence of the phase spectrum.

[4]The rms duration of a real function $f(t)$ is defined by $\Delta_t := [1/\|f(t)\|_2] \sqrt{\int_{-\infty}^{\infty} (t-t_0)^2 |f(t)|^2 dt}$, with $\|f(t)\|_2 = \sqrt{\int_{-\infty}^{\infty} f^2(t) dt}$ and the center $t_0 := [1/\|f(t)\|_2^2] \int_{-\infty}^{\infty} t |f(t)|^2 dt$. The rms bandwidth is defined analogously [see, e.g., Chui (1992)].

[5]The ages of the group of normal-hearing listeners and the group of hearing-impaired listeners did not match. From the normal-hearing listeners, 10 were in their twenties, 1 was in her thirties, and 1 was in her sixties. From the hearing-impaired listeners, 2 were in their twenties, 9 were in their thirties, 3 were in their forties, 3 were in their fifties, and 9 were in their sixties. In this study, it is assumed that differences in age do not affect the results. This assumption seems reasonable. Literature shows that, for listeners under 70 years of age and with normal hearing, speech intelligibility performance does not vary with age (Studebaker *et al.*, 1997).

[6]Also, an alternative fit was used, the combination of a horizontal line and a sloping line. The subthreshold perturbations (perturbations lower than the detection threshold) were fitted with a horizontal line and all data obtained with larger perturbations with a single sloping line that intercepts the horizontal line at the perturbation threshold. The slope of the sloping line is taken as an ''alternative'' measure of sensitivity. Using these ''alternative'' sensitivities, the correlation with SII_{SRTa} is significant [Spearman rank correlation: -0.4 ($p < 0.05$)], in contrast with the correlation with the ''original'' sensitivity that was not significant. The correlation between ''alternative'' sensitivity and the SII_{SRBT} is not significant [Spearman rank correlation: -0.3 ($p = 0.07$)].

[7]Negative sensitivities to spectral perturbations may be explained by measurement uncertainty and order/list effects.

[8]The ''alternative'' measure of sensitivity (see footnote 6) leads to the same interpretation of the data. No clear trend between SII_{SRTa} and ''alternative'' sensitivity is shown [Spearman rank correlation: -0.1 ($p = 0.5$)]; there is a correlation between SII_{SRBT} and sensitivity [Spearman rank correlation: -0.5 ($p < 0.05$)].

Allen, J. B. (**1977**). ''Short term spectral analysis, synthesis, and modification by discrete Fourier transform,'' IEEE Trans. Acoust., Speech, Signal Process. **25**, 235–238.

Allen, J. B., and Rabiner, L. R. (**1977**). ''A unified approach to short-time Fourier analysis and synthesis,'' Proc. IEEE **65**, 1558–1564.

ANSI (**1997**). ANSI S3.5-1997, *American National Standard Methods for Calculation of the Speech Intelligibility Index* (American National Standards Institute, New York).

Baer, T., and Moore, B. C. J. (**1993**). ''Effects of spectral smearing on the intelligibility of sentences in noise,'' J. Acoust. Soc. Am. **94**, 1229–1240.

Bosman, A. J., and Smoorenburg, G. F. (**1995**). ''Intelligibility of Dutch CVC syllables and sentences for listeners with normal hearing and with three types of hearing deficit,'' Audiology **34**, 260–284.

Buus, S., Florentine, M., and Zwicker, T. (**1995**). ''Psychometric functions for level discrimination in cochlearly impaired and normal listeners with equivalent-threshold masking,'' J. Acoust. Soc. Am. **98**, 853–861.

Chui, K. (**1992**). *Wavelet Analysis and its Applications. Volume 1. An Introduction to Wavelets* (Academic, London), pp. 54–57.

Dreschler, W. A., and Plomp, R. (**1985**). ''Relations between psychophysical data and speech perception for hearing-impaired subjects. II,'' J. Acoust. Soc. Am. **68**, 1261–1270.

Duquesnoy, A. J., and Plomp, R. (**1980**). ''Effect of reverberation and noise on the intelligibility of sentences in cases of presbyacusis,'' J. Acoust. Soc. Am. **78**, 537–544.

Festen, J. M., and Plomp, R. (**1983**). ''Relations between auditory functions in impaired hearing,'' J. Acoust. Soc. Am. **73**, 652–662.

Florentine, M., Reed, C. M., Rabinowitz, W. M., Braida, L. D., and Durlach, N. I. (**1993**). ''Intensity perception. XIV. Intensity discrimination in listeners with sensorineural hearing loss,'' J. Acoust. Soc. Am. **94**, 2575–2586.

Gabor, D. (**1947**). ''Acoustical quanta and the theory of hearing,'' Nature (London) **159**, 591–594.
Horst, J. W. (**1987**). ''Frequency discrimination of complex signals, frequency selectivity, and speech perception in hearing-impaired subjects,'' J. Acoust. Soc. Am. **82**, 874–885.
Houtgast, T. (**1995**). ''Psycho-acoustics and speech recognition of the hearing impaired,'' in *Proceedings of the European Conference on Audiology* (Noordwijkerhout, The Netherlands), pp. 165–169.
Houtgast, T., and Steeneken, H. J. M. (**1973**). ''The modulation transfer function in room acoustics as a predictor of speech intelligibility,'' Acustica **28**, 66–73.
ISO (**1961**). International Organization for Standardization, ISO R226-1961, ''Normal Equal-loudness Contours for Pure Tones and Normal Threshold of Hearing under Free Field Listening Conditions'' (available from American National Standards Institute, New York).
Loizou, P. C., Dorman, M., and Tu, Z. (**1999**). ''On the number of channels needed to understand speech,'' J. Acoust. Soc. Am. **106**, 2097–2103.
Moore, B. C. J. (**1995**). *Perceptual Consequences of Cochlear Damage* (University Press, Oxford).
Moore, B. C. J. (**1996**). ''Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids,'' Ear Hear. **17**, 133–161.
Moore, B. C. J., and Glasberg, B. R. (**1987**). ''Relationship between psychophysical abilities and speech perception for subjects with unilateral and bilateral cochlear hearing impairments,'' in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Nijhoff, Boston), pp. 449–460.
Noordhoek, I. M., Houtgast, T., and Festen, J. M. (**1999**). ''Measuring the threshold for speech reception by adaptive variation of the signal bandwidth. I. Normal-hearing listeners,'' J. Acoust. Soc. Am. **105**, 2895–2902.
Noordhoek, I. M., Houtgast, T., and Festen, J. M. (**2000**). ''Measuring the threshold for speech-reception by adaptive variation of the signal bandwidth. II. Hearing-impaired listeners,'' J. Acoust. Soc. Am. **107**, 1685–1696.
Noordhoek, I. M., Houtgast, T., and Festen, J. M. (**2001**). ''Relations between intelligibility of narrow-band speech and auditory functions, both in the 1-kHz frequency region,'' J. Acoust. Soc. Am. **109**, 1197–1212.
Nunnally, J. C. (**1967**). *Psychometric Theory* (McGraw–Hill, New York), pp. 172–235.
Oxenham, A. J., and Moore, B. C. J. (**1995**). ''Additivity of masking in normally hearing and hearing-impaired listeners,'' J. Acoust. Soc. Am. **98**, 1921–1934.
Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (**1982**). ''The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold,'' J. Acoust. Soc. Am. **72**, 1788–1803.
Pavlovic, C. V. (**1987**). ''Derivation of primary parameters and procedures for use in speech intelligibility predictions,'' J. Acoust. Soc. Am. **82**, 413–422.
Plomp, R., and Mimpen, A. M. (**1979**). ''Improving the reliability of testing the Speech Reception Threshold for sentences,'' Audiology **18**, 43–52.
Rioul, O., and Vetterli, M. (**1991**). ''Wavelets and signal processing,'' IEEE Signal Process. Mag. **8**, 14–38.
Scharf, B. (**1970**). ''Critical bands,'' in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), Vol. 1, pp. 157–202.
Schroder, A. C., Viemeister, N. F., and Nelson, D. A. (**1994**). ''Intensity discrimination in normal-hearing and hearing-impaired listeners,'' J. Acoust. Soc. Am. **96**, 2683–2693.
Smoorenburg, G. F. (**1992**). ''Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram,'' J. Acoust. Soc. Am. **91**, 421–437.
Studebaker, G. A., Sherbecoe, R. L., McDaniel, D. M., and Gray, G. A. (**1997**). ''Age-related changes in monosyllabic word recognition performance when audibility is held constant,'' J. Am. Acad. Audiol. Ser. 1 **8**, 150–162.
ter Keurs, M., Festen, J. M., and Plomp, R. (**1992**). ''Effect of spectral envelope smearing on speech reception. I,'' J. Acoust. Soc. Am. **91**, 2872–2880.
ter Keurs, M., Festen, J. M., and Plomp, R. (**1993**). ''Limited resolution of spectral contrast and hearing loss for speech in noise,'' J. Acoust. Soc. Am. **94**, 1307–1314.
Turner, C. W., Souza, P. E., and Forget, L. N. (**1995**). ''Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners,'' J. Acoust. Soc. Am. **97**, 2568–2576.
Tyler, R. S. (**1986**). ''Frequency resolution in hearing-impaired listeners,'' in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London), pp. 309–371.
Tyler, R. S., Summerfield, Q., Wood, E. J., and Fernandes, M. A. (**1982**). ''Psychoacoustic and temporal processing in normal and hearing-impaired listeners,'' J. Acoust. Soc. Am. **72**, 740–752.
van Rooij, J. C. G. M., and Plomp, R. (**1990**). ''Auditive and cognitive factors in speech perception by elderly listeners. II. Multivariate analyses,'' J. Acoust. Soc. Am. **88**, 2611–2624.
van Schijndel, N. H., Houtgast, T., and Festen, J. M. (**1999**). ''Intensity discrimination of Gaussian-windowed tones: Indications for the shape of the auditory frequency-time window,'' J. Acoust. Soc. Am. **105**, 3425–3435.
van Schijndel, N. H., Houtgast, T., and Festen, J. M. (**2001**). ''The effect of intensity perturbations on speech intelligibility for normal-hearing and hearing-impaired listeners,'' J. Acoust. Soc. Am. **109**, 2202–2210.
Versfeld, N. J., Daalder, L., Festen, J. M., and Houtgast, T. (**2000**). ''Method for the selection of sentence materials for efficient measurement of the speech reception threshold,'' J. Acoust. Soc. Am. **107**, 1671–1684.
Yantis, P. A. (**1994**). ''Puretone air-conduction threshold testing,'' in *Handbook of Clinical Audiology*, 4th ed., edited by J. Katz (Williams and Wilkins, Baltimore), Chap. 7, p. 106.

Acoustic radiation from bowed violins

Lily M. Wang[a] and Courtney B. Burroughs
Graduate Program in Acoustics, The Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

(Received 31 August 2000; revised 8 March 2001; accepted 9 April 2001)

Near-field acoustic holography (NAH) is applied to visualize the acoustic radiation from bowed violins across a frequency range from 294 Hz to 3 kHz. These visualizations are employed to localize regions of acoustic radiation from surfaces of violins. Three violins were tested: a common student instrument by Scherl and Roth; Hutchins violin SUS295, which has been the subject of many previous investigations; and a Hutchins mezzo violin from the Violin Octet set of instruments, which is longer, broader, and thinner than a standard instrument. The violins were bowed continuously with an open-frame mechanical bowing machine, while NAH measurements were made on four planes surrounding the instrument. Mappings of the acoustic intensity are presented that show locations of maximum radiation at low and high frequencies with a spatial resolution smaller than the acoustic wavelength. Comparisons are made of the radiation patterns between the two conventional instruments and the mezzo violin. Radiation patterns from SUS295 at frequencies near to known modal responses are also presented.
[DOI: 10.1121/1.1378307]

PACS numbers: 43.75.De, 43.75.Yy [RDA]

I. INTRODUCTION

Since its emergence in the sixteenth and seventeenth centuries, the violin has become one of the most loved and subsequently most studied musical instruments. One area that welcomes more study concerns the coupling of the body motion to sound energy radiation from the instrument. Violin mode shapes have been measured and categorized,[1–3] but they are not sufficient to characterize the radiation mechanisms of the instrument. Regions of high surface velocity may not be locations of high-energy output, but rather energy sinks. Furthermore, the sound radiation may have contributions from the air cavity through the *f* holes, not measured by modal analysis.

To understand sound energy radiation requires characterization of the full-field acoustic intensity. This includes magnitude and direction of energy flow at a number of locations and distances from the source for a wide range of frequencies, which amounts to a large quantity of data. A complete study of this kind is lacking in the current literature and has only been feasible in recent years with advances in computer technology. Some previous radiation studies have concentrated on frequency responses. The results, such as Saunders' loudness curves[4] and long-term-average spectra,[5,6] provide pressure magnitude as a function of frequency, but were made only at certain distances or averaged over specific locations around the violin.

Other researchers have measured acoustic intensity from stringed instruments directly, producing vector maps of the energy-flow magnitude and direction. Tro, Pettersen, and Kristiansen[7] used a single microphone with a reference signal to obtain complex pressure measurements, which were converted into intensity, on three planes around a double bass. Only two frequencies, 98 and 230 Hz, were studied, the lower one demonstrating monopole performance, while the higher one showed interesting near-field effects. Tachibana, Yano, and Hidaka[8] presented intensity measurements, as well, on violoncellos to illustrate the sound intensity technique in acoustic near-field. Again, however, the data are available at only a few frequencies and limited to certain planes.

Weinreich and Arnold[9] suggested a more efficient technique for measuring the sound field surrounding a violin using a spherical boom system. Pressure measurements on two spheres could be used with an eigenfunction expansion solution in Hankel functions and spherical harmonics. Solving for the coefficients of the expansion essentially solves for the whole sound field, although concerns were expressed at the time with convergence at small radii. This technique is analogous to that of planar near-field acoustical holography (NAH),[10,11] which Strong and co-workers used on guitars[12] and violins.[13] NAH, which is the basis of measurements taken in this work as well, reconstructs the three-dimensional sound field, including particle velocity and acoustic intensity, from one two-dimensional set of complex pressure measurements. Strong and co-workers sinusoidally excited the instruments at low frequencies (i.e., between 78 and 425 Hz on guitar) and measured two hologram planes, one facing the top plate and the other facing the back plate. Results show the importance of the sound holes and, again, near-field behavior with energy sinks on the source surfaces.

Past studies on violin sound radiation which include both magnitude and direction of energy flow have focused on the lower-frequency regime, but comprehensive investigations of the sound energy radiation from violins at frequencies above 1 kHz have not been performed. When violins are

[a]Current address: Architectural Engineering, University of Nebraska–Lincoln, 200B Peter Kiewit Institute, Omaha, NE 68182-0681; electronic mail: lwang@unl.edu

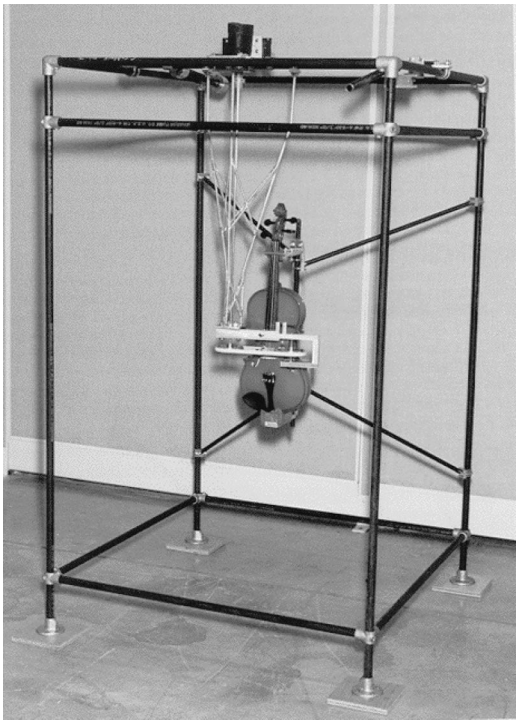


FIG. 1. Frontal view of the open-frame mechanical bowing machine. The violin is held at its neck and tailpiece in the center of the apparatus and driven by a belt sewn with horsehairs on a pulley system.

played, however, much of the sound energy is radiated at frequencies above 1 kHz. The research presented here, therefore, seeks to characterize the radiated sound field when the violin is played at frequencies extending from 294 Hz to 3 kHz, and more specifically, to localize the regions on the violin structure from which significant sound energy originates, with a resolution smaller than the acoustic wavelength. The computational algorithm for producing the surrounding radiated field is based on planar near-field acoustical holography measurements made on four planes which intersect and form a box around the violin.

A further goal of this study has been to determine how the radiation mechanisms differ between instruments of different construction. Hence, the techniques used in this research have been applied to three instruments: a Scherl and Roth full-size violin, rented from a local music store; Hutchins violin SUS295, which has been the subject of many previous experiments, including extensive modal analysis; and a mezzo violin from the Violin Octet instrument family, also made by Hutchins. With its longer and thinner shape, the mezzo violin differs quite significantly in its structure from a standard full-size violin.[14] The effects of the structural changes on the energy flow in the sound field may be established when visually comparing the results from the mezzo violin to those from the other two instruments.

II. METHOD

A. Violin excitation

To excite the violin, a steady-state bowing excitation is favored to simulate true playing circumstances as closely as possible. A custom-designed open-frame mechanical bowing machine has been constructed for this purpose (Fig. 1), and is described in detail in Ref. 15. The bowing machine applies force in a manner similar to the actual excitation of a played instrument, generating a complete set of harmonics as well as torsional motion of the string and coupling between the strings, neck, and fingerboard, which are not initiated by electromechanical excitations at the bridge.[16]

The open frame is constructed from $\frac{1}{2}$ in. black iron pipe shaft tubing. In the center of the frame, the violin rests on a foam pad at the base of the tailpiece and is held loosely at the neck with another foam pad. A driven pulley system running a continuous belt handsewn with horsehairs is placed against the violin strings when excitation is desired. Various bowing parameters may be controlled and measured, including which string is excited, bowing velocity, bow–bridge distance, and bowing force.

The nonlinear interaction between the bow and string strongly excites the motion of the string at the fundamental frequency and harmonics. The present work investigates the radiated fields measured at the frequencies of peak responses, each of which is due to a sum of violin modes. However, the contributions of individual modes to the radiated field are not separated in the analysis as in other studies,[17,18] since the interest lies not in determining the role of each mode but in assessing the combined effect on radiation. At any rate, modal overlap is high at higher frequencies, making continuous excitation of only one mode difficult. At lower frequencies, modal overlap is low, so one may expect modes which have frequencies of resonance close to the frequency of excitation to dominate the response.

B. Three test violins

Two of the three test violins were made by renowned violin maker Carleen M. Hutchins, while the other is a standard model, rented from a local music store. The rented violin is a full-size student instrument manufactured by Scherl and Roth (Model No. R270E4, Serial No. 434104) of mediocre quality, advertised as having a Stradivari design.

Hutchins SUS295 has a long experimental history, including modal analysis[2,19] and monopole radiativity tests,[20,21] both before and after modifications made to the back plate in 1988. A detailed review of experiments which have been performed on the instrument and a specific comparison of SUS295's radiation results from this study with previously documented modal analysis data are presented in Ref. 22.

The third test instrument is a Hutchins mezzo violin, SUS100, from the Violin Octet family. This family, designed by Saunders and Hutchins, consists of eight stringed instruments which are scaled so that the main air and main wood resonances match the frequencies of the middle strings, as is the case with the violin.[23] The mezzo violin was developed as a more powerful ally to the other octet instruments.[14] Compared to standard violins, it has a longer and thinner shape with larger top and back plate areas, which are expected to increase the sound radiation.

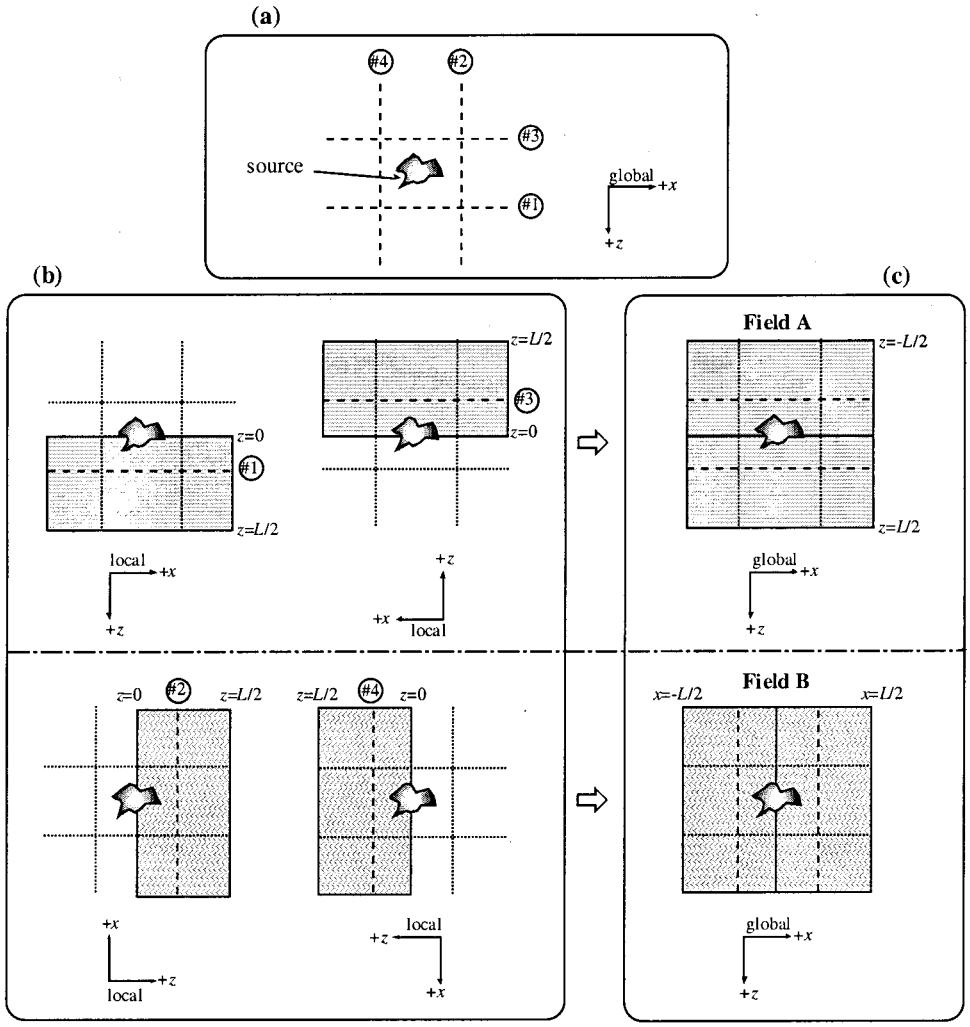


FIG. 2. Multiplanar NAH procedure, drawn in two dimensions: (a) hologram planes 1–4 box the source; (b) each hologram reconstructs a half space in local coordinates; (c) results from holograms 1 and 3 join to form field A, while data from holograms 2 and 4 join to form field B, with global coordinates shown. Finally, fields A and B are combined, either by taking the data in a certain region from the more accurate field or by averaging their values (see Ref. 22).

C. Multiplanar near-field acoustic holography (NAH) technique

To characterize the three-dimensional sound field around any source with high spatial resolution entails a large amount of data. In this study, near-field acoustical holography[10,11] is used to map the three-dimensional acoustic field around the violins from measurements made on four planar surfaces that ''box'' the violin.

Using cylindrical or spherical coordinate systems as the basis for NAH algorithms also produces three-dimensional reconstructions surrounding the source. However, planar NAH with its dependence on fast Fourier transform (FFT) algorithms is easier to implement. Also, for an arbitrary-shaped source such as a violin, the diameter of a sphere or cylindrical hologram would be dictated by the largest dimension of the violin, necessitating measurements at regions which are not in the near field of the violin above certain frequencies. Multiplanar NAH involves boxing the violin, obtaining measurements which are very close to its surface on all sides, thereby providing detailed reconstructions with high spatial resolution.

Four hologram planes of data with dimensions L by L are employed in the measurements. Marked as planes 1–4 in Fig. 2(a), these holograms box the violin and intersect each other. From each hologram, projections to parallel planes are made back towards the source center at $z=0$, and away from the source out to a distance of $z=L/2$, for each local coordinate system shown in Fig. 2(b). The discrete points on each hologram plane and the distances of projection planes are coordinated between all of the holograms, so that all together a lattice is filled with points which are equally spaced over a volume of L by L by L.

Planar NAH has traditionally been used to reconstruct field values only up to the first z plane that intersects the surface of the source. In this study, values at points that remain external to the source but lie on z planes that intersect the source have been reconstructed as well. These values are approximate; an assumption is made that the only contributions to the sound fields at these points are from parts of the source which lie 'beneath' them. This approximation is not inappropriate for the violin, whose shape closely matches the holograms. Additionally, tests of this procedure on simulated known sources showed that the approximations may be acceptable, within 3 dB on average.[22]

No values are reconstructed for points which lie within the surface of the source, since the planar NAH method used is inappropriate for determining the interior sound field. The interior field, moreover, is not the focus of this study. Therefore, as the plane of reconstruction crosses the curved surface of the top or back plate of the violin, values at points that fall

inside the surface are deleted, leaving values only at points that remain outside the surface. It is, therefore, necessary to know where these planes intersect the violin surface. A coordinate measuring machine was used to scan the surfaces of each violin and generate surface coordinates.[22] These coordinates are used not only to determine the surface location when projecting the sound field towards the source, but also to depict the violin surfaces in the visualization software.

After reconstructions from each hologram are processed, the final sound field is composed by first joining the half-space results from holograms 1 and 3, filling a field with planes that extend from $z=-L/2$ to $+L/2$ in the global coordinates shown in Fig. 2(c). Let us call this data set field A. Similarly, half-space results from holograms 2 and 4 are joined, forming a field with planes that extend from $x=-L/2$ to $+L/2$ in global coordinates; this data set is labeled as field B [Fig. 2(c)].

Fields A and B are then combined to form the final radiated sound field result, by taking values only from the more accurate field in some regions and averaging the values in other regions where reconstructions from two planes are expected to be of nearly equal accuracy.[22] A comparison of the accuracy of the multiplanar NAH combination against single planar NAH results has been performed in simulation using a transversely oscillating sphere. Overall, the multiplanar combination proves to be as, or more, accurate than the single planar NAH results in fields A and B.[22]

III. EXPERIMENT

A. Test environment and equipment

Measurements on the Scherl and Roth violin were conducted in a room with concrete walls of size 5.5 m by 7.3 m by 7.0 m. Reflections were minimized by stacking large fiberglass wedges around the apparatus. The experiments on the two Hutchins violins were performed in a semianechoic chamber with interior dimensions of approximately 5.5 m by 6.8 m by 9.3 m. In all cases, the bowing machine was placed on a concrete floor, which was covered with 2-in.-thick foam in the immediately surrounding area.

Four planes of hologram data were measured around each violin, parallel to the four sides of the apparatus frame. Since the recommended aperture size for planar holograms is at least twice the size of the source in each dimension, a hologram of 1.2 m by 1.2 m was chosen. Hologram measurements were made for the Scherl and Roth instrument over a 30 by 30 grid of points evenly spaced at 4 cm, while bowing the open A string. Thus, wave numbers for the first three partials up to 1320 Hz could be resolved. For the tests run on SUS295 and the mezzo violin, measurement positions were 1 cm apart, filling a 120 by 120 point hologram aperture; consequently, wave numbers for frequencies up to 3 kHz have been resolved. The open D and A strings were tested for SUS295, while only the open A string was measured for the mezzo violin, so this frequency range covers at least the first six partials for each string studied. Near-field holograms for the three violins were measured at distances of 1–4 cm away from the nearest exterior point on the surface of the violin; thus, for the highest measurement frequency of 3 kHz, most measurements were made at a distance of less than a quarter of a wavelength.[22]

TABLE I. Bowing parameters maintained during hologram measurements for the listed violins and string excitations.

Violin and string excitation	Bow speed (m/s)	Bow–bridge distance (m)	Bow force (N)
Scherl and Roth, open A	0.35	0.03	0.6
SUS295, open A	0.35	0.03	1.0
SUS295, open D	0.35	0.03	0.8
Mezzo violin SUS100, open A	0.35	0.03	1.0

A linear array of 15 electret microphones was used to obtain the complex sound pressure measurements. The diameter of each transducer is 1 cm; accordingly, the microphone size was much smaller than the wavelength for the highest frequency of 3 kHz, where $\lambda=11.4$ cm. The compact microphones were aligned vertically and attached to a one-dimensional positioning slide, situated horizontally on custom-built legs. While the array was moved to different positions in the horizontal direction by a computer-controlled stepper motor, its vertical placement was adjusted manually to eight positions, providing measurement locations that were vertically 1 cm apart. The positioning slide was kept at a fixed location in the testing rooms, while the bowing machine was rotated for each hologram measurement. A condenser microphone was clamped to the lower of the top bars on the bowing machine frame, facing the top plate of the violin. The signal from this microphone provided a phase reference for data collected by the traversing array.

The 16 channels of data were digitally acquired through a simultaneously sampling 16 bit analog-to-digital (A/D) board, mounted in a PC computer which ran the Hewlett Packard Visual Engineering Environment (HPVEE) software program. A low-pass filter with cutoff at 5 kHz was applied. Before measuring any hologram data, the electret microphones and their signal paths were calibrated relative to the reference microphone. Two standing-wave tubes were used to cover the frequency range from 100 Hz to 5 kHz.[22]

B. Hologram acquisition and field reconstruction

Prior to the start of each hologram measurement, the violin was tuned, and the various bowing parameters were adjusted to desired values within typical ranges.[24,25] Table I lists the parameters used for each violin test. Since the parameters were not the same between violins, comparisons between absolute levels are not made. Relative magnitudes are contrasted, however, since the violin behaves as a linear system.

A program written in HPVEE coordinated the acquisition of the hologram data. At each position, time signals from the 16 channels were simultaneously sampled for 1 s at 10 kHz for the Scherl and Roth tests, and at 15 kHz for the tests on the other two instruments. No additional averages were taken, since the data were deterministic and demonstrated sufficiently high spectral levels at the fundamental and harmonic frequencies. The signals from the reference microphone and each electret microphone were constantly monitored with a frequency analyzer to ensure the stability of the signal level.

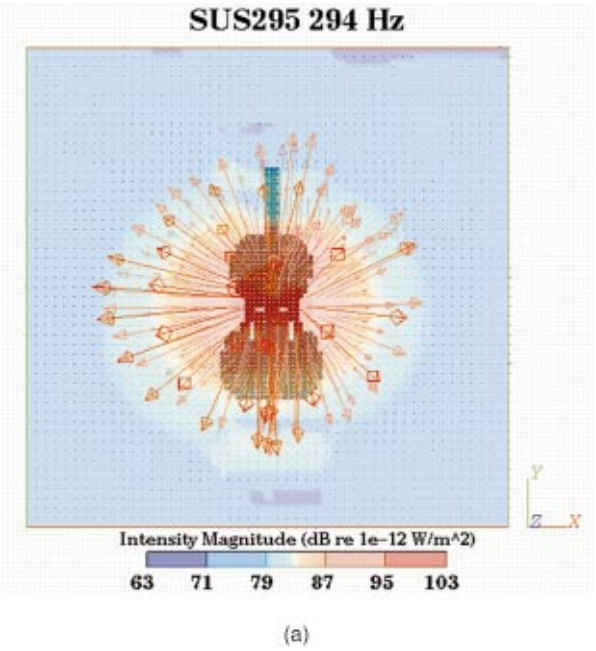


FIG. 3. NAH reconstructed intensity field for violin SUS295 at 294 Hz: (a) facing the top plate and (b) facing the back plate. The plotted data are from the closest local z plane that does not intersect the violin surface. The three-dimensional intensity vectors are shown, while the intensity vector magnitudes are denoted by color.

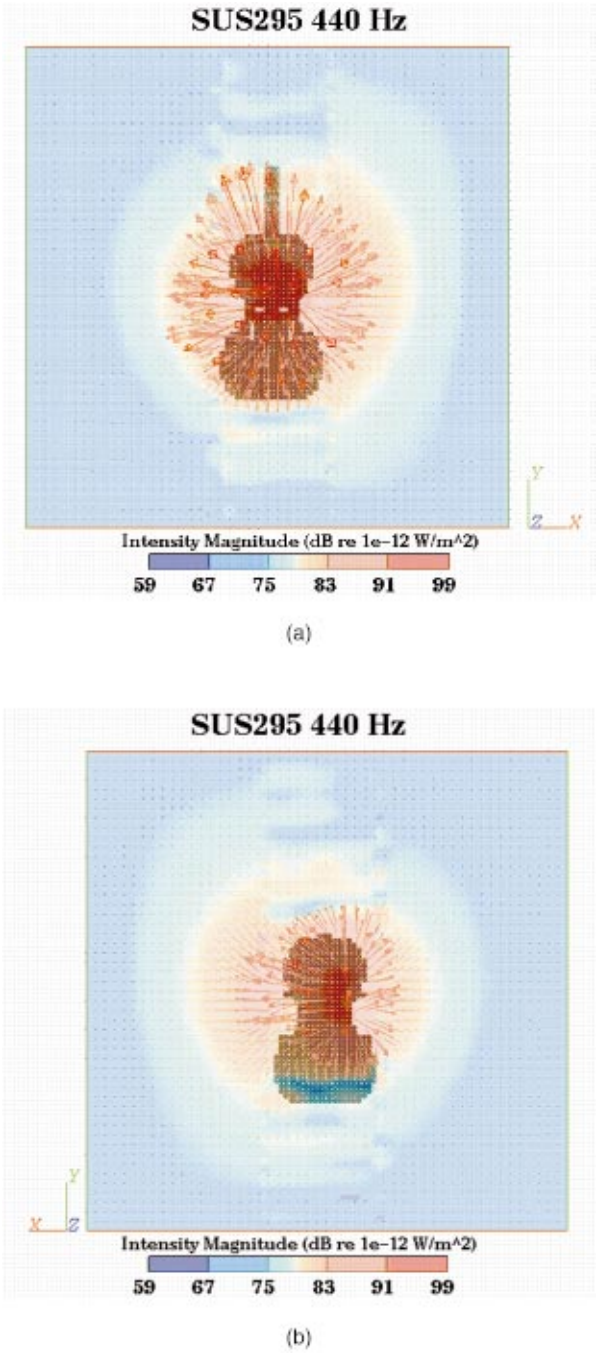
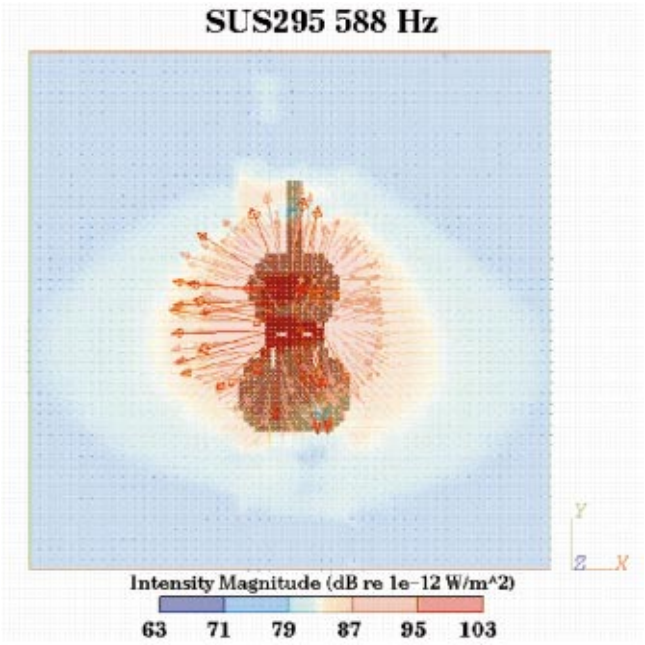


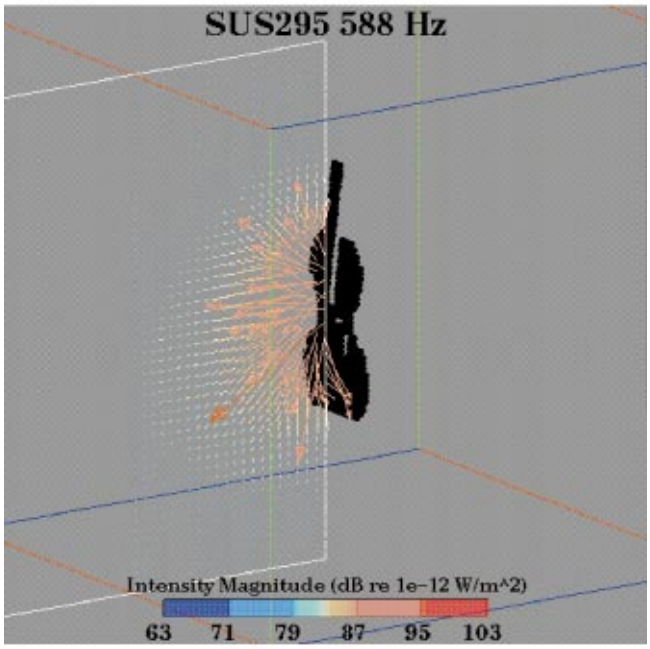
FIG. 4. NAH reconstructed intensity field for violin SUS295 at 440 Hz: (a) facing the top plate and (b) facing the back plate. The plotted data are from the closest local z plane that does not intersect the violin surface. The three-dimensional intensity vectors are shown, while the intensity vector magnitudes are denoted by color.

HPVEE processed the raw time signals at every position, first by applying a Hanning window, then taking an FFT to produce results in the frequency domain. Calibrations for amplitude and phase were applied. The resulting data demonstrated signal-to-noise ratios of at least 30 dB at the frequencies of each partial for each bowed string up to 3 kHz.

A program in *C* was used to consolidate selectively the data from all of the spectra files into the spatial holograms for each of the partial frequencies between 100 and 3000 Hz. A discrete implementation of planar NAH coded with MATLAB and *C* was subsequently applied for reconstruction in local coordinates, producing results for the pressure magnitude and the active intensity in three dimensions.[22] The N by M point hologram is spatially windowed with a broad Tukey window[26] and zero-padded to at least twice the aperture size in MATLAB; then, the *C* code Fourier transformed the spatial data to the wave-number domain, multiplied by the appropriate Green's function ''propagator'' for the desired reconstruction distance, applied an additional k-space window for inverse reconstruction, and finally, inverse transformed back to the real domain. The reconstructions were

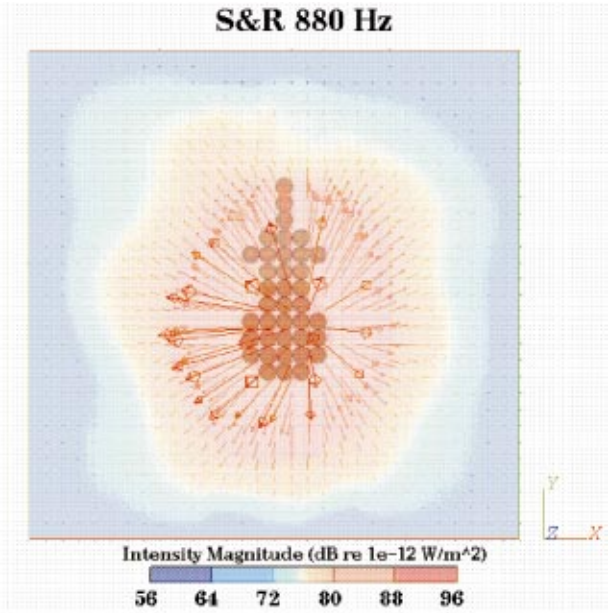


(a)

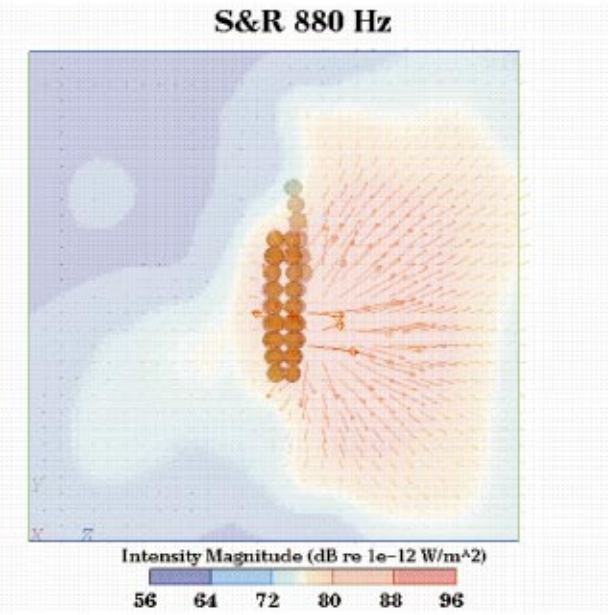


(b)

FIG. 5. NAH reconstructed intensity field for violin SUS295 at 588 Hz: (a) facing the top plate and (b) vectors on a slice through the field showing near-field effects. The three-dimensional intensity vectors are shown, while the intensity vector magnitudes are denoted by color.



(a)

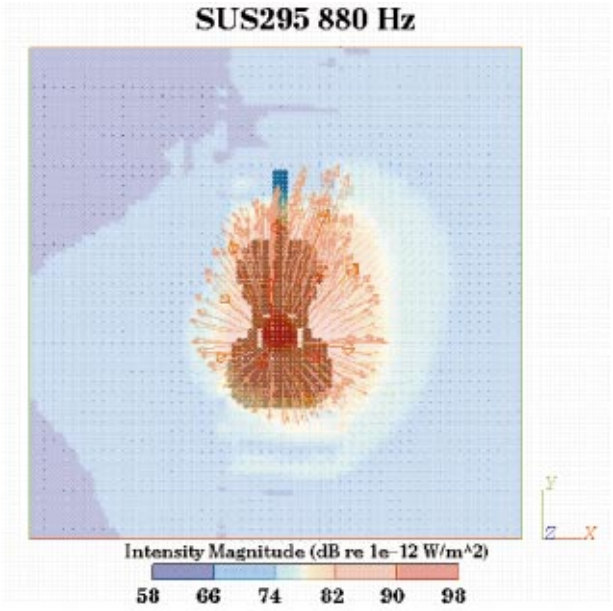


(b)

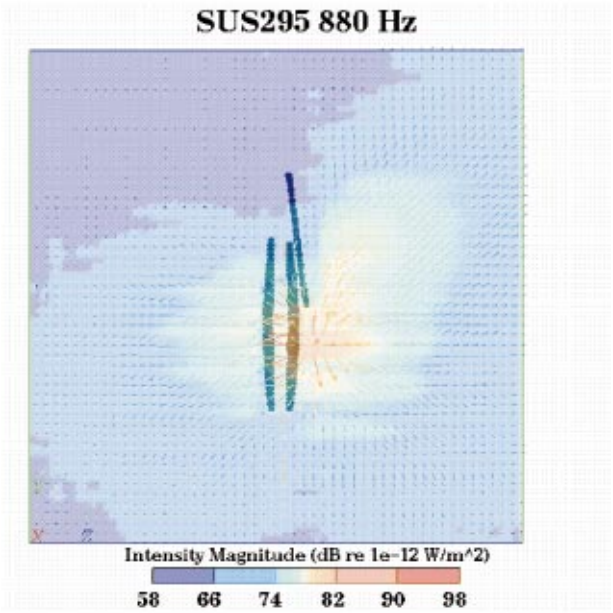
FIG. 6. NAH reconstructed intensity field for the Scherl and Roth violin at 880 Hz: (a) facing the top plate and (b) facing the bass bar side. The plotted data are from the closest local z plane that does not intersect the violin surface. The three-dimensional intensity vectors are shown, while the intensity vector magnitudes are denoted by color.

accomplished from the hologram onto parallel planes separated by a distance equal to the spacing between measurement points, filling a lattice of discrete points evenly spaced at 4 cm for the Scherl and Roth violin, and 1 cm for the other two violins. Planes of reconstruction from each hologram extended from the local $z=0$ center plane to $z=L/2$, where L is the size in one dimension of the reduced aperture (1 m).

After the half-space reconstructions were accumulated, they were combined according to the multiplanar NAH procedure reviewed above. Due to variations in bowing parameters and the situation of the bowing machine in the testing rooms, the absolute values from the four half-space reconstructions were not consistent in overlapping regions. It was clear from visual study, however, that the data in these areas did have the same relative values within each field. Thus, the half-space reconstructions were normalized to each other, prior to the combination procedure, resulting in an average magnitude difference of 1 dB in overlapping corners across

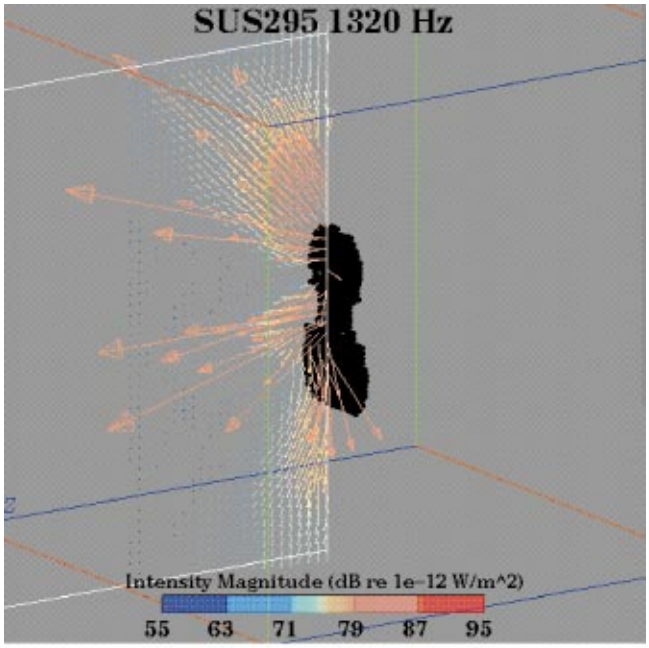


(a)

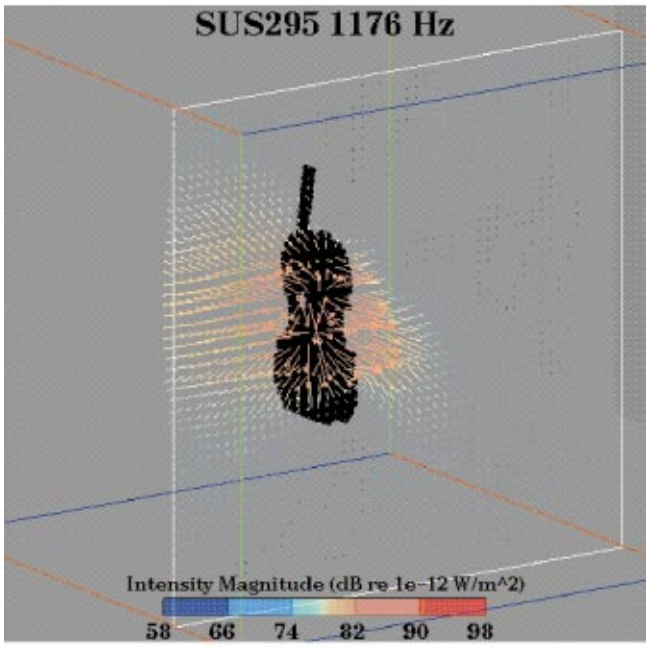


(b)

FIG. 7. NAH reconstructed intensity field for violin SUS295 at 880 Hz from excitation of the open A string: (a) facing the top plate and (b) facing the bass bar side. The plotted data are from the closest local z plane that does not intersect the violin surface. The three-dimensional intensity vectors are shown, while the intensity vector magnitudes are denoted by color.



(a)



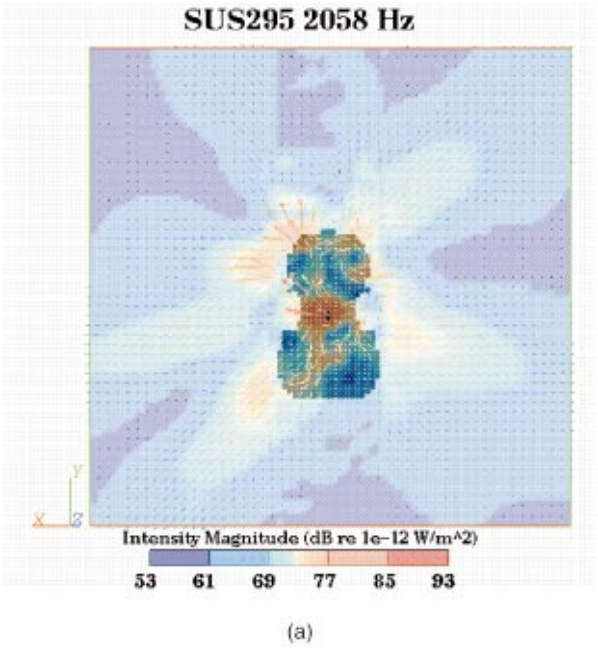
(b)

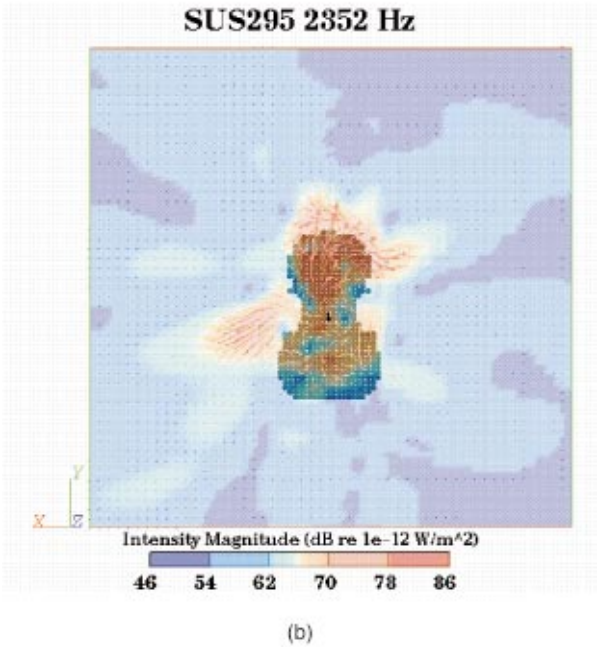
FIG. 8. Slices through the NAH reconstructed intensity vector fields for violin SUS295 at (a) 1320 Hz, showing directions of energy radiation not normal to the violin, and (b) 1176 Hz, showing vectors which curve around the edges of the violin. The three-dimensional intensity vectors are shown, while the intensity vector magnitudes are denoted by color.

the frequencies studied.[22] The pressure and active intensity results were finally visualized using the software Application Visualization System (AVS).

IV. RESULTS

Examination of the radiated sound field in terms of the active intensity vectors provides knowledge of the location of energy sources on the violin, as well as the direction of energy radiation from the instrument. For both the Scherl and Roth and SUS295 violins, what is most apparent across all frequencies studied is the dominance of the top plate as a radiator of sound energy (Figs. 3–10). The vectors which are greatest in magnitude always originate from the top plate, producing lobes radiating strongly outwards. Even at the lowest frequencies which demonstrate omnidirectional behavior,[15] such as 294, 440, and 588 Hz for SUS295, most of the intensity vectors originate from the top plate.





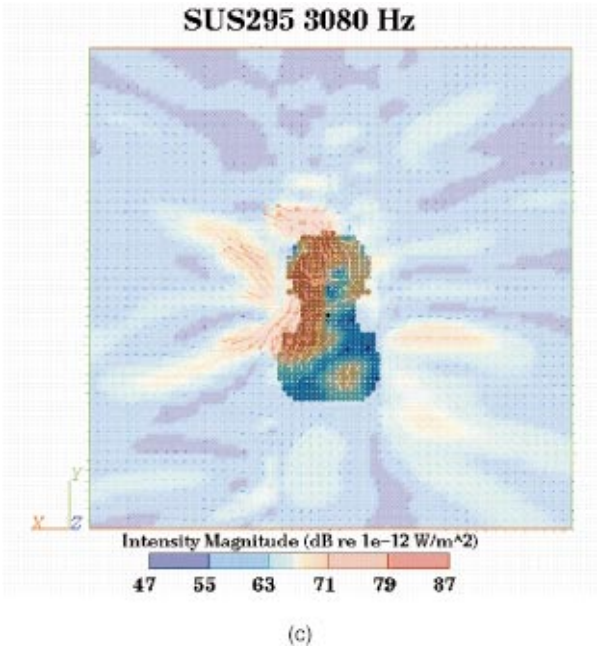
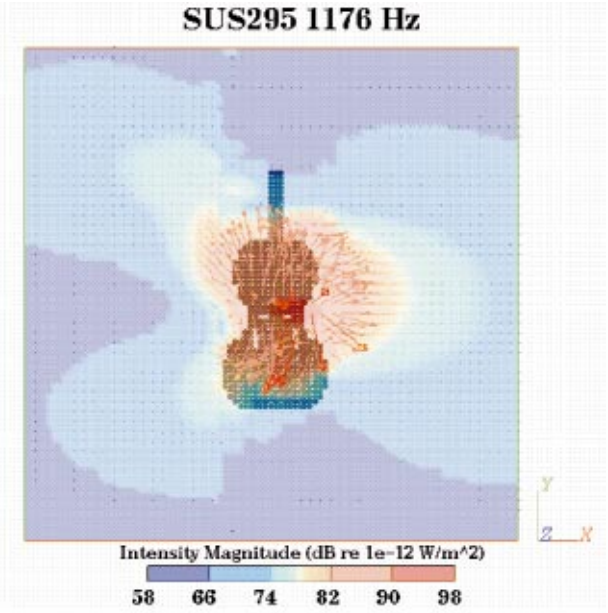
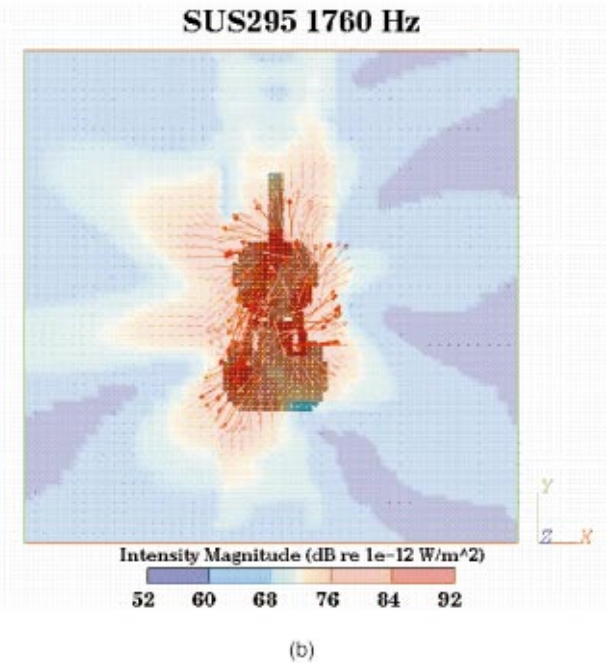


FIG. 9. Bottom edge of the back plate is rarely found to be a region of radiating energy, as evidenced by the NAH reconstructed intensity fields for violin SUS295, facing the back plate at (a) 2058 Hz, (b) 2352 Hz, and (c) 3080 Hz. The plotted data are from the closest local z plane that does not intersect the violin surface. The three-dimensional intensity vectors are shown, while the intensity vector magnitudes are denoted by color.

At 294 Hz, the *f*-hole region dominates [Fig. 3(a)], while the back plate does not display any significant sources of sound radiation [Fig. 3(b)]. This behavior corresponds to modal analysis because the first main resonance or the ''main air'' mode, which was found to be at 276 Hz for SUS295,[19] exhibits air pumping in and out of the *f* holes. The frequency of resonance for this mode is expected to shift slightly, depending on test conditions,[27] but certainly it is located close to the 294 Hz excitation. Thus, the air mode is a major contributor to the radiation at this frequency, with sound energy generated by air motion in and out of the *f*-hole region. Directivity patterns,[15] however, show omnidirectional radiation at 294 Hz. This is explained by the fact that the sound waves have long wavelengths at this frequency compared to the size of the violin, and therefore, energy is diffracted around the instrument to produce the far-field omnidirectional behavior.

At 440 Hz for SUS295, the region from which the strongest intensity vectors emanate is slightly above the *f* holes, from the wood in the upper half of the top plate, or upper bout [Fig. 4(a)]. At this frequency, the back plate shows more activity than was apparent at 294 Hz, but still less than the top plate [Fig. 4(b)]. Marshall[19] found the ''main wood'' mode, or the first bending mode with motion predominantly in the top plate, at 472 Hz. Now, instead of radiation from an air mode in which the body plays a small role, the motion of the body plays a dominant role in the radiation at 440 Hz. Thus, the source is located higher on the top plate in the wood, rather than at the *f*-hole region, and the back plate plays a greater role at 440 Hz than it did at 294 Hz. The





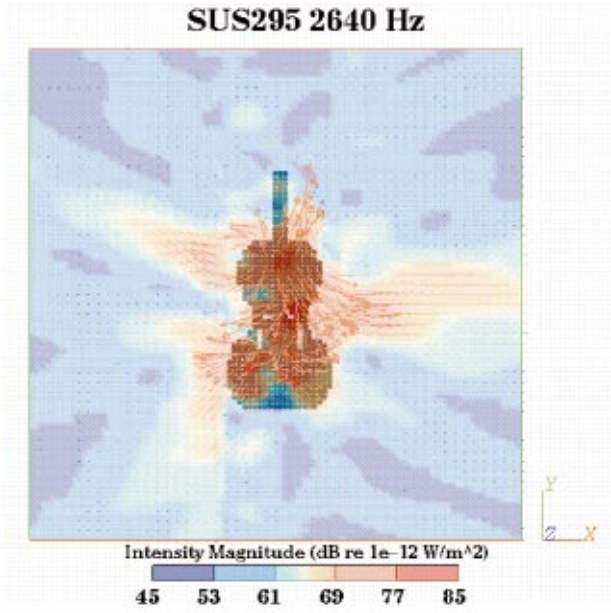


FIG. 10. Area around the soundpost is consistently a region of sound radiation activity, as supported by these NAH reconstructed intensity fields for violin SUS295, facing the top plate at (a) 1176 Hz, (b) 1760 Hz, and (c) 2640 Hz. The plotted data are from the closest local z plane that does not intersect the violin surface. The three-dimensional intensity vectors are shown, while the intensity vector magnitudes are denoted by color.

radiation mechanisms described so far for 294 and 440 Hz both correspond with modal behavior. Further comparisons of the radiation mechanisms under 1000 Hz with SUS295 mode shapes, determined from modal analysis by other researchers, are presented in Ref. 22.

The role of the violin's plates in the radiation of sound energy continues at the next lowest frequency studied of 588 Hz, which is located near a bending mode. Regions of energy radiation now appear more localized on the violin [Fig. 5(a)]. Also, some near-field effects, where the sound energy travels back towards the surface of the violin, are apparent at this frequency [Fig. 5(b)]. Overall, the radiated field at this frequency is still omnidirectional.

For the Scherl and Roth violin, the intensity vector plots at 440 Hz show similar omnidirectional results stemming primarily from the top plate. The locations of sound energy sources are not as precise, though, since the resolution is not as high as the results on SUS295. The 880 Hz intensity result for the Scherl and Roth violin similarly does not provide great detail, but it does demonstrate the predominance of the top plate in radiating sound energy and follows a cardioid radiation pattern [Figs. 6(a) and (b)], as the pressure magnitude does.[22] The cardioid pattern seems to indicate that at this frequency, the baffling effect of the violin body becomes significant, i.e., less diffraction of the radiation from the top plate occurs around the body of the violin while the sound energy propagates directly outwards from the top plate. The same cardioid shape is found in greater spatial resolution

with the SUS295 results at 880 Hz [Figs. 7(a) and (b)]. Apparently, one can summarize that at frequencies under 600 Hz, where omnidirectional radiation patterns result, the top plate is primarily responsible for radiating energy, and the sound energy diffracts around the violin, producing omnidirectional far-field results. Around the frequency 880 Hz, though, the radiation becomes baffled by the violin body, producing cardioid radiation patterns. If one assumes that the baffling begins to occur physically when ka is approximately equal to 1, where k is the acoustic wave number and a is the effective radius of the source, then for a violin a is approximately 6 cm. This corroborates theoretical estimates of $a=7$ cm by Weinreich.[28]

At frequencies above 1 kHz, the asymmetry of the violin becomes more apparent when looking at the radiation vectors emanating from the instrument (Figs. 8–10). Energy sources are more localized and distinct on the structure. Smaller source regions are expected, of course, with the shorter wavelength at higher frequencies. The asymmetry is also anticipated, since the violin is an asymmetrical instrument, due to the internal placement of the soundpost and bass bar, as well as externally with the tuning of the strings and inhomogeneities in the wood.[2] Along with the more complicated radiation mechanisms appearing above 1 kHz, the paths of the intensity vectors no longer extend directly normal to the surface of the violins, as at lower frequencies. Instead, the intensity vectors are often directed up or down into the space, mostly from the top and back plates [Fig. 8(a)]. Sometimes vectors are also found curving around the C bout and the edges of the violin, acting like line sources [Fig. 8(b)]. At frequencies below the critical frequency but above the frequencies of the lowest modes, such radiation is expected from edges of finite, unbaffled structures.

The regions of dominant sound energy radiation vary across the top and back plates for different frequencies. Two generalizations may be made, however, concerning the locations of ''hot spots.'' One is that the bottom edge of the back plate is rarely found to be a region of radiating energy [Figs. 9(a)–(c)]. Upon first glance, one may consider this to be a result of the violin mounting; however, the violin was not clamped at that location but simply held in foam at the base, and the same result is not apparent at the bottom edge of the top plate. The second trend that is noted from these studies is that the area surrounding the soundpost on the top plate consistently seems to be a region of some energy radiation [Figs. 10(a)–(c)]. The soundpost, lodged between the top and back plates, often forces vibration patterns to have a nodal point in that region. The nodal point emphasizes the violin asymmetry and leads to less cancellation of the sound field from the top plate, thereby increasing sound radiation.[29] The region around such a constraint is expected to be a source of radiating energy, as confirmed in this study.

Radiation mechanisms from the violin of different construction, Hutchins' mezzo violin SUS100, have also been characterized and compared with those discussed above for violins of standard construction, primarily SUS295. At the lowest frequency of comparison, 440 Hz, the radiated fields from both instruments are omnidirectional,[22] and the radiation mechanisms between the two instruments are also similar. The energy radiates primarily from the center region of the top plate in both instruments, while the back plate has more emphasis on the bass bar side, projecting towards the soundpost side. At the next partial, 880 Hz, it was previously noted that the two standard violins exhibit cardioid radiation patterns, with most radiation coming off of the center of the top plate leaning towards the soundpost side and more defined regions of energy radiation off the back plate which are less in magnitude [Figs. 7(a) and (b)]. The mezzo violin's behavior at this frequency is distinctly different. Although the center of the mezzo violin's top plate does radiate some

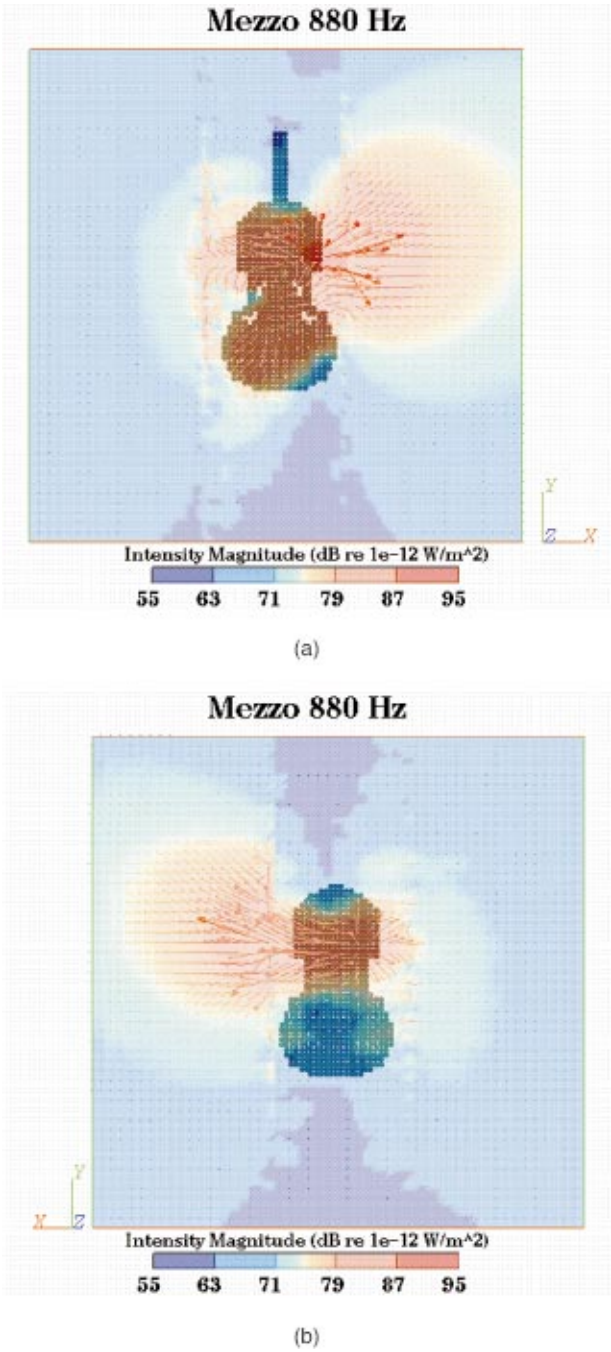


FIG. 11. NAH reconstructed intensity field for the mezzo violin SUS100 at 880 Hz: (a) facing the top plate and (b) facing the back plate. The plotted data are from the closest local z plane that does not intersect the violin surface. The three-dimensional intensity vectors are shown, while the intensity vector magnitudes are denoted by color.

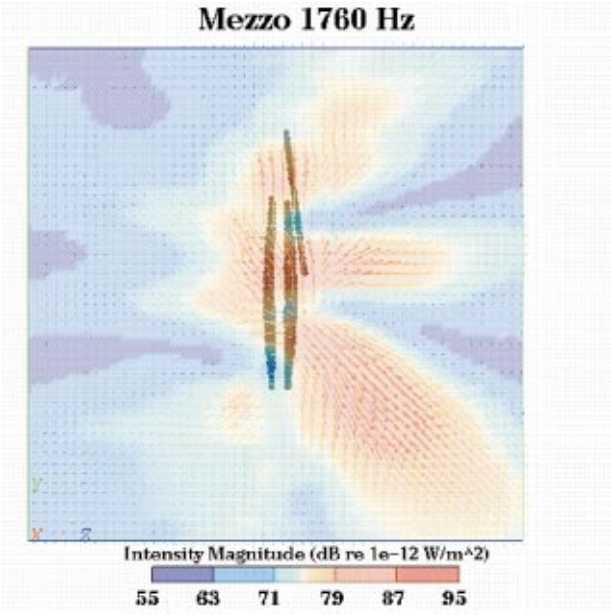


FIG. 12. NAH reconstructed intensity field for the mezzo violin SUS100 at 1760 Hz: (a) facing the bass bar side and (b) facing the top plate. The plotted data are from the closest local z plane that does not intersect the violin surface. The three-dimensional intensity vectors are shown, while the intensity vector magnitudes are denoted by color.

sound energy, the cardioid with a maximum directly normal to the top plate is not found. Instead, other regions along the soundpost side dominate the response, with the most significantly radiating area located in the upper soundpost quadrant near the upper C bout [Fig. 11(a)]. Meanwhile, the upper bass bar quadrant of the back plate shows radiating activity, as it did on SUS295, but the lower bout does not [Fig. 11(b)]. These differences may be due to differences in the modal behavior between SUS295 and the mezzo violin.[30] Unfortunately, no modal analysis data are available for the mezzo violin to assist in the analysis of the differences shown in the acoustic fields between the two violins at 880 Hz.

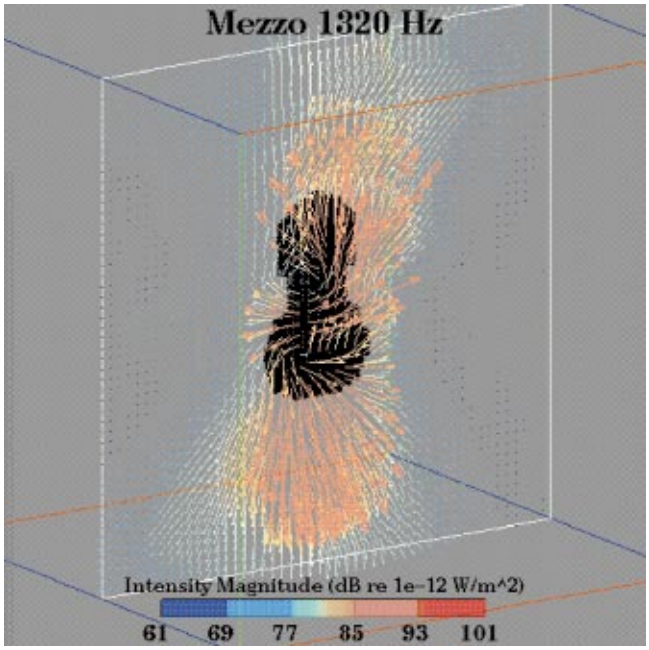


FIG. 13. Slice through the NAH reconstructed intensity vector field for the mezzo violin at 1320 Hz, showing near-field effects. The three-dimensional intensity vectors are shown, while the intensity vector magnitudes are denoted by color.

At higher frequencies, a few generalities about the mezzo violin's radiation mechanisms remain the same as for SUS295. The top plate is the dominant radiator of sound energy. Also, asymmetry is clearly apparent, with energy traveling in a variety of directions, not necessarily normal to the top plate [Fig. 12(a)]. There is evidence of near-field effects, where the sound energy from a local source turns back towards the surface, falling into a local energy sink (Fig. 13). Actual locations and the relative magnitudes of energy radiating regions differ quite significantly between the mezzo violin and the standard SUS295. However, some similarities are found between regions which do not have a large response. For example, the radiation results at 1760 Hz from the top plate of SUS295 [Fig. 10(c)] and the mezzo violin [Fig. 12(b)] both show low-intensity magnitudes in the lower soundpost quadrant.

One significant difference is apparent at the three highest frequencies studied: 2200, 2640, and 3080 Hz. Relative to the top plate's radiation, the back plate on the mezzo violin radiates much less than on the standard violin SUS295, as noted when contrasting Figs. 9(a)–(c) with Figs. 14(a) and (b). This seems rather contradictory to the instrument's purpose, since one of the main reasons for the development of the mezzo violin was to produce more sound power by increasing the area of the upper and lower quadrants. The structural changes, though, have lessened the impact of the back plate at these higher frequencies.

V. SUMMARY AND CONCLUSIONS

Regions responsible for sound energy radiation from bowed violins have been characterized for frequencies from

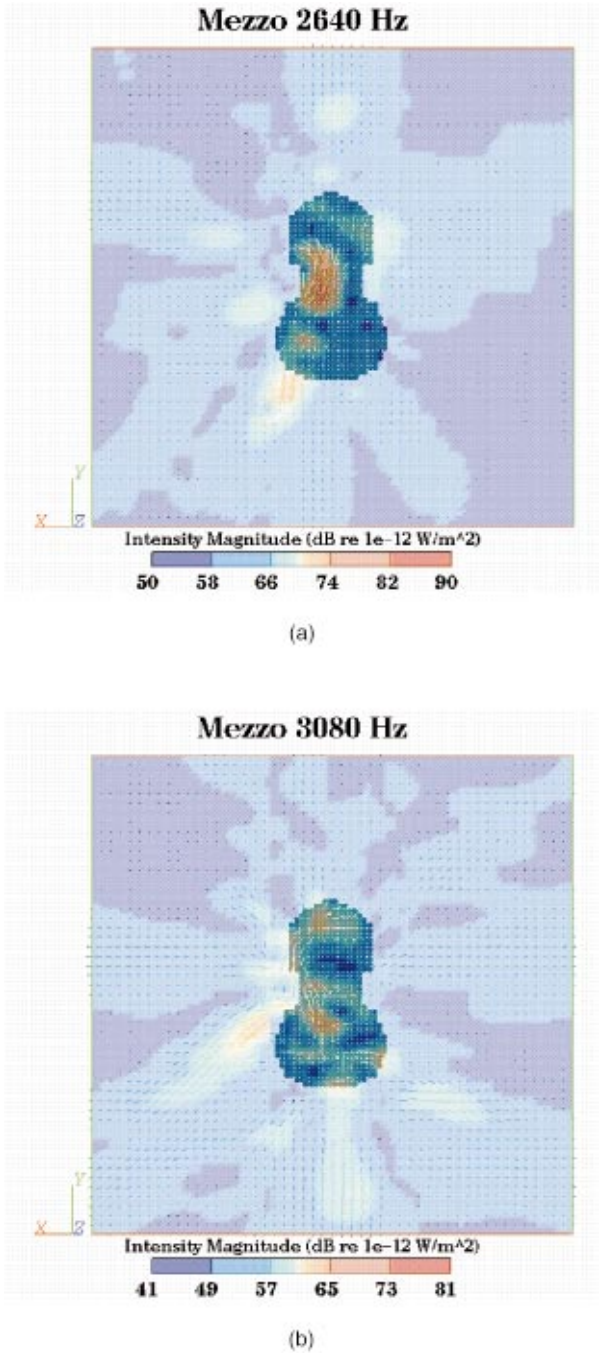


FIG. 14. NAH reconstructed intensity field for the mezzo violin SUS100 facing the back plate at (a) 2640 Hz and (b) 3080 Hz. The plotted data are from the closest local *z* plane that does not intersect the violin surface. The three-dimensional intensity vectors are shown, while the intensity vector magnitudes are denoted by color.

294 Hz to 3 kHz, by applying multiplanar NAH. Certain trends in the sound energy radiation across frequencies were noted. First, the top plate was found to be the dominant source of sound energy radiation from the violins. At low frequencies, diffraction of sound radiation from the top plate around the violin body produces far-field omnidirectional radiation. Around 880 Hz, baffling by the violin body begins to produce far-field cardioid patterns since the energy is radiating primarily from the top plate. At higher frequencies, the regions of energy sources become more localized and are distributed asymmetrically on the violin body. Evidence shows that sound radiation rarely originates from the bottom edge of the back plate, but often radiation is found stemming from the area around the soundpost of the top plate.

The significant conclusions of these results are varied. Because of the clear dominance of the violin's top plate in sound energy production, the tuning of the top plate may be more crucial than that of the back plate during violin construction, in terms of the instrument's strength. The fact that the violin top plate faces out towards the audience in performance indicates an intuitive understanding over time of where most of the sound energy comes from, and this study has confirmed it. Additionally, the long-standing belief among violin makers and players that the placement of the soundpost critically affects an instrument's sound is justified, since the soundpost region has been found to be a predominant source region at most of the frequencies studied.

Comparison of results between the three violins has indicated similarities between the two violins of normal construction, but significant differences with the mezzo violin. The mezzo violin results do demonstrate similar general trends across frequencies, namely, that the top plate is the main radiator. However, the mezzo violin's back plate plays a more minor role relative to the top plate at frequencies between 2 and 3 kHz than is true for standard violins. It appears that the coupling to the back plate is reduced in the mezzo violin, perhaps due to the combination of thinner ribs and greater back plate area across which to transmit the vibrational energy. The lesser amount of sound radiation from the back plate may not affect the overall sound level from the violin which reaches the audience; indeed, there is evidence that the mezzo violin is more powerful than ordinary violins at certain low frequencies.[22] However, the lack of back plate radiation may affect how the instrument is subjectively experienced by the player.

ACKNOWLEDGMENTS

The authors are grateful to Carleen Hutchins for the loan of her violins, SUS295 and SUS100. This work has been supported by a National Science Foundation (NSF) Graduate Research Fellowship, American Association for Women (AAUW) Selected Dissertations Fellowship, and a Bell Laboratories Graduate Research Program for Women (GRPW) grant.

[1] J. A. Moral and E. V. Jansson, ''Eigenmodes, input admittance, and the function of the violin,'' Acustica **50**, 329–337 (1982).
[2] K. D. Marshall, '''Modal analysis of a violin,'' J. Acoust. Soc. Am. **77**, 695–709 (1985).
[3] E. V. Jansson, N. E. Molin, and H. O. Saldner, ''On eigenmodes of the violin—Electronic holography and admittance measurements,'' J. Acoust. Soc. Am. **95**, 1100–1105 (1994).
[4] F. A. Saunders, ''The mechanical action of violins,'' J. Acoust. Soc. Am. **9**, 91–98 (1937).
[5] E. V. Jansson, ''Long-time-average spectra applied to analysis of music. Part III: A simple method for surveyable analysis of complex sound sources by means of a reverberation chamber,'' Acustica **34**, 275–280 (1976).
[6] A. Gabrielsson and E. V. Jansson, ''Long-time-average spectra and rated qualities of twenty-two violins,'' Acustica **42**, 47–55 (1979).
[7] J. Tro, O. Kr. Ø. Pettersen, and U. K. Kristiansen, ''Sound radiation from a double bass visualized by intensity vectors,'' J. Catgut Acoust. Soc. **40**, 7–9 (1983).

[8] H. Tachibana, H. Yano, and Y. Hidaka, ''Visualization of sound fields by the sound intensity technique,'' Proceedings of the 2nd Symposium on Acoustic Intensity, Tokyo, Japan, pp. 117–126 (1988) (in Japanese).
[9] G. Weinreich and E. B. Arnold, ''Method for measuring acoustic radiation fields,'' J. Acoust. Soc. Am. **68**, 404–411 (1980).
[10] E. G. Williams, J. D. Maynard, and E. Skudrzyk, ''Sound source reconstructions using a microphone array,'' J. Acoust. Soc. Am. **68**, 340–344 (1980).
[11] J. D. Maynard, E. G. Williams, and Y. Lee, ''Near-field acoustic holography: I. Theory of generalized holography and the development of NAH,'' J. Acoust. Soc. Am. **78**, 1395–1413 (1985).
[12] W. Y. Strong, Jr., T. B. Beyer, D. J. Bowen, E. G. Williams, and J. D. Maynard, ''Studying a guitar's radiation properties with near-field holography,'' J. Guitar Acoustics **6**, 50–59 (1982).
[13] W. Y. Strong, Jr., E. Torick, and J. D. Maynard, ''Experimental study of vibration and radiation characteristics of a violin,'' J. Acoust. Soc. Am. **72**, S83 (1982).
[14] C. M. Hutchins and J. C. Schelling, ''A new concert violin,'' J. Audio Eng. Soc. **15**, 432–436 (1967).
[15] L. M. Wang and C. B. Burroughs, ''Directivity patterns of acoustic radiation from bowed violins,'' J. Catgut Acoust. Soc. **3**, 7–15 (1999).
[16] H. Dünnwald, ''Deduction of objective quality parameters on old and new violins,'' J. Catgut Acoust. Soc. **1**, 1–5 (1991).
[17] G. Bissinger, ''Some mechanical and acoustical consequences of the violin soundpost,'' J. Acoust. Soc. Am. **97**, 3154–3164 (1995).
[18] G. Bissinger, ''Modeling the sound of the violin: The V–R model and the role of the soundpost,'' J. Catgut Acoust. Soc. **3**, 29–37 (1998).
[19] K. D. Marshall, personal research notes (1990).
[20] O. E. Rodgers, ''Effect on plate frequencies of local wood removal from violin plates supported at the edges,'' J. Catgut Acoust. Soc. **1**, 7–11 (1991).
[21] O. E. Rodgers, private communications (1997).
[22] L. M. Wang, ''Radiation mechanisms from bowed violins,'' Ph.D., Pennsylvania State University (1999).
[23] C. M. Hutchins, ''A 30-year experiment in the acoustical and musical development of violin-family instruments,'' J. Acoust. Soc. Am. **92**, 639–650 (1992).
[24] A. Askenfelt, ''Measurement of bow motion and bow force in violin playing,'' J. Acoust. Soc. Am. **80**, 1007–1015 (1986).
[25] A. Askenfelt, ''Measurement of the bowing parameters in violin playing. II: Bow–bridge distance, dynamic range, and limits of bow force,'' J. Acoust. Soc. Am. **86**, 503–516 (1989).
[26] E. G. Williams, *Fourier Acoustics: Sound Radiation and Near-Field Acoustical Holography* (Academic Press, San Diego, CA, 1999), p. 105.
[27] M. Roberts and T. D. Rossing, ''Normal modes of vibration in violins,'' J. Catgut Acoust. Soc. **3**, 3–9 (1998).
[28] G. Weinreich, ''Directional tone color,'' J. Acoust. Soc. Am. **101**, 2338–2346 (1997).
[29] C. M. Hutchins, ''A history of violin research,'' J. Acoust. Soc. Am. **73**, 1421–1440 (1983).
[30] Admittance plots for SUS295 and the mezzo violin SUS100 are distinctly different in the low-frequency regime, as shown in Ref. 22, implying different modal behavior.

Inclusion of wave steepening in a frequency-domain model of trombone sound production

Michael W. Thompson and William J. Strong
Department of Physics and Astronomy, Brigham Young University, Provo, Utah 84602

(Received 31 July 2000; revised 30 January 2001; accepted 21 March 2001)

A frequency-domain model of trombone sound production that includes the effects of wave steepening is proposed. This model builds upon the work of Msallam *et al.* [Proceedings of the Institute of Acoustics **19**(5), 419–424 (1997)] by including thermoviscous wall losses in a more realistic manner, by applying wave steepening systematically to the entire instrument, and by making quantitative comparisons to the experimentally measured output of the same trombone that the model is based on. The trombone is approximated as a set of contiguous cylindrical tubes with superposition of incoming and outgoing waves in each cylinder and with continuity of pressure and flow at each cylinder junction. The far-field radiated pressure spectrum is calculated on the basis of the spectrum of a measured pressure wave in the mouthpiece. This calculation includes the effects of wave steepening for the outgoing wave in each cylinder. The equations describing the model are given. Mouthpiece spectra are processed both with and without the effects of wave steepening. The predicted far-field spectra are compared to the corresponding measured far-field spectra. In all cases analyzed, the inclusion of wave steepening greatly reduces the error between predicted and measured spectra. [DOI: 10.1121/1.1371759]

PACS numbers: 43.75.Fg [RDA]

I. INTRODUCTION

For several years, wave steepening has been known to be partially responsible for the typically bright timbre of loudly blown trombone tones. Beauchamp (1980) observed that differences in sound pressure levels between radiated harmonics and corresponding harmonics in the mouthpiece are not constant for all dynamics, as a linear model would predict. An example of such a nonlinear response is shown in Fig. 1, where the divergence of the two curves is the result of amplitude-dependent behavior.

Hirschberg *et al.* (1996), documented shock waves at the bell of the trombone using a flow visualization technique. They also suggested that the change in timbre due to this observed wave steepening could be modeled by calculating the outgoing pressure wave in the mouthpiece based on linear theory, steepening the outgoing wave appropriately, and filtering the wave with a linear model of the horn to obtain the radiated pressure.

Msallam *et al.* (1997), developed a simple time-domain model that includes wave steepening in the slide of the trombone. The slide was treated as a single 3 m long cylindrical tube, and thermoviscous losses at the walls of the slide were applied prior to the wave steepening calculation. This simplified method of handling wall losses may have been adopted in order to avoid the numerous convolution sums encountered if the slide had been treated instead as a set of many shorter elements. To obtain the radiated sound, the steepened wave was processed by a linear filter representing the bell. Although their results were encouraging, Msallam *et al.*, did not quantitatively compare the output of their model to that of a real instrument.

The present research attempts to build on the work of Msallam *et al.*, by exploring the effect of wave steepening on sustained trombone tones in the frequency domain, both computationally and experimentally. Thermoviscous losses are more easily handled in the frequency domain, allowing the slide to be broken up into many smaller elements without incurring excessive computational load. Also, wave steepening is more systematically applied throughout the entire instrument—not just in the slide. A model has been developed that predicts well the radiated spectrum of the trombone at several combinations of dynamic and pitch. The output of the model is compared to experimentally measured tones from the same trombone that the model is based on, with good success.

II. THE FREQUENCY-DOMAIN MODEL

The instrument being modeled was a King Cleveland 605 tenor trombone with a Vincent Bach 12C mouthpiece. The shape of the instrument was approximated with a set of 152 cylindrical tubes; the lengths and radii of the cylinders were varied to provide a match to the trombone bore (Plitnik and Strong, 1979). All dimensions except those of the mouthpiece were taken from Copley's previous work (1995) on the same trombone. (Detailed dimensions are available in Appendix I of Thompson, 2000.)

A frequency-domain model of sound production was derived by considering lossy wave propagation in the cylinders and pressure reflection at the cylinder junctions. Since only steady, periodic waves were considered, the waves were represented as sums of harmonic components. Various parameters, including reflection coefficients and impedances, were specified only at harmonic frequencies. (A detailed derivation of the model is given in Appendices C and D of Thompson, 2000.)

Throughout this paper, the model is called linear when wave steepening is ignored and nonlinear when wave steep-

 0001-4966/2001/110(1)/556/7/$18.00

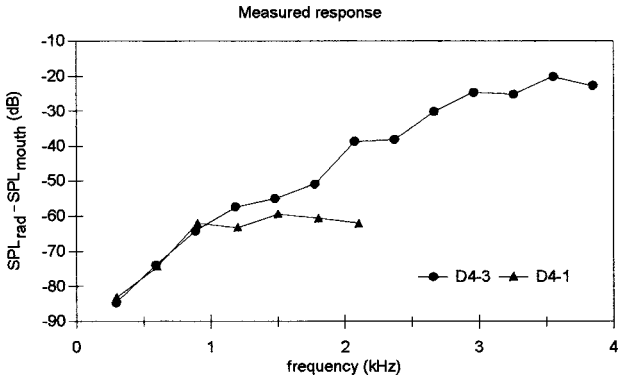


FIG. 1. Difference between measured radiated SPLs (see Figs. 3 and 4) and measured mouthpiece SPLs (see Fig. 2) at the pitch D4. D4-3 is a loud tone and D4-1 is a soft tone. The divergence of the two curves is evidence of amplitude-dependent behavior. Only values at harmonic frequencies are included.

ening is included. Also, a time dependence of $e^{j\omega t}$ is assumed throughout and has been omitted in order to simplify the notation.

A. The linear model

The equations for the linear model are given for a single frequency to simplify the notation. The subscript n refers to cylinder number, where $1 \leqslant n \leqslant N$ and $N=152$.

The trombone can be characterized acoustically in terms of the complex pressure reflection coefficients at the input ends (the ends nearest to the mouthpiece) of the cylinders. At the input of a given cylinder, the reflection coefficient is

$$R_n = \frac{P_n^-}{P_n^+}, \tag{1}$$

where P_n^+ and P_n^- are the outgoing and incoming complex pressure amplitudes, which can be treated as complex Fourier coefficients. Let us also define a dimensionless input impedance

$$Z_n' = \frac{A_{n-1}k_{n-1}}{\rho_0 \omega} Z_n, \tag{2}$$

where Z_n is the acoustic input impedance, A_{n-1} and k_{n-1} are the cross-sectional area and complex wave number of the adjacent cylinder nearest to the mouthpiece, ρ_0 is the ambient density of air, and ω is the angular frequency. At the junction of two cylinders, pressure and flow are continuous, which implies that acoustic impedance is also continuous. From these conditions (and after some algebra), the following equations can be obtained:

$$R_n = e^{-2jk_nL_n}\left(\frac{Z_{n+1}'-1}{Z_{n+1}'+1}\right) \tag{3}$$

and

$$Z_n' = \frac{A_{n-1}k_{n-1}}{A_n k_n}\left(\frac{1+R_n}{1-R_n}\right), \tag{4}$$

where L_n is the cylinder length. The dimensionless radiation impedance at the bell is

$$Z_{N+1}' = \frac{A_N k_N}{\rho_0 \omega} Z_{\text{bell}}, \tag{5}$$

where Z_{bell} is the acoustic radiation impedance of an unbaffled, plane circular piston in the end of a long cylindrical tube (Beranek, 1996). The reflection coefficients of all cylinders can be obtained by starting with Eq. (5) at the bell and alternately calculating Eqs. (3) and (4) back to the mouthpiece.

An experimentally measured pressure wave in the mouthpiece serves as the input to the model. Let the subscript ''mouth'' represent the index of the cylinder whose input end is nearest to the position where the mouthpiece pressure is known, and let P_{mouth} be a complex Fourier coefficient of the measured pressure wave. This Fourier coefficient can be separated into outgoing and incoming parts using linear theory. The outgoing pressure in the mouthpiece is

$$P_{\text{mouth}}^+ = \frac{P_{\text{mouth}}}{1+R_{\text{mouth}}}, \tag{6}$$

and the outgoing pressure at each of the subsequent cylinder inputs is

$$P_{n+1}^+ = P_n^+ e^{-jk_nL_n}\left(\frac{1+R_n e^{2jk_nL_n}}{1+R_{n+1}}\right). \tag{7}$$

Equation (7) is calculated repeatedly to model the filtering that occurs as the outgoing wave propagates through each cylinder in the model. The total pressure at the bell is

$$P_{\text{bell}} = P_N^+ e^{-jk_NL_N}\left(\frac{2Z_{N+1}'}{Z_{N+1}'+1}\right). \tag{8}$$

At a sufficiently large distance from the bell, the on-axis pressure is approximately equal to the pressure that would be radiated by a simple source having the same source strength. The radiated pressure can be expressed in terms of the total pressure at the bell and the radiation impedance as

$$P_{\text{rad}} = \frac{P_{\text{bell}}}{Z_{\text{bell}}}\frac{j\rho_0\omega}{4\pi x_{\text{rad}}}\exp\left(-j\frac{\omega}{c_0}x_{\text{rad}}\right), \tag{9}$$

where x_{rad} is the distance from the plane of the bell, and c_0 is the sound speed in air. From these predicted Fourier coefficients, the far-field radiated pressure wave can be synthesized.

B. Inclusion of wall losses

Thermoviscous losses at the walls of the cylinders can be included if we allow the wave number to be complex valued (Kinsler *et al.*, 1982):

$$k_n = \frac{\omega}{c_n} - j\alpha_n, \tag{10}$$

where

$$\alpha_n = \frac{1}{r_n c_0}\sqrt{\frac{\eta_{\text{eff}}\omega}{2\rho_0}} \tag{11}$$

and

$$c_n = c_0\left(1 - \frac{1}{2r_n}\sqrt{\frac{2\eta_{\text{eff}}}{\rho_0\omega}}\right) \tag{12}$$

are the attenuation coefficient and sound speed associated with these losses. In these expressions, r_n is the cylinder radius; and

$$\eta_{\text{eff}} = \eta\left[1 + (\gamma - 1)\sqrt{\frac{\kappa}{C_P\eta}}\right]^2 \tag{13}$$

is an effective viscosity, where η is the shear viscosity of air, γ is the adiabatic constant of air, κ is the thermal conductivity of air, and C_P is the specific heat of air at constant pressure.

C. The nonlinear model

The discussion of wave steepening that follows applies to the case of an outgoing wave in a single cylinder; the cylinder number index, *n*, is omitted and a harmonic number index, *h*, is used instead to handle the nonlinear interaction of harmonics, where $1 \leqslant h \leqslant H$ and $H = 200$. The choice of 200 harmonics was somewhat arbitrary and limited by the available computational resources. A much greater number of harmonics would be required in order to accurately model shock formation.

The Burgers equation predicts the wave steepening that occurs as an outgoing pressure wave propagates through each cylinder in the linear model. In the frequency domain, this equation is (Ginsberg and Hamilton, 1998; Appendix D of Thompson, 2000)

$$\frac{d}{dx}P_h^+ = -\frac{\delta\omega_h^2}{2c_0^3}P_h^+ + \frac{\beta\omega_k}{\rho_0 c_0^3}\frac{j}{4}\left\{\sum_{h'=1}^{h-1}(P_{h'}^+ P_{h-h'}^+) + 2\sum_{h'=h+1}^{\infty}[P_{h'}^+(P_{h'-h}^+)^*]\right\}, \tag{14}$$

where x is the distance from the input end of the cylinder; ω_h is the angular frequency of harmonic h;

$$\delta = \frac{1}{\rho_0}\left(\frac{4}{3}\eta + \eta_B\right) + \frac{\kappa}{\rho_0}\left(\frac{1}{C_V} - \frac{1}{C_P}\right) \tag{15}$$

is the diffusivity of sound (Hamilton and Morfey, 1998), where η_B is the bulk viscosity of air, and C_V is the specific heat of air at constant volume;

$$\beta = \frac{\gamma + 1}{2} \tag{16}$$

is the coefficient of nonlinearity; and * represents the complex conjugate. The first term on the right-hand side of Eq. (14) is proportional to the attenuation in the bulk of the fluid (Ginsberg and Hamilton, 1998),

$$\theta_h = \frac{\delta\omega_h^2}{2c_0^3}, \tag{17}$$

which reduces the amplitudes of the high-frequency harmonics as they increase due to wave steepening. The second term on the right-hand side of Eq. (14) is inversely related to the shock formation distance (Blackstock *et al.*, 1998)

$$x_{\text{shock}} = \frac{\rho_0 c_0^3}{\beta \max\left[\frac{d}{dt}p_{\text{mouth}}^+(t)\right]}. \tag{18}$$

The shock formation distance is inversely related to frequency and to acoustic pressure amplitude through the maximum value of the time derivative of the outgoing mouthpiece pressure wave in the denominator. This frequency dependence appears in the second term on the right-hand side of Eq. (14) as ω_h, and the amplitude dependence appears in the sums as pressure squared. The nonlinearities described by this equation are maximal at large pressure amplitudes and at high frequencies, or equivalently, at small shock formation distances.

It should be noted that Eq. (14) accounts for changes in the outgoing Fourier coefficients due only to wave steepening and thermoviscous losses in the bulk of the fluid; phase change due to propagation and losses at the wall of the cylinder were already included in Eq. (7). Wall and bulk losses are handled asymmetrically because wall losses are best applied to both the outgoing and the incoming waves via the complex wave number, while bulk losses are needed in the steepening of the outgoing wave in order to control the amplitude of the high-frequency harmonics during shock formation.

A good numerical method for solving Eq. (14) is fourth order Runge–Kutta. Since this method requires real equations of real variables, this equation must be separated into real and imaginary components:

$$\frac{d}{dx}X_h = -\theta_h X_h - \phi_h\left[\sum_{h'=1}^{h-1}(X_{h'}Y_{h-h'} + Y_{h'}X_{h-h'}) + 2\sum_{h'=h+1}^{H}(Y_{h'}X_{h'-h} - X_{h'}Y_{h'-h})\right] \tag{19}$$

and

$$\frac{d}{dx}Y_h = -\theta_h Y_h + \phi_h\left[\sum_{h'=1}^{h-1}(X_{h'}X_{h-h'} - Y_{h'}Y_{h-h'}) + 2\sum_{h'=h+1}^{H}(X_{h'}X_{h'-h} + Y_{h'}Y_{h'-h})\right] \tag{20}$$

where the substitutions

$$X_h + jY_h = P_h^+ \tag{21}$$

and

$$\phi_h = \frac{\beta\omega_h}{\rho_0 c_0^3}\frac{1}{4} \tag{22}$$

have been made to simplify the notation, and the Fourier series has been truncated at H harmonics.

Wave steepening can be included in the linear algorithm as follows: At the input of the cylinder in the mouthpiece where the measured pressure wave is known, the outgoing pressure coefficients are computed from Eq. (6). If desired, the shock formation distance can then be computed from Eq.

(18). The maximum value of the time derivative in this equation is best obtained by evaluating the analytical time derivative of the Fourier synthesis equation using a large number of samples per period and taking the maximum sample value. The outgoing mouthpiece pressure coefficients are modified by Eqs. (19) and (20), which models the wave steepening that would result from a propagation distance equal to the length of the cylinder. In this calculation, the algorithm is allowed to take a maximum spatial step of 5% of the shock formation distance in order to prevent large numerical error. Equation (7) is then computed to account for phase change, losses at the wall, and filtering at the cylinder junction. The process is repeated for each cylinder in the model. After wave steepening has been calculated for the final cylinder at the bell, the pressures at the bell and in the far field are computed from Eqs. (8) and (9).

Although the superposition of outgoing and incoming waves is not strictly valid for a nonlinear system, Msallam *et al.* (1997) concluded that errors due to this oversight would be less than 5% for harmonics below the 30th. They also reported that steepening the incoming wave did not significantly affect their results. This supports the suggestion of Hirschberg *et al.* (1996), that the gain in high-frequency energy due to wave steepening is mostly radiated at the bell, implying that steepening need be calculated for only the outgoing wave.

D. Testing the model

Several tests were performed to determine the reliability of the model; each test confirmed that the model was properly implemented. Errors on the order of a few dB are tolerable since the phenomenon of interest—the gain in high-frequency energy due to wave steepening—is on the order of tens of dB (see Fig. 1).

The input impedances of three theoretical tube shapes—a cylindrical pipe, a conical horn, and an exponential horn—were calculated based on approximations using sets of contiguous cylinders. The impedance magnitudes and phases were compared to the analytical results (Fletcher and Rossing, 1991) with good agreement. The error in the cylindrical case was negligible. The exponential case had the largest error—on the order of a few dB for frequencies below 20 kHz when using 150 cylinders. Admittedly, a more accurate (and mathematically complex) model could be implemented using cones instead of cylinders.

The input impedance of the trombone at the mouthpiece was calculated and compared to measured and calculated results from a different model of the same instrument (Copley, 1995). The resonance frequencies and corresponding impedance magnitudes agreed well.

Wave steepening of a sinusoid was calculated over one shock formation distance using 10 and 20 spatial steps. The differences in harmonic amplitudes between the two cases were a fraction of a dB. This suggests that allowing a maximum spatial step of 5% of the shock formation distance keeps the numerical error of the Runge–Kutta algorithm within acceptable limits.

III. EXPERIMENTAL VERIFICATION

A. Collecting the data

The model was thoroughly tested with experimentally measured tones from the same trombone that the model is based on. In an anechoic chamber, the trombone was mounted at a comfortable height for playing, with the bell oriented horizontally. The diaphragm of a PCB 112A23 microphone was mounted flush with the inside wall of the mouthpiece by means of a hole drilled through the wall. The signal from this microphone was recorded to one channel of a digital audio tape. A 1/2″ Larson–Davis 2540 microphone was mounted on-axis 2.85 m in front of the bell, and the signal from this microphone was recorded to the second channel of the tape. The bandwidth of the Larson–Davis microphone is approximately 20 Hz to 20 kHz, and the usable bandwidth of the PCB microphone was measured to be approximately 100 Hz to 16 kHz. Both microphones were calibrated so that accurate pressure values could be obtained from the recorded data. Sustained tones were recorded at the pitches B♭2, F3, B♭3, and D4 at various dynamics. The sample rate was 44.1 kHz. (Details of the experimental apparatus, calibration, and procedure can be found in Appendix G of Thompson, 2000.)

B. Preparing and processing the data

Six tones at the lowest pitch and three tones at each of the other pitches were chosen for analysis. These tones ranged from the softest to the loudest that the player could produce comfortably. Each of the tones was labeled with its pitch name and octave number and with a number representing the dynamic at which it was played. The dynamics were numbered in increasing order, from soft to loud. For example, the tone D4-1 is the softest tone analyzed at the pitch D in the fourth octave.

A steady segment of 10 periods was extracted from each of the B♭2 tones. The numbers of periods extracted from the tones at the other pitches (see Table I) were chosen so that all segments had approximately the same duration ($\sim$ 8.5 ms). The measured fundamental frequencies (see Table I) of these tones agreed well with the fundamental frequencies predicted by linear theory (Copley, 1995), and were therefore used in the model computations. The complex Fourier coefficients of the mouthpiece and the far-field radiated waves of each segment were calculated. Because all harmonics above a certain frequency (see Table I) had amplitudes less than the noise floor, data above that frequency were discarded. Figure 2 shows the mouthpiece spectra for the tones D4-1 and D4-3.

The Fourier series of each segment of mouthpiece data was processed by the model to predict the measured pressure at the far-field location. Each tone was processed twice—once neglecting wave steepening and once including wave steepening. The shock formation distances (see Table I), the SPLs of all Fourier coefficients (relative to 20 μPa), the errors in SPLs between the measured and predicted far-field coefficients, and the synthesized far-field pressure waves were calculated for all tones. (Complete experimental and computational data are given in Appendix H of Thompson, 2000.)

TABLE I. Tone parameters. The shock formation distances [see Eq. (18)] were computed from the outgoing mouthpiece pressures. For comparison, the length of the trombone is 2.72 m. The mouthpiece signal cutoff frequencies were the frequencies above which noise in the mouthpiece signal prevented accurate estimation of the harmonic amplitudes. They are also the maximum frequencies used in computing the errors in Fig. 7.

Tone	Fundamental frequency (Hz)	Number of periods analyzed	Shock formation distance (m)	Mouthpiece signal cutoff frequency (kHz)
B♭2-1	116.98±0.04	10	44	0.936
B♭2-2	117.07±0.04	10	11	1.874
B♭2-3	118.90±0.04	10	4.5	2.735
B♭2-4	119.58±0.04	10	3.0	2.990
B♭2-5	118.68±0.04	10	2.0	4.154
B♭2-6	117.04±0.04	10	0.90	4.916
F3-1	177.68±0.05	15	34	1.422
F3-2	177.16±0.05	15	3.3	2.658
F3-3	176.78±0.05	15	1.5	3.182
B♭3-1	238.83±0.07	20	20	1.672
B♭3-2	237.03±0.07	20	4.7	2.134
B♭3-3	235.51±0.07	20	1.5	3.298
D4-1	300.82±0.09	25	16	2.106
D4-2	297.65±0.09	25	3.4	2.679
D4-3	296.13±0.09	25	1.6	3.850

C. Selected results

The results from the softest and loudest tones at the pitch D4 (tones D4-1 and D4-3) typify the results obtained for all other tones. They will be used as examples in the discussion that follows.

Both the linear and the nonlinear models predict approximately the same radiated spectrum for the soft tone (see Fig. 3). In the case of the loud tone (see Fig. 4), however, the nonlinear model predicts the measured spectrum reasonably well, while the linear model severely underestimates the amount of high frequency energy in the measured spectrum.

A comparison of the measured and predicted pressure waves confirms the superiority of the nonlinear model. For the soft tone [see plots (b) through (d) of Fig. 5], the waves predicted by both models have amplitudes and phases similar to the measured wave. In the case of the loud tone [see plots (b) through (d) of Fig. 6], however, the nonlinear model predicts an amplitude that is reasonably correct, while the linear model predicts an amplitude that is much too small.

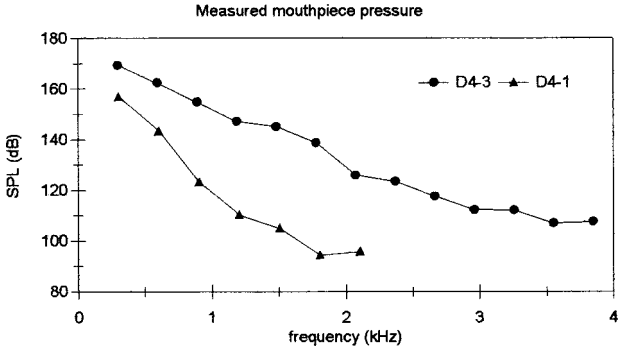


FIG. 2. Measured mouthpiece pressure spectra for the tones D4-1 and D4-3. Only values above the noise floor and at harmonic frequencies are included. Mouthpiece wave forms are shown in plots (a) of Figs. 5 and 6.

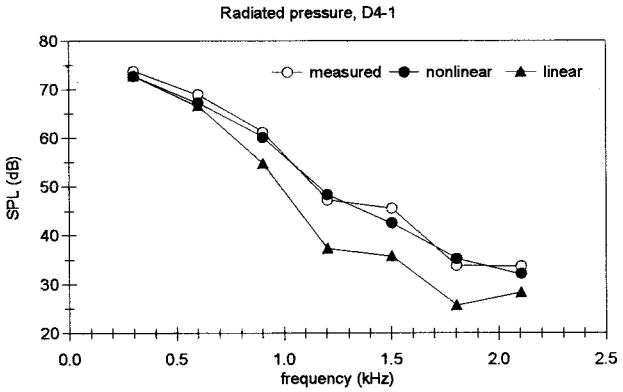


FIG. 3. Radiated pressure spectra for the tone D4-1. The SPLs predicted by the nonlinear and linear models are compared to the SPLs measured on-axis 2.85 m in front of the bell. Only values at harmonic frequencies are included. Radiated wave forms are shown in plots (b)–(d) of Fig. 5.

At extremely loud dynamics corresponding to shock formation distances significantly less than the length of the instrument, the nonlinear model tends to overestimate the amplitudes of the high-frequency harmonics. This problem has been reported previously for a similar computation (Anderson and Vaidya, 1991). This error is mostly due to the truncation of the Fourier series at 200 harmonics. Another contributing factor may be that the calculation of wave steepening and the calculation of wall losses are not done simultaneously for the outgoing wave in a given cylinder. This factor may be lessened by dividing the longest cylinders in the approximation into several shorter cylinders, thereby distributing the wall losses more uniformly.

Figure 7 compares the errors in SPLs from both models for all 15 tones. The nonlinear model clearly does better than the linear model at predicting the pressure at the far-field location. The anomalies in the data (i.e., B♭ 2-4 and F3-3) can be attributed to noise in the mouthpiece data that was not completely removed before processing.

In informal listening tests, the tones generated by the nonlinear model were found to be similar in timbre to the experimentally measured tones, but the tones generated by the linear model were found to be lacking in brightness.

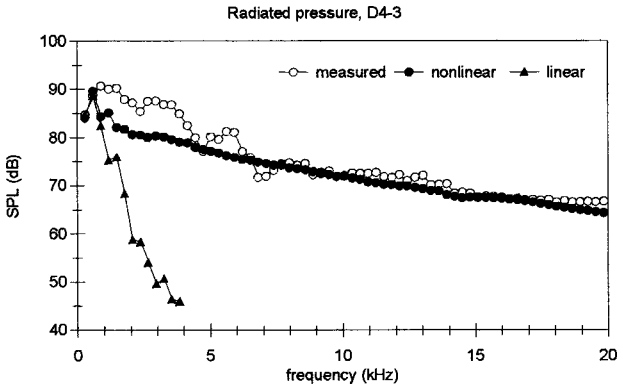


FIG. 4. Radiated pressure spectra for the tone D4-3. The SPLs predicted by the nonlinear and linear models are compared to the SPLs measured on-axis 2.85 m in front of the bell. Only values at harmonic frequencies are included. Radiated wave forms are shown in plots (b)–(d) of Fig. 6.

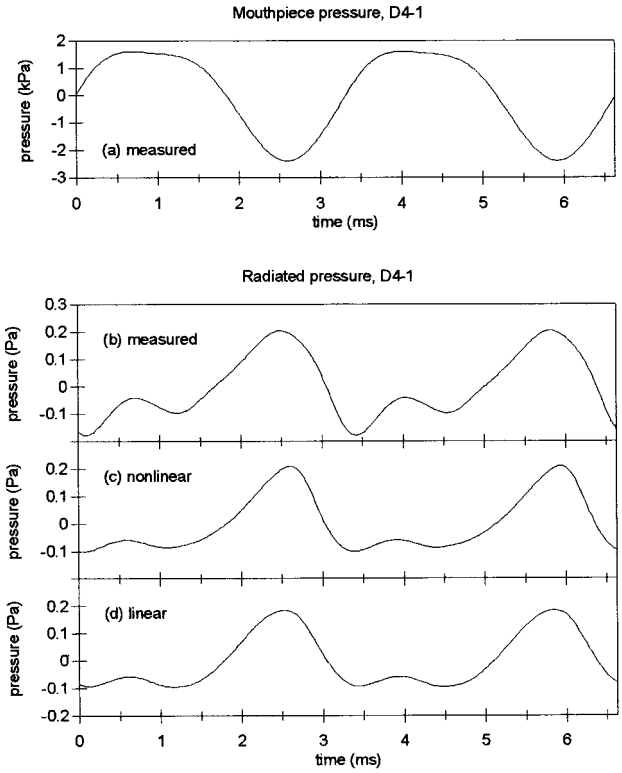


FIG. 5. Acoustic pressure wave forms for the tone D4-1. (a) Measured wave in the mouthpiece. (b) Measured wave on-axis 2.85 m in front of the bell. (c) Radiated wave computed from the nonlinear model. (d) Radiated wave computed from the linear model. DC offset has been neglected. The mouthpiece spectrum is shown in Fig. 2. The radiated spectra are shown in Fig. 3.

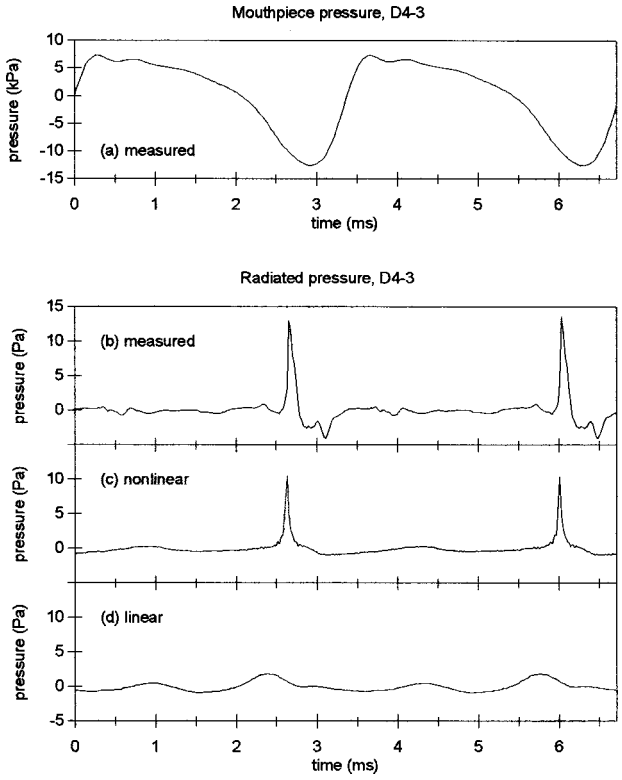


FIG. 6. Acoustic pressure wave forms for the tone D4-3. (a) Measured wave in the mouthpiece. (b) Measured wave on-axis 2.85 m in front of the bell. (c) Radiated wave computed from the nonlinear model. (d) Radiated wave computed from the linear model. DC offset and frequencies above 20 kHz have been neglected. The mouthpiece spectrum is shown in Fig. 2. The radiated spectra are shown in Fig. 4.

IV. CONCLUSIONS

The results are indeed encouraging. For small-amplitude tones, both the linear and the nonlinear models predict well the measured data. At large amplitudes, the nonlinear model is reasonably accurate in its predictions, while the linear model severely underestimates the amplitudes of the high-frequency harmonics. However, at shock formation distances significantly less than the length of the instrument, the nonlinear model tends to overestimate the amplitudes of the high-frequency harmonics.

The model may be improved upon by using a greater number of cylinders in the approximation, which would effectively distribute the losses at the walls of the cylinders more uniformly. Including more harmonics in the computation would help to reduce numerical error in the wave steepening calculation. Using mouthpiece data with a lower noise floor would permit evaluation of the model at higher frequencies. A more accurate model could be formulated using a set of contiguous cones instead of cylinders.

Although direct comparisons between the present research and the work of Msallam *et al.* (1997), cannot be made conclusively, this research builds upon their work by providing an alternate method for including wave steepening in a computational model of trombone sound production: Thermoviscous wall losses are handled in a more realistic manner, wave steepening is applied systematically to the entire instrument, and the accuracy of the model is analyzed

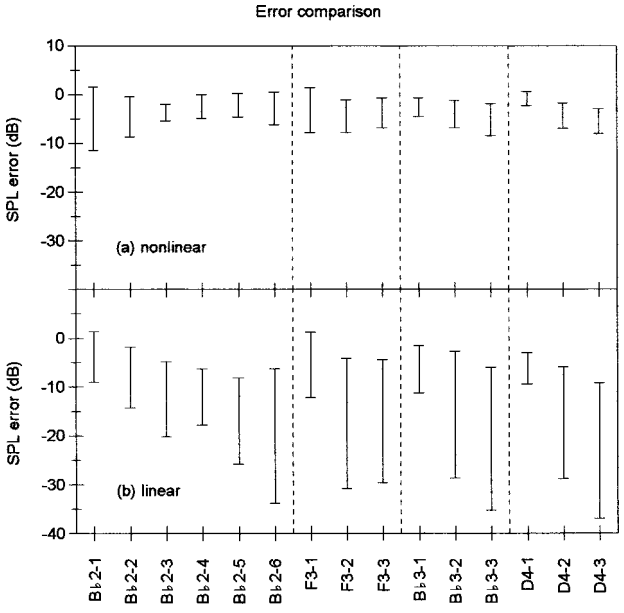


FIG. 7. Comparison of the average errors between measured and computed harmonic SPLs from both models. Each line ranges from the mean error in SPLs minus the standard deviation of the error to the mean plus the standard deviation. Only frequencies below the mouthpiece signal cutoff frequencies (see Table I) were included. The different dynamics at each pitch are numbered in increasing order, from soft to loud.

quantitatively through comparison of its output to experimentally measured tones.

ACKNOWLEDGMENTS

The trombone was played by Gordon Dix. This research was funded in part by a grant from the Institute for Scientific Research in Music.

APPENDIX A: VALUES OF PHYSICAL CONSTANTS

The following values were used for the physical constants that appear in the equations of the model. Approximations were made where precise values could not be determined. Although the values chosen may not perfectly reflect the conditions under which the experiment was conducted, a small change in any of the values does not significantly alter the computational results:

$$c_0 = 350\,\frac{\mathrm{m}}{\mathrm{s}}, \tag{A1}$$

$$C_P = 957.6\,\frac{\mathrm{J}}{\mathrm{kg\,K}}, \tag{A2}$$

$$C_V = \frac{C_P}{\gamma} = 683.0\,\frac{\mathrm{J}}{\mathrm{kg\,K}}, \tag{A3}$$

$$\beta = 1.201 \quad [\text{see Eq. (16)}], \tag{A4}$$

$$\gamma = 1.402, \tag{A5}$$

$$\delta = 3.80\times 10^{-5}\,\frac{\mathrm{m}^2}{\mathrm{s}} \quad [\text{see Eq. (15)}], \tag{A6}$$

$$\eta = 1.81\times 10^{-5}\ \mathrm{Pa\,s}, \tag{A7}$$

$$\eta_B = 0.6\,\eta = 1.09\times 10^{-5}\ \mathrm{Pa\,s}, \tag{A8}$$

$$\eta_{\mathrm{eff}} = 3.90\times 10^{-5}\ \mathrm{Pa\,s} \quad [\text{see Eq. (13)}], \tag{A9}$$

$$\kappa = 0.0234\,\frac{\mathrm{W}}{\mathrm{m\,K}}, \tag{A10}$$

$$\rho_0 = 1.18\,\frac{\mathrm{kg}}{\mathrm{m}^3}. \tag{A11}$$

Anderson, M. J., and Vaidya, P. G. (**1991**). ''Thermo-viscous effects on finite amplitude sound propagation in a rectangular waveguide,'' J. Acoust. Soc. Am. **90**, pp. 1056–1067.

Beauchamp, J. W. (**1980**). ''Analysis of simultaneous mouthpiece and output waveforms of wind instruments,'' Audio Engineering Society preprint No. 1626.

Beranek, L. L. (**1996**). *Acoustics*, 5th ed. (Acoustical Society of America, Woodbury, NY), pp. 116–128.

Blackstock, D. T., Hamilton, M. F., and Pierce, A. D. (**1998**). ''Progressive waves in lossless and lossy fluids,'' in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, San Diego), pp. 65–76.

Copley, D. C. (**1995**). M. S. thesis, Brigham Young University.

Fletcher, N. H., and Rossing, T. D. (**1991**). *The Physics of Musical Instruments* (Springer-Verlag, New York), pp. 183–194.

Ginsberg, J. H., and Hamilton, M. F. (**1998**). ''Computational methods,'' in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, San Diego), pp. 311–319.

Hamilton, M. F., and Morfey, C. L. (**1998**). ''Model equations,'' in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, San Diego), pp. 53–54.

Hirschberg, A. J., Gilbert, J., Msallam, R., and Wijnands, A. P. J. (**1996**). ''Shock waves in trombones,'' J. Acoust. Soc. Am. **99**, 1754–1758.

Kinsler, L. E., Frey, A. R., Coppens, A. B., and Sanders, J. V. (**1982**). *Fundamentals of Acoustics*, 3rd ed. (Wiley, New York), pp. 206–210.

Msallam, R., Dequidt, S., Tassart, S., and Caussé, R. (**1997**). ''Physical model of the trombone including nonlinear propagation effects,'' Proceedings of the Institute of Acoustics, **19**(5), pp. 419–424.

Plitnik, G. R., and Strong, W. J. (**1979**). ''Numerical method for calculating input impedances of the oboe,'' J. Acoust. Soc. Am. **65**, 816–825.

Thompson, M. W. (**2000**). M. S. thesis, Brigham Young University.

Melody lead in piano performance: Expressive device or artifact?

Werner Goebl[a)]
Austrian Research Institute for Artificial Intelligence, Schottengasse 3, A-1010 Vienna, Austria

(Received 3 July 2000; revised 2 March 2001; accepted 6 April 2001)

As reported in the recent literature on piano performance, an emphasized voice (the melody) tends to be played not only louder than the other voices, but also about 30 ms earlier (*melody lead*). It remains unclear whether pianists deliberately apply melody lead to separate different voices, or whether it occurs because the melody is played louder (*velocity artifact*). The velocity artifact explanation implies that pianists initially strike the keys simultaneously; it is only different velocities that make the hammers arrive at different points in time. The measured note onsets in these studies, mostly derived from computer-monitored pianos, represent the hammer-string impact times. In the present study, the finger-key contact times are calculated and analyzed as well. If the velocity artifact hypothesis is correct, the melody lead phenomenon should disappear at the finger-key level. Chopin's Ballade op. 38 (45 measures) and Etude op. 10/3 (21 measures) were performed on a Bösendorfer computer-monitored grand piano by 22 skilled pianists. The hammer-string asynchronies among voices closely resemble the results reported in the literature. However, the melody lead decreases almost to zero at the finger-key level, which supports the velocity artifact hypothesis. In addition to this, expected onset asynchronies are predicted from differences in hammer velocity, if finger-key asynchronies are assumed to be zero. They correlate highly with the observed melody lead.
[DOI: 10.1121/1.1376133]

PACS numbers: 43.75.St, 43.75.Mn [RDA]

I. INTRODUCTION

Simultaneous noted in the printed score (chords) are not played strictly simultaneously by pianists. An emphasized voice is not only played louder, but additionally precedes the other voices typically by around 30 ms; this phenomenon is referred to as *melody lead* (Hartmann, 1932; Vernon, 1937; Palmer, 1989, 1996; Repp, 1996b). It is still unclear whether this phenomenon is part of the pianists' deliberate expressive strategies and used independently from other expressive parameters (Palmer, 1996), or whether it is mostly due to the timing characteristics of the piano action (*velocity artifact*, Repp, 1996b), a result of the dynamic differentiation of different voices. Especially in chords played by the right hand, high correlations between hammer velocity differences and melody lead times (between melody notes and accompaniment) seem to confirm this velocity artifact explanation (Repp, 1996b).

The data used in previous studies, derived mostly from computer-monitored pianos, represent asynchronies at the hammer-string contact points. The present study examined asynchrony patterns at the finger-key contact points as well. Finger-key asynchronies represent what pianists initially do when striking chords. If the velocity artifact explanation is correct, the melody lead phenomenon should disappear at the finger-key level. This means that pianists tend to strike the keys almost simultaneously, and it is only the different dynamics (velocities) that result in the typical hammer-string asynchronies (*melody lead*).

a)Electronic mail: werner.goebl@ai.univie.ac.at

A. Background

In considering note onset asynchronies, one has to differentiate between asynchronies that are indicated in the score (arpeggios, apoggiaturas) and asynchronies that are performed but not especially marked in the score. The latter come in two kinds: (1) The melody precedes other voices by about 30 ms on average (*melody lead*), or (2) the melody lags behind the other voices. Asynchronies of the second type occur mainly between the two hands and usually show much larger timing differences (over 50 ms). A typical example would be when a bass note is played clearly before the melody (*melody lag* or *bass lead*), which is well known from old recordings of piano performances, but has been observed in contemporary performances too (Palmer, 1989; Repp, 1996b). Asynchronies of the first type are common within one hand (especially within the right hand, as the melody often is the highest voice), but may also occur between the hands.

Note asynchronies have been studied since the 1930s, when Hartmann (1932) and the Seashore group (Vernon, 1937) conducted the first objective investigations of piano performances. Hartmann used piano rolls as a data source and found mostly asynchronies of the second type. Vernon (1937) differentiated between asynchronies *within* one hand and asynchronies *between* different hands. For the former he observed melody lead (type 1), whereas the latter mostly showed bass note anticipation (type 2).

In the recent literature, Palmer (1989, 1996) and Repp (1996b) have studied the melody lead phenomenon. Palmer (1989) used electronic keyboard recordings to analyze chord asynchronies, among other issues. Six pianists played the beginning of the Mozart Sonata K. 331 and of Brahms' In-

termezzo op. 117/1 (''Schlaf sanft, mein Kind...''). The melody led by about 20 to 30 ms on average; this effect decreased for deliberately ''unmusical'' performances and for melody voices in the middle of a chord (Brahms op. 117/1). In a second study, melody lead was investigated exclusively (Palmer, 1996). Six pianists played the first section of Chopin's Prelude op. 28/15 and the initial 16 bars of Beethoven's Bagatelle op. 126/1 on a Bösendorfer computer-monitored grand piano (SE290, as in the current study). Again, melody lead was found to increase with intended expressiveness, also with familiarity with a piece (the Bagatelle was sight-read and repeated several times), and with skill level (expert pianists showed a larger melody lead than student pianists).

In another study published at the same time, in part with the same music, Repp (1996b) analyzed 30 performances by 10 pianists of the whole Chopin Prelude op. 28/15, a Prelude by Debussy, and ''Träumerei'' by Schumann on a Yamaha upright Disklavier. To reduce random variation, Repp averaged over the three performances produced by each pianist. He then calculated timing deviations between the (right hand) melody and each other voice, so that asynchronies within the right hand and between hands could be treated separately. He argued that melody lead could be explained mostly as a consequence of dynamic differences between melody and accompaniment. Dynamic differences (differences in MIDI velocity) were positively correlated with timing differences between the melody and each of the other voices, and these correlations were generally higher for asynchronies within the right hand than for those between hands.

Palmer (1996) also computed correlations between melody lead and the average hammer velocity difference between melody and accompaniment, but her correlations were mostly nonsignificant. In her view, the anticipation of the melody voice is primarily an expressive strategy that is used independently from other performance parameters such as intensity, articulation, and pedal use. In a perception test, listeners had to identify the intended melody in a multivoiced piece by rating different artificial versions: one with intensity differences and melody lead, one with melody lead only, and one without any such differences. Melody identification was best for the original condition (melody lead and intensity difference), but the results in the melody lead condition did not differ much from the results in the neutral condition, especially for nonpianist listeners. Only pianist listeners showed some success in identifying the intended melody from melody leads alone. A condition with intensity differences only was not included (Palmer, 1996, p. 47).

B. Piano action timing properties

To fully explain the melody lead phenomenon, it is necessary first to clarify its physical underpinnings. The temporal parameters of the piano action have been described by Askenfelt (1990) and Askenfelt and Jansson (1990, 1991, 1992). When a key is depressed, the time from its initial position to the bottom contact ranges from 25 ms (*forte* or 5-m/s final hammer velocity, FHV) to 160 ms (*piano* or 1-m/s FHV; Askenfelt and Jansson, 1991, p. 2385).[1] In a grand piano the hammer impact times (when the hammer excites the strings) are shifted in comparison to key bottom contact times. According to measurements by Askenfelt (1990, p. 43), the hammer impact time is 12 ms before the key bottom contact at a *piano* touch (hammer velocity 1 m/s), but 3 ms *after* the key bottom contact at a *forte* attack (5 m/s). The timing properties of a grand piano action are outlined by these data, but more detailed data were not available (Askenfelt, 1999, personal communication).

The timing properties of the piano action can be modified by changing the regulation of the action. Modifications, e.g., in the hammer-string distance or in the *let-off distance* (the distance of free flight of the hammer, after the jack is released by the escapement dolly) affect the timing relation between hammer-string contact and key-bottom contact (Askenfelt and Jansson, 1990). Greater hammer mass in the bass (Conklin, 1996, p. 3287) influences the hammer-string contact durations (Askenfelt and Jansson, 1990), but not the timing properties of the action.

Another measurement was made by Repp (1996b) on a Yamaha Disklavier on which the ''prelay'' function was not working.[2] This gave him the opportunity to measure roughly a grand piano's timing characteristics in the middle range of the keyboard. He measured onset asynchronies at different MIDI velocities in comparison to a note with a fixed MIDI velocity. The time deviations extended over a range of about 110 ms for MIDI velocities between 30 and 100 and were fit well by a quadratic function (Repp, 1996b, p. 3920).

The timing characteristics of electronic keyboards vary across manufacturers and are rarely well documented. Each key has a spring with two electric contacts that define the off and on states. When a key is depressed, the spring contact is moved from the off position to the on position (see Van de Berghe *et al.*, 1995, p. 16). The time difference between the breaking of the off contact and the on contact determines the MIDI velocity values; the note onset is registered near the key bottom contact.

We now have to distinguish between asynchronies at the beginning of the attack movement (finger-key contact) and asynchronies at its end (hammer-string impact or key-bottom contact). Computer-monitored grand pianos, like those of Yamaha or Bösendorfer, store time points of hammer-string impact, which are essentially equivalent to the beginnings of the sound events (see Fig. 1).[3]

II. AIMS

Almost nothing is known about asynchronies at the finger-key level, because none of the instruments used for acquiring performance data measure this parameter. However, to clarify the origin of melody lead, it is important to consider exactly those finger-key asynchronies. When pianists stress one voice in a chord, do they hit the keys asynchronously or do their fingers push the keys down at the same time but with different velocities, so that the hammers arrive at the strings at different points in time?

To examine this question, it is necessary to determine the finger-key contact times. One possibility might be to observe finger key contacts by using a video camera or by special electronic measurements at the keyboard. In this study, the finger-key contacts were inferred from the time the

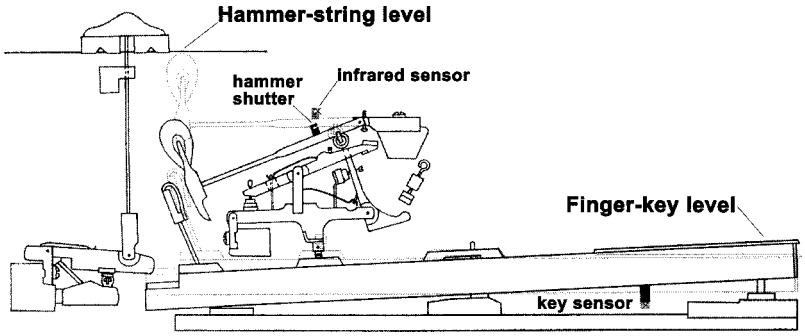


FIG. 1. Grand piano action, with the measurement points of the Bösendorfer SE system sketched. The infrared sensor mounted on the hammer flange rail captures two trip points of the shutter on the hammer shank: one 5 mm below the strings, the second immediately at the strings. The sensor at the key reacts when the key is depressed more than about 2 mm. (Figure prepared with computer software by the author.)

hammer travels from its resting position to the strings at different final hammer velocities (timing correction curve). With the help of this function, the finger-key contacts could be accurately estimated; also the size of the expected melody lead effect in milliseconds could be predicted from the velocity differences between the voices, assuming simultaneous finger-key contacts.

III. METHOD

A. Materials and participants

The *Etude* op. 10/3 (first 21 measures, Fig. 2) and the *Ballade* op. 38 (initial section, bars 1 to 45, Fig. 3) by Frédéric Chopin were recorded on a Bösendorfer SE290 computer-monitored concert grand piano by 22 skilled pianists (9 female and 13 male). They were professional pianists, graduate students or professors at the Universität für Musik und darstellende Kunst (University of Music and Performing Arts) in Vienna. They received the scores several days before the recording session, but were nevertheless allowed to use the music scores during recording. Their average age was 27 years (the youngest was 19, the oldest 51). They had received their first piano lesson at 6 and a half years of age on average. They had received piano instruction for a mean of 22 years (s.d.=7); 8 of them had already finished their studies; about half of them played more than 10 public concerts per year.

After the recording, the pianists were asked to play the initial 9 bars of the Ballade in two additional versions: first with a particularly emphasized highest voice (voice 1, see Fig. 3) and second with an emphasized third voice (the lowest voice in the upper stave, played also by the right hand, see Fig. 3). The purpose of these special versions was to investigate how pianists change melody lead and dynamic shaping of the voices when they were explicitly advised to emphasize one particular voice.

All performance sessions were recorded onto digital audio tape (DAT), and the performance data from the Bösen-

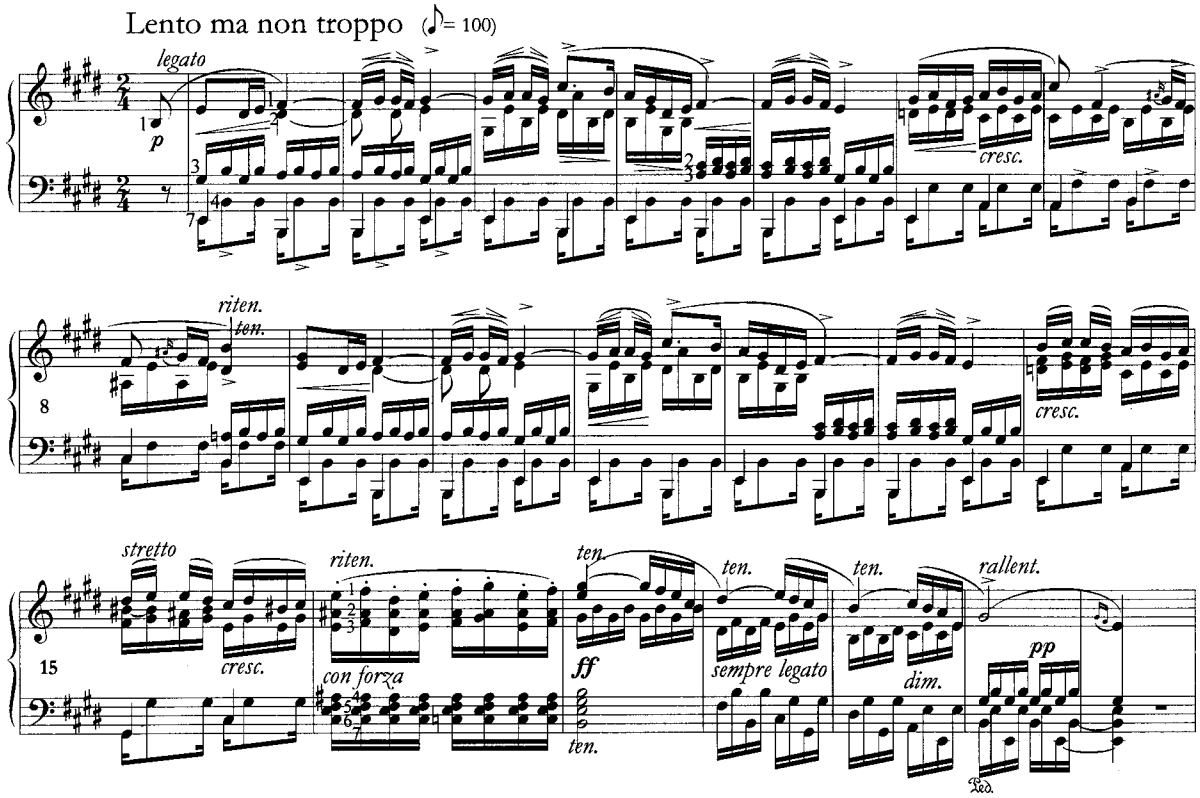


FIG. 2. Frédéric Chopin. Beginning of the Etude in E major, op. 10/3. The numbers against the note heads are voice numbers (soprano=1,...,bass=7). (Score prepared with computer software following Paderewski Edition.)

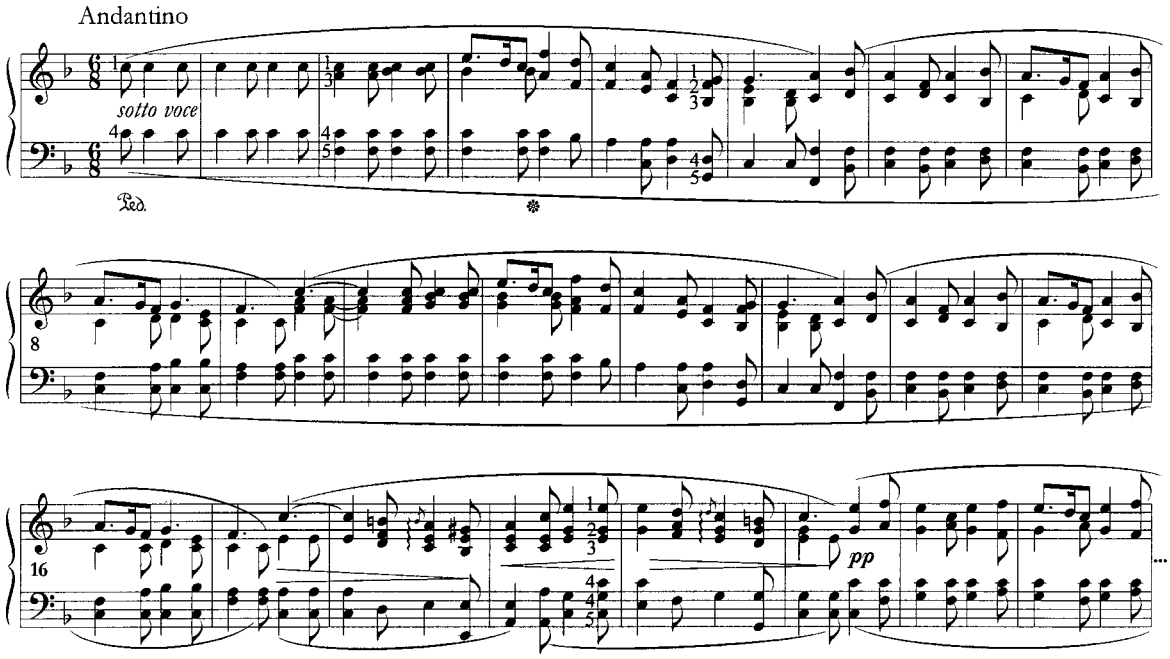


FIG. 3. Frédéric Chopin. The beginning of the second Ballade op. 38 in F major (the first 23 bars). The voices are numbered as in Fig. 2, but the highest number is now 5 for the bass. (Score prepared with computer software following Henle Urtext Edition.)

dorfer grand piano were stored on a PC's hard disk. The performances were consistently of a very high pianistic and musical level.[4] At the end of the session, the participants had to fill in a questionnaire

B. Apparatus

To provide accurate performance data, a Bösendorfer SE290 *Imperial* computer-monitored concert grand piano[5] was used. The exact recording and playback functionality of the Bösendorfer SE system is insufficiently described in the manual and literature (Moog and Rhea, 1990; Palmer and Brown, 1991; Alcedo and Schäfer, 1992; Repp, 1993; Palmer, 1996). Additional information was obtained from W. Stahnke (private communication, see Note 3) and from the Bösendorfer technician F. Lachnit.

The SE system (see Fig. 1) is equipped with two sets of shutters, one at the hammers, another one under the keys. The hammer shutters provide two trip points, one as the hammer crown just starts to contact the string and the other 5 mm lower. These two trip points capture two instants in time as the hammer travels upward, and the time difference between these instants yields an estimate of the *final hammer velocity* (FHV, in meters per second).[6]

The instant at which the trip point at the strings is passed is taken as the note onset time. The note onset times are taken 800 times per second, thus they have a timing resolution of 1.25 ms. The SE system has another set of shutters about 2–3 mm under the keys, which provides the note offset times and—only in the case of silent notes (if the hammer does not touch the strings)—notes onsets.

To avoid timing distortions in reproduction, the Bösendorfer SE290 uses a timing correction similar to the Yamaha Disklavier's ''prelay'' function (cf. Repp, 1996b). The Bössendorfer SE system recalculates the precise timing characteristics for each key individually by running a calibration program on demand. Among other parameters, the calibration function records the travel time interval from the key-shutter response (2–3 mm below key zero position) to the hammer string impact for seven final hammer velocities and all 97 keys (the Bösendorfer SE290 has 9 additional keys in the bass). This data matrix is stored in a hardware chip (EEPROM X2816AP). The memory of this hardware chip of the particular grand piano in Vienna used in the experiments was transferred into a file. The timing correction matrix (TCM) derived from these data is shown in Fig. 4. The data include both irregularities of the piano action and the electronic playback system. What can be seen from this matrix is that travel time does not depend on hammer mass which becomes greater in the bass. It seems moreover that the properties of the individual keys do not vary much in the middle range of the keyboard, except at low velocities.

After eliminating outliers that may be due to irregularities in the electronic equipment, the TCM was averaged across the keyboard. The (inverse power) curve interpolated to these seven averaged data points gave a stable representation of the travel times as a function of final hammer velocity (Fig. 5).[7]

C. Procedure

Note onsets and the hammer velocity information were extracted from the performance data. These data were matched to a symbolic score in which each voice was individually indexed, beginning with 1 as the highest voice[8] (see Figs. 2 and 3). Wrong notes (substitutions) or missing notes (deletions) were marked as such. The rate of not-played or wrongly played notes was very low: for all pianists 0.43% for the Etude (of $n_{\text{total}}=9988$), 0.69% for the Ballade (of $n_{\text{total}}=16\,082$), and 1.75% for the two repeated versions of the Ballade (of $n_{\text{total}}=5764$).[9]

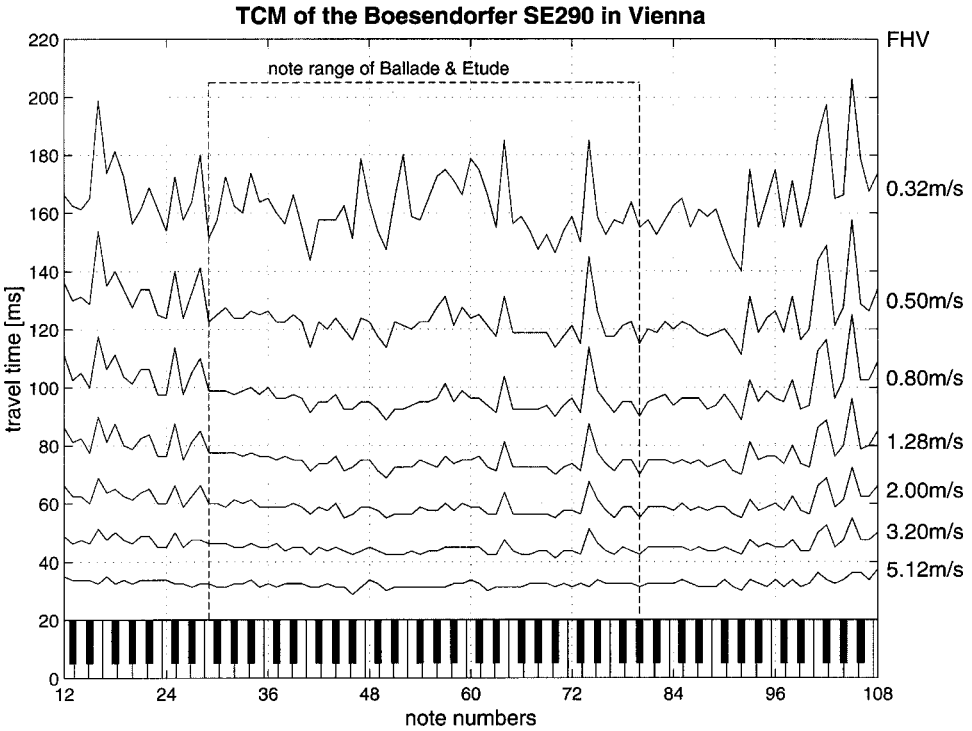


FIG. 4. The travel time correction matrix (TCM) of the Bösendorfer SE290 in Vienna used for the experiments. The seven functions display the measured travel time for seven final hammer velocities. This matrix is derived from a hardware chip (EEPROM) included in the Bösendorfer SE system.

Timing differences and hammer velocity differences between the first voice (=melody) and each other voice were calculated separately for all nominally simultaneous events in the score. All missing or wrong notes, as well as chords marked in the score as arpeggio (Ballade) or as appoggiatura (Etude) were excluded.[10] The finger-key contact times were calculated for each note by subtracting from the hammer-string impact time the corresponding travel time, which was determined by the TCC (see Fig. 5). From this, finger-key asynchronies were calculated, again between voice 1 and all other voices separately for all nominally simultaneous events in the score.

IV. RESULTS

Figure 6 shows the mean velocity profiles (top graphs) as well as the mean asynchrony profiles (bottom graphs) of the 22 performances of the Ballade and the Etude and their overall averages. All pianists played the first voice consistently louder than the other voices. None of the pianists chose another voice to be played as the loudest voice. The velocity levels of the individual voices were fairly constant in the performances of the Ballade, so averaging over all notes in a voice made sense. For the performances of the Etude the dynamic climax of bar 17 caused a strong increase in the velocity values. Therefore, in Fig. 6 the section from bar 14 to 18 was averaged separately and was not included in the overall average. Again, the first voice clearly showed the highest velocity values.

The two bottom graphs in Fig. 6 show the hammer-string and the finger-key asynchrony profiles for the two pieces. The thicker lines with the standard deviation bars represent the average of the mean asynchrony profiles of the 22 performances (thin lines without symbols).

In the hammer-string domain, the melody preceded other voices, as expected, by about 20–30 ms. In the Ballade the asynchrony profiles of the individual performances were very similar to each other, and the melody lead was slightly greater relative to the left hand voices than to the right hand voices. The individual chord profiles for the Etude showed more variability among pianists, especially in the left hand, where the bass voice (7) tended to lead for some pianists (for an example, see the following).

The asynchronies at the finger-key level (Fig. 6, broken lines, average with circles) were consistently smaller than those at hammer-string level. In particular, the melody lead within the right hand is reduced to about zero, whereas the left hand tends to lead the right hand. Two repeated-measure analyses of variance (ANOVA) on the average melody leads for each voice in each performance with type of asynchronies (hammer-string and finger-key) and voice (2 to 5 in the Ballade and 2 to 7 in the Etude) as within-subject factors separately for the two pieces (Etude, Ballade) showed significant main effects of type of melody lead and significant interactions between type and voice.[11]

A real outlier was pianist 3, who played the melody 40–70 ms before the accompaniment, as shown in Fig. 7. This was a deliberate strategy that pianist 3 habitually uses to emphasize melody. In a private communication with pianist 3, he confirmed this habit and called it a ''spleen.'' His finger-key profiles still showed a melody lead of about 20 ms and more. A similar but smaller tendency was shown by two other pianists. This finding suggests that melody lead can be applied deliberately and used as an expressive device—in addition to a dynamic differentiation—to highlight the melody. We argue here that, when melody lead is used as a conscious expressive device, it should be observable at the finger-key level. This strategy seems to be fairly rare.

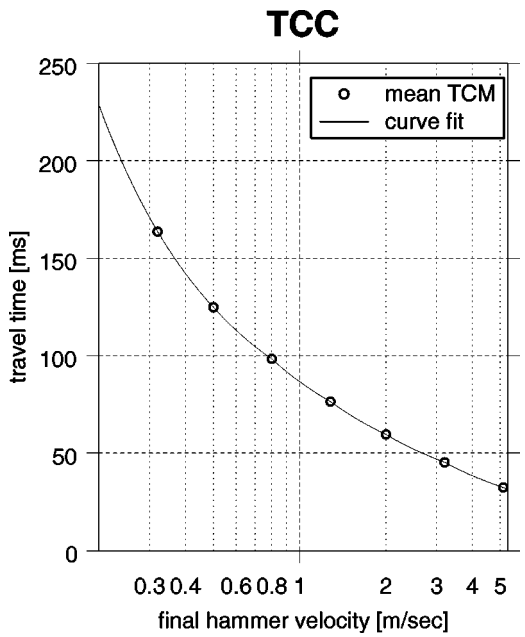


FIG. 5. The timing characteristics of a grand piano action: the hammer travel times as a function of final hammer velocity. This timing correction curve (TCC) was fitted to average data derived from a Bösendorfer SE EPROM chip (see Fig. 4). The *y* axis represents the time interval between finger-key contact times (measured 2–3 mm below the key surface) and the hammer-string contact times.

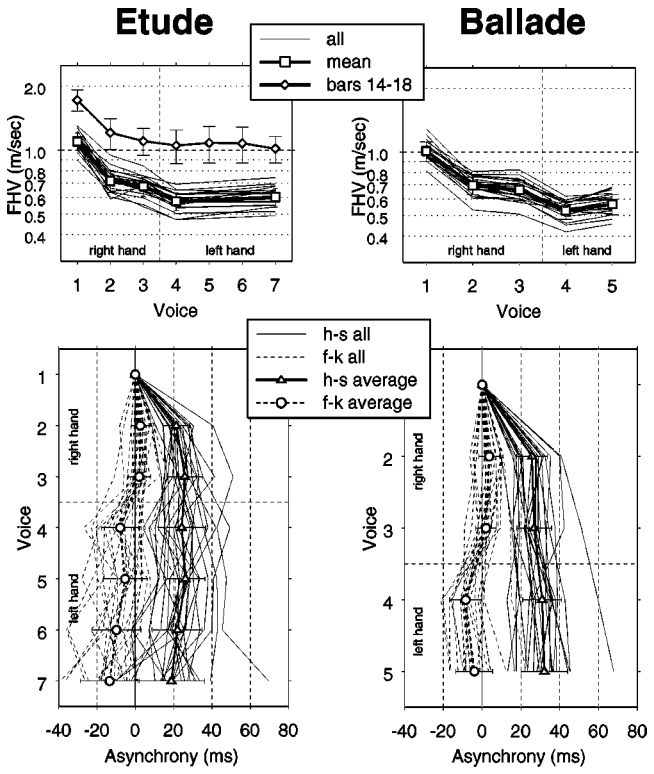


FIG. 6. The individual and mean final hammer velocity (FHV) and asynchrony profiles (with standard deviation bars) of 22 performances for the Etude (left-hand panel) and the Ballade (right). In the top panel, the mean intensity values by pianists and voice are plotted. The thicker lines with squares indicate the average across pianists. In the Etude, bars 14–18 are averaged separately. The profiles at the bottom show the averaged timing delays of voices relative to voice 1. Solid lines represent hammer-string (*hs*) asynchronies, broken lines inferred finger-key (*fk*) asynchronies. The horizontal bars are standard deviations, computed across individual performers.

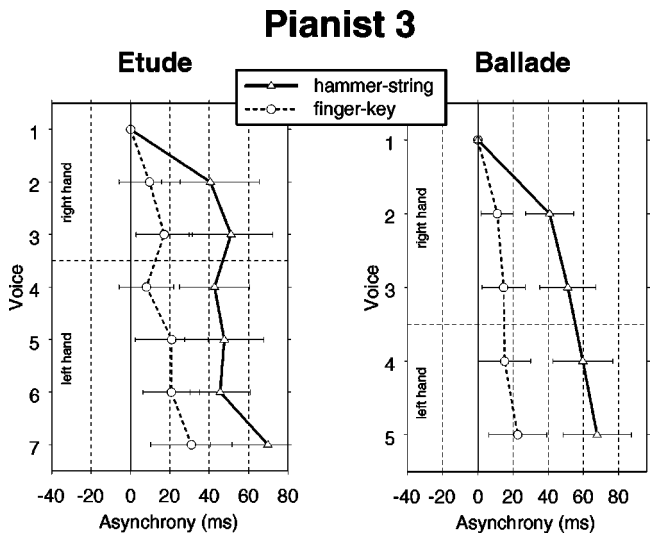


FIG. 7. The asynchrony profiles of pianist 3 (with standard deviation bars) at hammer-string contact (closed lines with triangles) and finger-key contact (broken lines with circles).

The results of the two emphasized versions of the first nine bars of the Ballade are shown in Fig. 8. In the top graphs, the mean intensity values are plotted by voices. In the first voice version (top left graph), the emphasized voice was played louder than in the normal version (mean FHV 1.28 m/s vs 1.01 m/s), while the accompaniment maintained its dynamic range. The melody lead increased up to 40–50 ms (Fig. 8, bottom left graph).

When the third voice was emphasized, that voice was played loudest (at about FHV 1.12 m/s on average), with the melody somewhat attenuated (0.84 m/s) and the other voices as usual (top right graph). The third voice led by about 20 ms compared to the first voice, while the left hand lagged by about 40 ms (Fig. 8). Thus, when pianists are asked to emphasize one voice, they play this voice louder, and the timing difference changes correspondingly.

The first nine bars of the (normal version of the) Ballade were compared with these two special versions (Ballade first voice, Ballade third voice) with regard to hammer velocity and melody lead. A repeated-measure ANOVA on the average hammer velocities of each voice in each performance with instruction (normal, first, third) and voice (1–5) as within-subject factors was conducted. Significant effects on instruction $[F(2,21)=4.98, p<0.05]$, voice $[F(4,84)=466.2, p<0.001]$, and a significant interaction between instruction and voice $[F(8,168)=88.58, p<0.001]$ indicate that pianists changed the dynamic shaping of the individual voices significantly. Another repeated-measure ANOVA was conducted on the melody leads averaged for each voice in each performance, again with instruction (normal, first, and third) and voice (2–5) as within-subjects factors. It showed significant effects of instruction $[F(2,42)=114.41, p<0.001]$ and voice $[F(3,63)=24.12, p<0.001]$, and an interaction between instruction and voice $[F(6,126)=31.29, p<0.001]$.

Relationship between velocity and timing. Generally, it was the case that the larger the dynamic differences, the greater the extent of melody lead. The velocity differences

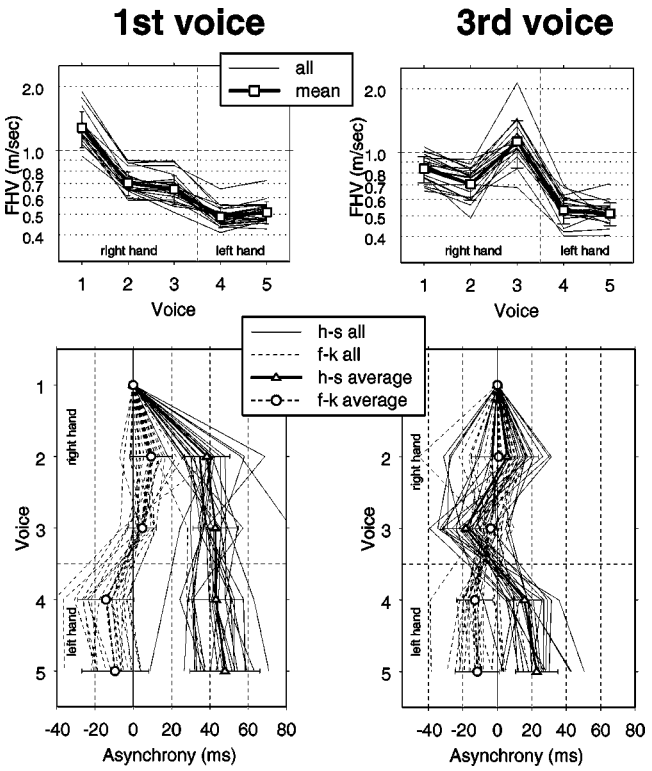


FIG. 8. Average velocity and asynchrony profiles of the 22 individual performances in the Ballade's emphasized melody conditions. On the left-hand side, the first voice was emphasized, on the right, the third voice. The solid lines indicate hammer-string (*hs*) contacts, broken lines finger-key (*fk*) contacts.

between the first voice and the other notes were negatively correlated with the timing differences. The mean correlation coefficients across the 22 pianists are shown in Table I(a), separately for each piece and for right-hand (within-hand) and left-hand (between-hand) comparisons.[12]

The within-hand coefficients were substantially higher than the between-hand coefficients. This suggests a larger independence between the hands than between the fingers of a single hand. Especially for the Etude, almost all of the between-hand coefficients were nonsignificant (with the exception of two pianists). The coefficients for the special versions were slightly higher than those for the ''normal'' versions.

These correlation coefficients assume a linear relationship between melody leads and the velocity differences. However, the expected effect resulting from the piano action timing properties (velocity artifact) does not represent a linear, but rather an inverse power relation (see Fig. 5). To test the presence of this effect in the data, the observed timing differences were correlated with the timing differences predicted by the TCC [Table I(b)]. These correlations were generally higher than the correlations between timing differences and final hammer velocity differences. Eighty-seven out of 88 individual coefficients were highly significant for the right hand. This result shows that the connection of melody lead and intensity variation is even better explained by the velocity artifact than by a linear correlation, as done in literature (Repp, 1996b; Palmer, 1996).

Some of the individual left-hand correlation coefficients between observed and predicted melody lead were nonsignificant in the Etude, but not in the Ballade or in the special versions [Table I(b)]. This suggests not only the general trend of larger between-hand asynchrony variability, but is also due to large bass anticipations—the type 2 asynchronies mentioned previously—played by some pianists, who clearly struck some bass notes earlier. To illustrate these bass anticipations, the beginning of the Etude performed by pianist 5 is shown in Fig. 9. In the bottom graph of Fig. 9, we can observe five bass leads. Two are quite small (bars 6 and 7 about 35–40 ms), two are somewhat larger (bars 2 and 8 about 75 ms) and one is huge (bar 9, 185 ms). All bass leads are even larger in the finger-key domain (see Fig. 9, open symbols). In this example, most of the large bass leads occur at metrically important events. These bass leads are well perceivable and often exceed the range of the melody leads.

V. DISCUSSION

In this study, a large and high quality set of performance data was analyzed. In addition to the measuring of asynchronies at the hammer-string impact level, we estimated the asynchronies at the start of the key acceleration (finger-key

TABLE I. (a) Mean correlation coefficients, with standard deviations (s.d.), between melody lead and final hammer velocity differences across 22 pianists. n_{max} indicates the maximum number of note pairs that went into the computation of each correlation (missing notes reduced this number at some individual performances). $\#r^{**}$ indicates the number of highly significant ($p<0.01$) individual correlations ($\#r^{**}_{max}=22$). (b) The mean correlation coefficients, with standard deviations (s.d.), between observed and predicted melody lead across 22 pianists, and the number of highly significant ($p<0.01$) correlations of the pianists ($\#r^{**}$).

	Etude		Ballade		Ballade first voice		Ballade third voice	
	right hand	left hand	right hand	left hand	right hand	left hand	right hand	left hand
(a)								
n_{max}	126	103	181	269	29	58	29	58
Mean	−0.45	−0.15	−0.42	−0.31	−0.55	−0.29	−0.73	−0.53
s.d.	0.12	0.20	0.13	0.12	0.17	0.22	0.14	0.17
$\#r^{**}$	21	2	22	20	16	12	22	18
(b)								
n_{max}	126	103	181	269	29	58	29	58
Mean	0.66	0.34	0.58	0.50	0.72	0.55	0.79	0.63
s.d.	0.10	0.23	0.13	0.13	0.17	0.22	0.11	0.13
$\#r^{**}$	22	14	22	22	21	21	22	22

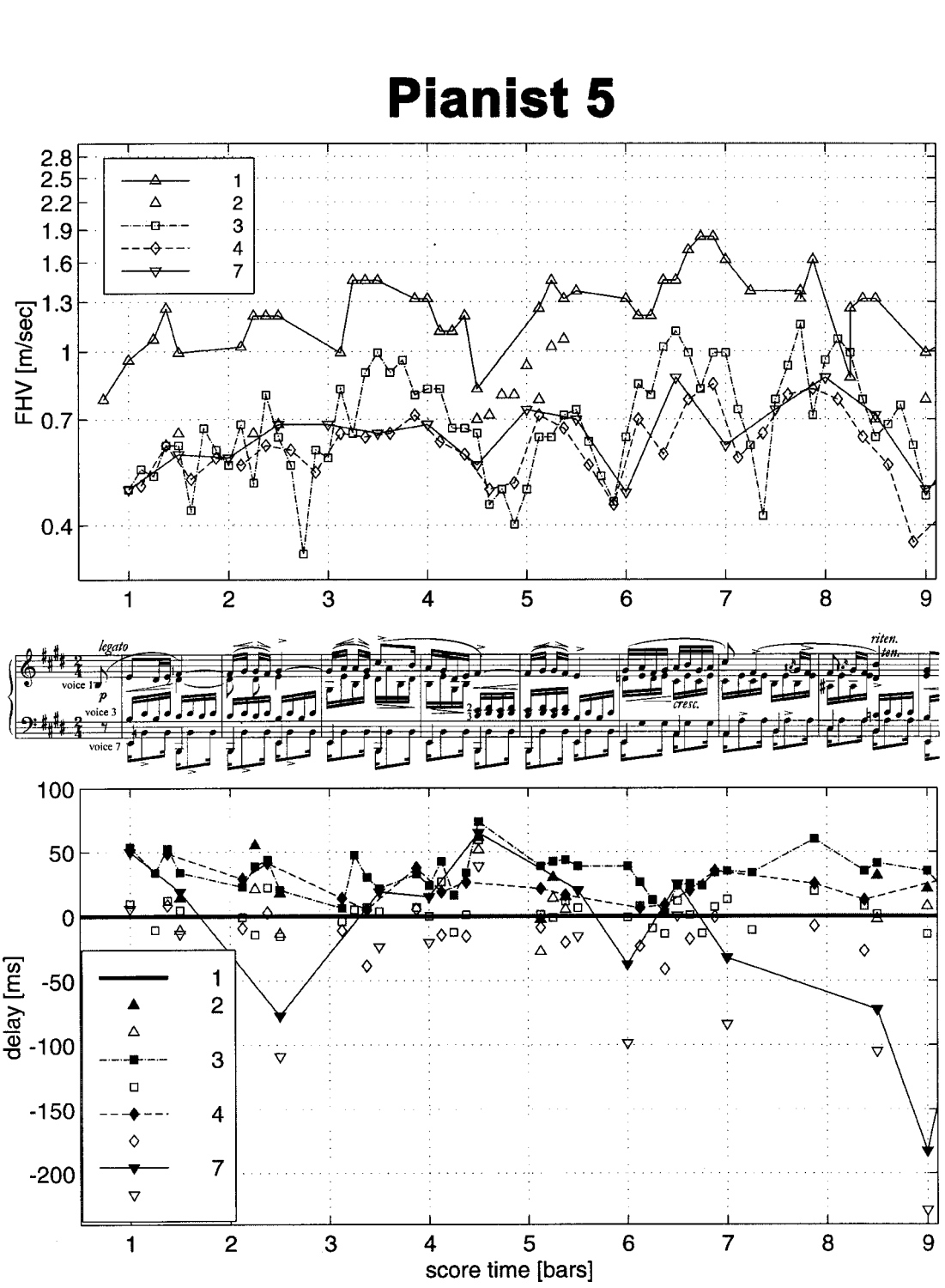


FIG. 9. The dynamic profiles (top) and the note asynchronies (bottom) for the first bars of the Etude op. 10/3 for pianist 5. Top graph: The final hammer velocity (FHV) is plotted against nominal time according to the score. Each voice is plotted separately. The melody is played clearly more loudly than the other voices. The bottom graph shows the time delay of each note relative to onset time of the corresponding melody note (voice 1). The closed symbols represent hammer-string asynchronies, the open symbols the estimated finger-key contact times.

level) through calculation. The hypothesis that melody lead occurs as a consequence of dynamic differentiation was supported in three ways.

(1) The consistently high correlations between hammer-string asynchronies and dynamic differences show the overall connection of melody and velocity difference. The more the melody is separated dynamically from the accompaniment, the more it precedes it. These findings replicate Repp's (1996b) results.

(2) In addition to these findings, the estimated finger-key asynchronies show that, with few exceptions, the melody lead phenomenon disappears at finger-key level. Pianists start to strike the keys almost synchronously, but different velocities cause the hammers to arrive at the strings at different points in time.

(3) With the help of the timing correction curve (TCC), melody lead was predicted in ms. The correlations between this predicted and the observed melody lead were

even higher than the correlations between velocity differences and melody lead. Differences in hammer velocity account for about half of the variance in asynchronies in the data. The other variance could be due to deliberate expression, or motor noise.

The findings of this study are consistent with interpretations of Repp (1996b, velocity artifact explanation) rather than those of Palmer (1989, 1996), who regarded melody lead to be produced independently of other expressive parameters (e.g., dynamics, articulation). Of course it remains true that melody lead can help a listener identify the melody in a multivoiced music environment. Temporally offset elements tend to be perceived as belonging to separate streams (*stream segregation*, Bregman, 1990), and spectral masking effects are diminished by asynchronous onsets (Rasch, 1978, 1979). But in light of the present data, perceptual segregation is not the main reason for melody lead. Primarily, the temporal shift of the melody is a result of the dynamic differentiation of the voices, but both phenomena have similar perceptual results, that is, separating melody from accompaniment.

Nevertheless, pianists clearly played asynchronously in some cases. Some bass notes are played before the melody. Bass lead time deviations are usually around 50 ms and extend up to 180 ms in some cases. These distinct anticipations seem to be produced intentionally, although probably without immediate awareness. This bass lead is well documented in the literature, not only as a habit of an older pianists's generation, but also in some of today's young pianists' performances (Repp, 1996b; Palmer, 1989).

The case of pianist 3 suggests that pianists can enlarge the melody lead deliberately if they wish to do so. In this case, even in the finger-key domain melody lead is observable. However, it does not seem possible for pianists to dynamically differentiate voices in a chord without producing melody lead in the hammer-string domain. At least there is no example in the present data that would prove this.

In the examples of deliberately produced asynchronies (bass lead and enlarged melody lead), the extent of the asynchrony usually exceeded 30 ms. Such asynchronies may be regarded as a deliberate expressive device under direct control of the pianists. According to the pianists in this study, they are produced in a somewhat subconscious way (private communication with the pianists), but pianists report a general awareness of the use of these asynchronies and that they could suppress them if they wanted to. However, the use of the ''normal'' melody lead that was produced by *all* pianists was unconscious. Pianists reported that they emphasize one voice by playing it dynamically louder, but not earlier (the same was reported by Palmer, 1989, p. 335).

The asynchronies in the finger-key domain were computed by using a timing correction curve which provides the time interval from key-press to hammer-string impact as a function of final hammer velocity. The key shutter reacts when the key is depressed about 2 to 3 mm (the ''touch depth'' of the key is usually about 9.5 mm, Askenfelt and Jansson, 1991, p. 2383, varying slightly across pianos). Thus, to be precise, the finger-key domain represents points in time when keys are depressed by 2 to 3 mm. However, almost nothing is known about how keys are accelerated and released in reality. In very precise acceleration measurements by Van den Berghe (Van den Berghe *et al.*, 1995, p. 17) it can be seen that sometimes keys are not released entirely, especially in repetitions. The modern piano action has the double repetition feature that allows a second strike without necessarily releasing the key entirely. If the system measured onsets close to the zero position, some onsets would not be detected as such. Nevertheless, the 2 to 3 mm below zero level still gives a good impression about the asynchronies at the start of a key acceleration. For more accurate statements about played and perceived onset asynchronies in piano performance, evaluation of acceleration measurements at different points in the piano action would be necessary.

This paper was concerned with the particular properties of the piano. Other keyboard actions (harpsichord, organ) may have similar timing properties as far as the key itself is concerned (a key that is depressed faster reaches the keybed earlier than a slower one), but their actions respond differently due to their different way of producing sound: the harpsichord plucks the strings, and on the organ a pipe valve is opened or closed. Additionally, they do not allow continuous dynamic differentiation like a piano does, and therefore performers may choose timing as a means to separate voices. However, we note a difference in the played repertoire: homophonic textures, like the Chopin excerpts used in this study, are seldom seen in the harpsichord or organ repertoires.

According to Vladimir Horowitz, when accenting a tone within a chord one should ''raise the whole arm with as little muscular effort as possible, until the fingers are between three and five inches above the key. During the up and down movements of the arm, prepare the fingers by placing them in position for the depression of the next group of notes and by holding the finger which is to play the melody-note a trifle lower and firmer than the other fingers which are to depress the remaining keys of the chord'' (Eisenberg, 1928).[13] This would suggest that an asynchrony at the key is intended, but Horowitz goes on: ''The reason for holding the finger a trifle lower is only psychological in effect; in actual practice, it isn't altogether necessary. Experience shows that in the beginning it is almost impossible to get a student to hold one finger more firmly than the others unless he is also permitted to hold it in a somewhat different position from the others. Holding it a little lower does not change the quality or quantity of tone produced and does not affect the playing in any way but it does put the student's mind at greater ease'' (Eisenberg, 1928). As the pianists in the present study, Horowitz is aiming at intensity differences here, but not at differences in timing: ''The finger which is held a trifle lower and much firmer naturally strikes the key a much firmer blow than do the more relaxed fingers which do not overcome the resistance of the key as easily as does the more firmly held finger. The tone produced by the key so depressed is therefore stronger than the others'' (Eisenberg, 1928). The quote suggests that Horowitz was unaware of the consequences of his recommendation for onset synchrony or that he considered onset asynchrony unimportant.

ACKNOWLEDGMENTS

This research was supported by the START program of the Austrian Federal Ministry of Education, Science and Culture (Grant No. Y99-INF to Gerhard Widmer). I thank Wayne Stahnke, who generously shared his insight into the functionality of the Bösendorfer SE system. Thanks are due to the Bösendorfer Company in Vienna, which provided the SE290 *Imperial* grand piano for experimental use, and especially to Fritz Lachnit, who carefully maintained the grand piano; to the Department for Microelectronics at the Vienna University of Technology, and to Simon Dixon for reading and converting the data of the EEPROM hardware chip. Special thanks to Nigel Nettheim for posting the Horowitz interview on the web. Emilios Cambouropoulos, Simon Dixon, Caroline Palmer, Richard Parncutt, Bruno Repp, Oliver Vitouch, and Gerhard Widmer provided helpful comments on earlier drafts of this paper. Finally, I am indebted to all the participating pianists who provided the data for this study.

[1]Askenfelt and Jansson (1990) used a Hamburg Steinway & Sons grand piano, model B (7 ft, 211 cm) for their measurements.

[2]The ''prelay'' function compensates for the different travel times of the action at different hammer velocities. In order to prevent timing distortions in reproduction, the MIDI input is delayed by 500 ms. The solenoids (the linear motors moving the keys) are then activated earlier for softer notes than for louder notes, according to a pre-programmed function.

[3]Palmer's (1996, pp. 27 and 29) assumption that the Bösendorfer SE290 grand piano records key-bed impact times as note onsets is contradicted by information given by W. Stahnke (private communication), who developed the SE system in the 1980s (Moog and Rhea, 1990, p. 53),

[4]All recordings can be downloaded in MP3 format at *http://www.ai.univie.ac.at/~wernerg*

[5]''SE'' stands for Stahnke Electronics, 290 indicates the length of the piano in cm.

[6]This time difference is by definition called the *inverse hammer velocity* (IHV), and is measured by a counter operating at 25.6 kHz. Therefore, the relationship between the *IHV* and the *final hammer velocity*, expressed in meters per second, is determined via the function IHV=128/FHV. The IHV values can range from 0 (128 m/s) to 1023 (0.125 m/s). The performing pianists typically produce IHV values between 30 (4.26 m/s) and about 600 (0.21 m/s), where at 600 the hammers will not reach the strings anymore (silent note).

[7]In an earlier analysis (Goebl, 2000), the author made use of average TCM data provided by Stahnke (private communication), derived from his Bösendorfer SE290. It was preferred in the present study to use the mean TCM from the Vienna SE grand. However, the differences between these two inverse power curves were very small and did not affect the results much.

[8]The lowest voice played by the right hand was called 3. If there were three simultaneous notes in the right hand, the middle one was labeled 2. The highest voice played by the left hand was indexed 4, the base line 5 in the Ballade, and 7 in the Etude. Voices 5 and 6 in the Etude occurred only in measures 16 and 17. In the Ballade, there was only one chord (bar 19) with three simultaneous notes in the left hand. Here, the two higher notes were labeled 4, the bass 5.

[9]Additional notes (insertions) that were so soft (or silent) that they did not disturb the performance and were apparently not perceived as mistakes, were not counted as errors. In the Etude we observed 181 such notes over the 22 performances (+1.8%), in the Ballade 189 (+1.17%). Similar observations were made also by Repp (1996a).

[10]The excluded events for the Etude were ([bar number]·[relative position in the bar]): 7.75, 8.25, and 21.0, for the Ballade: 18.5, 20.5, 40.0, and 45.0.

[11]The repeated-measure ANOVA for the Ballade: significant effect of type $[F(1,21)=718.2, p<0.001]$, no significant effect of voice $[F(3,63)=1.2, p>0.05]$, and a significant interaction between type and voice $[F(3,63)=112.3, p<0.001]$; for the Etude: significant effects of type $[F(1,21)=603.9, p<0.001]$, and voice $[F(5,105)=5.59, p<0.002]$, and an interaction between type and voice $[F(5,105)=34.83, p<0.001]$.

[12]The negative sign of the correlation coefficients stems from the way of calculating timing and velocity differences and has no relevance for data interpretation: from the onset time of each accompanying note (t_n) the onset time of the corresponding melody note (t_1) is subtracted (t_n-t_1), so the melody lead is positive. Similarly, the velocity differences are calculated as: v_n-v_1, which results in negative values. Therefore, the correlation coefficients between melody leads and velocity differences are negative, whereas the coefficients between observed and predicted melody leads are positive.

[13]This article may be found at *http://users.bigpond.net.au/nettheim/horo28.htm*

Alcedo, C., and Schäfer, S. (**1992**). ''Computer-controlled player pianos,'' Comput. Music J. **16**, 104–111.

Askenfeld, A. (Ed). (**1990**). *Five lectures on the acoustics of the piano*. Stockholm: Publications issued by the Royal Swedish Academy of Music, Vol. 64.

Askenfeld, A. (**1999**). Private communication.

Askenfeld, A., and Jansson, E. V. (**1990**). ''From touch to string vibrations. I. Timing in grand piano action,'' J. Acoust. Soc. Am. **88**, 52–63.

Askenfelt, A., and Jansson, E. V. (**1991**). ''From touch to string vibrations. II. The motion of the key and hammer,'' J. Acoust. Soc. Am. **90**, 2383–2393.

Askenfelt, A., and Jansson, E. V. (**1992**). ''From touch to string vibrations. III. String motion and spectra,'' J. Acoust. Soc. Am. **93**, 2181–2196.

Bregman, A. S. (**1990**). *Auditory Scene Analysis. The Perceptual Organization of Sound* (MIT, Cambridge).

Conklin, H. A. (**1996**). ''Design and tone in the mechanoacoustic piano. I. Piano hammers and tonal effects,'' J. Acoust. Soc. Am. **99**, 3286–3296.

Eisenberg, J. (**1928**). ''Horowitz explains accenting a melody while playing chords,'' The Musician **June 1928**, 11.

Goebl, W. (**2000**). ''Skilled piano performance: Melody lead caused by dynamic differentiation,'' in Proceedings of the Sixth International Conference on Music Perception and Cognition,'' edited by C. Woods, G. Luck, R. Brochard, F. Seddon, and J. A. Sloboda, Keele, Keele University Department of Psychology, pp. 1165–1176.

Hartmann, A. (**1932**). ''Untersuchungen über das metrische Verhalten in musikalischen Interpretationsvarianten,'' Arch. Gesamte Psychologie **84**, 103–192.

Moog, R. A., and Rhea, T. L. (**1990**). ''Evolution of the keyboard interface: The Bösendorfer 290 SE recording piano and the Moog multiply-touch-sensitive keyboards,'' Comput. Music J. **14**, 52–60.

Palmer, C. (**1989**). ''Mapping musical thought to musical performance,'' J. Exp. Psychol. **15**, 331–346.

Palmer, C. (**1996**). ''On the assignment of structure in music performance,'' Music Perception **14**, 23–56.

Palmer, C., and Brown, J. C. (**1991**). ''Investigations in the amplitude of sounded piano tones,'' J. Acoust. Soc. Am. **90**, 60–66.

Rasch, R. A. (**1978**). ''The perception of simultaneous notes such as in polyphonic music,'' Acustica **40**, 21–33.

Rasch, R. A. (**1979**). ''Synchronization in performed ensemble music,'' Acustica **43**, 121–131.

Repp, B. H. (**1993**). ''Some empirical observations on sound level properties of recorded piano tones,'' J. Acoust. Soc. Am. **93**, 1136–1144.

Repp, B. H. (**1996a**). ''The art of inaccuracy: Why pianists' errors are difficult to hear,'' Music Perception **14**, 161–184.

Repp, B. H. (**1996b**). ''Patterns of note onset asynchronies in expressive piano performance,'' J. Acoust. Soc. Am. **100**, 3917–3932.

Stahnke, W. Private communication.

Van den Berghe, G., De Moor, B., and Minten, W. (**1995**). ''Modeling a grand piano key action,'' Comput. Music J. **19**, 15–22.

Vernon, L. N. (**1937**). ''Synchronization of chords in artistic piano music,'' in *Objective Analysis of Musical Performance*, edited by C. E. Seashore (University Press), Vol. IV, pp. 306–345.

Relationships among calcaneal backscatter, attenuation, sound speed, hip bone mineral density, and age in normal adult women

Keith A. Wear[a)]
U.S. Food and Drug Administration, Center for Devices and Radiological Health, HFZ-142, 12720 Twinbrook Parkway, Rockville, Maryland 20852

David W. Armstrong III
National Naval Medical Center, Endocrinology Clinic, 8901 Wisconsin Avenue, Bethesda, Maryland 20889-5600

(Received 16 February 2001; revised 27 March 2001; accepted 11 April 2001)

The present study was undertaken in order to investigate the use of calcaneal ultrasonic backscatter for the application of diagnosis of osteoporosis. Broadband ultrasonic attenuation (BUA), speed of sound (SOS), the average backscatter coefficient (ABC), and the hip bone mineral density (BMD) were measured in calcanea in 47 women (average age: 58 years, standard deviation: 13 years). All three ultrasound variables had comparable correlations with hip BMD (around 0.5). As reported previously by others, BUA and SOS were rather highly correlated with each other. The logarithm of the ABC was only moderately correlated with the other two. The three ultrasound parameters exhibited similar moderate negative correlations with age. These results taken collectively suggest that the ABC may carry important diagnostic information independent of that contained in BUA and SOS and, therefore, may be useful as an adjunct measurement in the diagnosis of osteoporosis. [DOI: 10.1121/1.1378343]

PACS numbers: 43.80.Cs, 43.80.Qf [FD]

I. INTRODUCTION

Ultrasound-based methods for bone assessment are less expensive, faster, simpler, and more portable than their x-ray counterparts: dual-energy x-ray absorptiometry (DEXA) and quantitative computed tomography (QCT). In addition, they produce no ionizing radiation. The diagnostic capability of ultrasound to predict fracture risk has been documented in many studies. Calcaneal broadband ultrasonic attenuation[1–4] (BUA) and speed of sound[2,3,5–11] (SOS) correlate highly with the calcaneal bone mineral density (BMD), which is in turn an indicator of osteoporotic fracture risk in the hip.[12] Linear combinations of BUA and SOS are thought to have greater diagnostic capability than either parameter by itself and have been demonstrated to be predictive of hip and other fractures in women in prospective[13,14] and retrospective[15–18] studies.

In general, ultrasonic backscatter is known to provide information regarding size, shape, number density, and elastic properties of scatterers.[19–23] In trabecular bone applications, trabeculae are likely candidates for scattering sites due to the high contrast in acoustic properties between mineralized trabeculae and marrow.[24–27] The diminished number and thicknesses of trabeculae within bone that are associated with increased fracture risk would be expected to reduce backscatter. Several preliminary studies have established diagnostic promise and/or clinical feasibility of this measurement.[28–34] A recent study has suggested that backscatter increases significantly upon decollagenization (with 7% sodium hypochlorite) and decreases or remains the same upon demineralization (with ethylenediaminetetraacetic acid, disodium salt dihydrate, or EDTA).[35]

The present study was undertaken to further investigate the use of ultrasonic backscatter as a diagnostic measurement. This study is an extension to a larger population of subjects of an earlier preliminary investigation of ten subjects. The initial study was performed using a clinical imaging system designed for abdominal applications. A center frequency of 2.25 MHz was used. In the present work a lower-frequency (1 MHz, more suitable for highly attenuating bone) system, but without imaging capability, was employed and applied to a much larger set of subjects.

II. METHODS

Forty-seven normal adult women (3 African American, 1 African European, 1 Asian, 1 Hispanic, and 41 non-Hispanic Caucasian; average age 58 years with a standard deviation of 13 years) were studied at the National Naval Medical Center (NNMC) in Bethesda, MD. These nonpregnant women were free of known disorders of skeletal metabolism, with no known history of nontraumatic bone fractures, and free of disorders whose treatment or medication promotes loss of bone such as diabetes, emphysema, and asthma. In addition, women with a history of bilateral hip replacement, Paget's disease of the bone, renal failure, hyperthyroidism, or treated hypothyroidism were excluded. All participants provided informed consent (approved by both the FDA and NNMC).

Ultrasound measurements were performed on each subject's left calcaneus. While sitting upright in a chair, the subject rested her left foot on her right knee, as shown in Fig.

[a)]Electronic mail: kaw@cdrh.fda.gov

FIG. 1. Data acquisition. The subject rested her left heel upon her right knee. A standoff pad was used so that the region of interest coincided with the focal zone of the transducer. Coupling gel was used at the transducer/standoff-pad and standoff-pad/skin interfaces.

1. A standoff pad (Cone Instruments, Solon, OH) was placed on the foot in order that data could be acquired in the focal zone of the transducer. The standoff pad also created a situation in which multiple reverberations between the transducer and the calcaneal surface were not received until well after the backscatter from the calcaneus and, therefore, could be excluded from the analysis.

Backscatter measurements were performed using a Panametrics (Waltham, MA) 5800 pulser/receiver and Panametrics V302 1-in.-diam, focused (focal length=2 in.), broadband transducer with center frequency of 1 MHz. Received signals were digitized (8 bit, 10 MHz) using a LeCroy (Chestnut Ridge, NY) 9310C Dual 400 MHz oscilloscope and stored on a computer (via GPIB) for off-line analysis.

Backscatter coefficients were measured using a reference phantom method.[36] Good agreement between experimental measurements using this method and theoretical predictions based on Faran's theory of scattering[19] for ultrasonic backscatter coefficients from phantoms consisting of glass spheres embedded in gelatin has previously been reported by this laboratory.[37] Over a band of frequencies corresponding to the system usable bandwidth (700 kHz–1.5 MHz), backscatter coefficient versus frequency data were least-squares fit to a power-law relationship. The average backscatter coefficient (ABC) was computed from the midband (1 MHz) value of the fit. In order to compute backscatter coefficients it was necessary to compensate backscattered spectra for the effects of attenuation. For this compensation, the attenuation coefficient was taken to be linear with frequency and computed as the ratio of BUA (dB/MHz) and an assumed constant value for calcaneal thickness for all subjects (3 cm).

TABLE I. Means and standard deviations of measurements. Acronyms are BUA (broadband ultrasonic attenuation), SOS (speed of sound), ABC (average backscatter coefficient), and BMD (bone mineral density).

Parameter	Mean±standard deviation
BUA (dB/MHz):	87.7±13.6
SOS (m/s):	1539.1±34.2
Stiffness:	89±18
Log (ABC):	−0.3±0.24
BMD (g/cm^2):	0.873±0.120
Age (years):	58±12.8
Height (in.):	64.8±2.8
Weight (lb):	148.6±27.3

A Lunar Achilles Plus bone sonometer was used for measurements of broadband ultrasonic attenuation and speed of sound. The Achilles estimate of SOS was based on transit time of the *leading edge* rather than the *center* of the pulse propagating through the calcaneus and, therefore, provided an estimate closely related to *signal velocity* rather than *group velocity*.[38,39] This device also provided measurements of a parameter commonly (but not rigorously) referred to as ''stiffness,'' which is a linear combination of BUA and SOS and is thought to be a more robust indicator of bone status than either parameter by itself:

$$\text{Stiffness} = 0.68*\text{BUA(dB/MHz)} + 0.28*\text{SOS(m/s)} - 420.$$

Based on a wealth of empirical data showing that the BMD is positively correlated with both BUA and SOS, the first two coefficients in this equation were obtained by doing regression analysis to obtain those coefficients which make the contributions of the first two terms roughly equal in magnitude (on the average) over the range of clinical values. The constant term was designed to make the average value of stiffness for young normal adult women to be equal to 100. (This terminology should not be confused with the more standard and formal use of the term stiffness, which corresponds to the coefficient of proportionality between displacement and restorative force,[40] as in a spring, and may be expressed in units of N/m.) The Achilles has been reported to yield estimates of BUA which are somewhat higher than those provided with comparable clinical calcaneal ultrasound

TABLE II. Correlation matrix for measurements and age. Acronyms are BUA (broadband ultrasonic attenuation), SOS (speed of sound), ABC (average backscatter coefficient), BMD (bone mineral density), HGT (height), and WGT (weight).

Parameter	BUA	SOS	Stiffness	ABC	BMD	Age	HGT	WGT
BUA	1.00	0.81	0.95	0.43	0.56	−0.44	0.17	0.19
SOS	0.81	1.00	0.96	0.46	0.51	−0.46	−0.06	−0.04
Stiffness	0.95	0.96	1.00	0.47	0.56	−0.47	0.05	0.08
ABC	0.43	0.46	0.47	1.00	0.50	−0.42	0.17	0.08
BMD	0.56	0.51	0.56	0.50	1.00	−0.31	0.06	0.30
Age	−0.44	−0.46	−0.47	−0.42	−0.31	1.00	−0.29	0.03
Height	0.17	−0.06	0.05	0.17	0.06	−0.29	1.00	0.35
Weight	0.19	−0.04	0.08	0.08	0.30	0.03	0.35	1.00

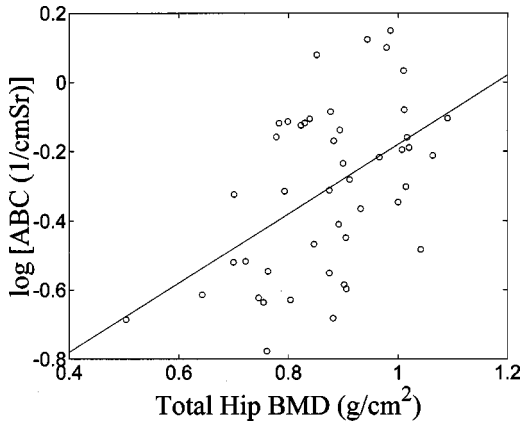


FIG. 2. Logarithm of average backscatter coefficient (ABC) versus total hip bone mineral density (BMD).

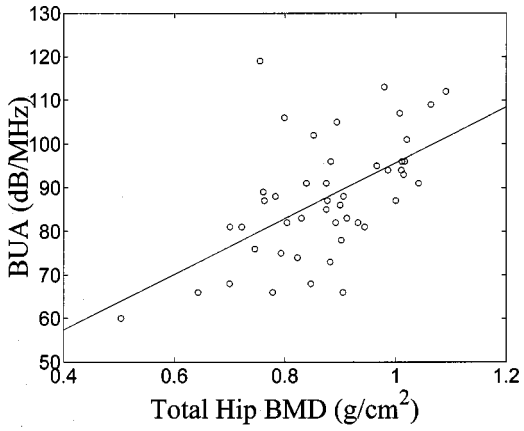


FIG. 3. Broadband ultrasonic attenuation (BUA) versus total hip bone mineral density (BMD).

devices.[41] In order to calibrate the Achilles, separate measurements of BUA were performed on two commercial phantoms using the Achilles and a laboratory apparatus previously described.[42] The phantoms were manufactured by CIRS, Inc. (Norfolk, VA). One was designed to mimic normal bone (normalized BUA = 20.4 dB/cm MHz, SOS = 1562 m/s). The other was designed to mimic osteoporotic bone (normalized BUA = 14.2 dB/cm MHz, SOS = 1503 m/s). The normalized BUA values specified by the manufacturer were obtained over the frequency range from 250 to 550 kHz. Six repeated measurements were acquired for each phantom for each measurement system. In the case of the laboratory system, each measurement was based on eight rf lines. The average offset of the Achilles obtained from this comparison was subtracted from each of the clinical measurements of BUA.

A Hologic QDR-2000 DEXA scanner was used to measure the BMD of the total hip, femoral neck, trochanter, and Ward's triangle. Only total hip BMD measurements are reported here.

III. RESULTS

The average values and standard deviations for the measurements performed on the 47 subjects are given in Table I. The correlations between pairs of measurements are given in Table II. As reported by others, BUA and SOS were rather highly correlated with each other. The logarithm of the average backscatter coefficient was only moderately correlated with the other two. All three ultrasound variables had comparable correlations with hip BMD (around 0.5). The three ultrasound parameters exhibited similar moderate negative correlations with age. These results taken together suggest that the ABC may carry important diagnostic information independent of that contained in BUA and SOS. Scatter plots of the three ultrasound parameters against hip BMD and age are given in Figs. 2–7.

The sharpest decline in BMD occurs shortly following menopause. The age-related drop in BUA between ranging in age from 45 to 54 and those ranging from 55 to 64 was 93.3 dB/MHz (average of 12 subjects) to 84.6 dB/MHz (average of 11 subjects). These values are quite similar to those reported by Cheng *et al.*,[43] namely, 88.5 dB/MHz (average of 20 subjects) to 78.0 dB/MHz (average of 20 subjects) in the same age brackets. Similarly, a drop in SOS from 1561 to 1523 m/s was observed in the present study (Cheng *et al.* did not measure SOS).

A linear regression between BUA and age was applied to the subset of subjects over 55, yielding a postmenopausal average annual rate of decline of 0.49 dB/MHz. This is similar to the value of 0.35 dB/MHz per year reported by Frost and co-workers.[44] The corresponding values for SOS are 0.30 m/s per year (present study) and 0.56 m/s per year.[44]

IV. DISCUSSION

This study indicates that calcaneal ultrasonic backscatter exhibits correlations with hip BMD and age that are comparable to those exhibited by BUA and SOS. The correlation between backscatter and the other two ultrasonic variables is only about 0.5, suggesting that backscatter may reveal substantial information not contained in the other two and, therefore, may have some potential value as a substitute or an adjunct for BUA and SOS for assessment of bone status. This study represents the first clinical evaluation of a contact-based (as opposed to water-bath-based[30]) method for measurement of calcaneal backscatter near 1 MHz. The correlation between backscatter and hip BMD in these and the contact-based and water-bath-based studies are essentially identical (0.50).[30]

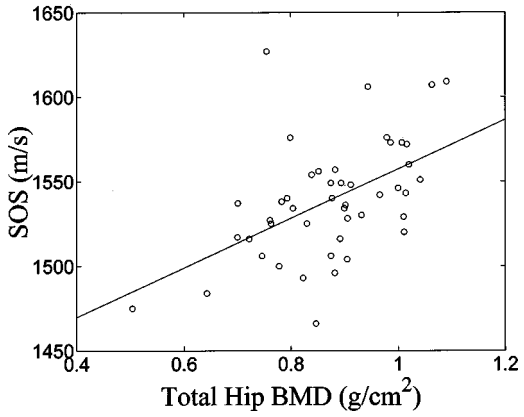


FIG. 4. Speed of sound (SOS) versus total hip bone mineral density (BMD).

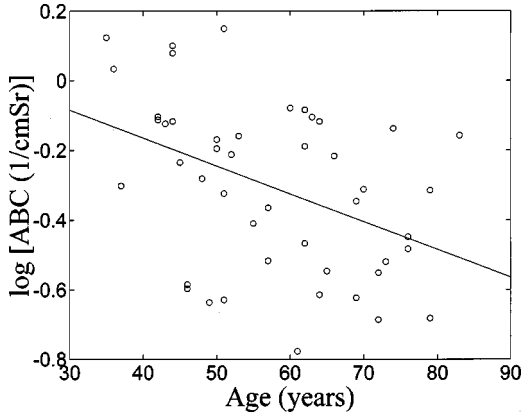


FIG. 5. Logarithm of average backscatter coefficient (ABC) versus age.

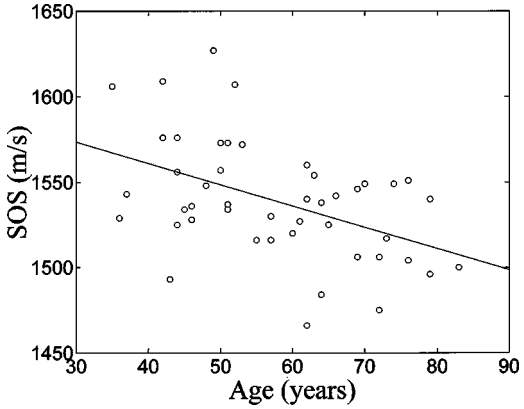


FIG. 7. Speed of sound (SOS) versus age.

The correlation between backscatter and BMD observed here at 1 MHz ($r=0.50$; 95% CI: 0.24–0.69) is smaller than previously reported at 2.25 MHz ($r=0.87$).[29] In addition to the difference in center frequencies, there are numerous differences between these two studies that may account for this. Most importantly, the earlier study correlated calcaneal backscatter with *calcaneal* BMD (*volumetric* BMD, in g/cm^3, measured with QCT), while the current investigation correlated calcaneal backscatter with *hip* BMD (*areal* BMD, in g/cm^2, measured with DEXA). It is not surprising that choosing two different sites rather than using the same site resulted in a decrease in the correlation. In addition, the orientation of the transducer relative to the calcaneus was different in the two studies. In the earlier study, the transducer was placed on the inferior surface of the foot, below the calcaneus, while in the present study, data were acquired in the mediolateral orientation. In addition, unlike the earlier study, the current study incorporated compensation for calcaneal attenuation. Finally, applying the method of Hoel,[45] the 95% CI for the correlation coefficient in the earlier study may be computed to be 0.54–0.97, indicating some overlap in the 95% CI for the two studies.

Preliminary *in vitro* studies[28,29,32] have demonstrated that ultrasonic backscatter increases with BMD in human calcaneus. Variations in BMD accounted for 66%–68% (correlation coefficient near 0.8) of variations in the backscatter coefficient, suggesting that backscatter measurements give some indication of BMD, but also may convey some structural information not already contained in BMD measurements. Another study in bovine tibia trabecular bone[34] reported ''apparent'' (not compensated for attenuation) integrated (averaged over a range of frequencies) backscatter to actually decrease with density in the mediolateral (ML) and superoinferior orientations and to show no change with density in the anteroposterior (AP) orientation. It is plausible that compensation for attenuation (which demonstrated a significant *increase* with BMD in the AP and ML directions) in this study[34] could have yielded backscatter results more similar to the two above-mentioned *in vitro* studies on human calcaneus. Differences between human calcaneus and bovine tibia trabecular bone[46] as well as different frequency ranges of analysis [700 kHz–1.5 MHz in the present study, 200–600 kHz,[28] and 1–3 MHz (Ref. 34)] may also partially account for some disparity in the results.

Fracture risk, not BMD, is the fundamental clinical end point in the diagnosis of osteoporosis. Ideally, the utility of any diagnostic measurement for the prediction of fracture risk should be compared in populations with and without fractures, either prospectively or retrospectively. However, such a study can be somewhat more difficult as individuals either with fractures or with high propensity for fracture can be harder to find and recruit than normals. The current study demonstrates feasibility and some clinical promise for the contact method.

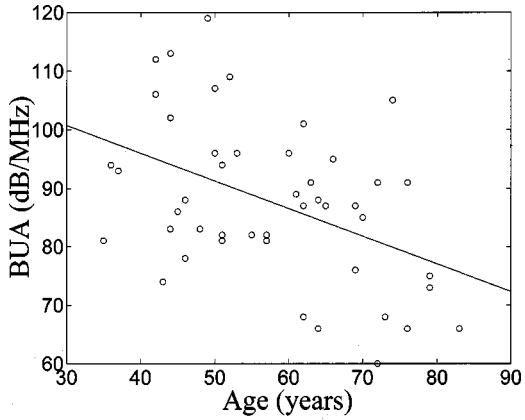


FIG. 6. Broadband ultrasonic attenuation (BUA) versus age.

ACKNOWLEDGMENTS

The guidance of Dr. Mona S. Calvo, U.S. Food and Drug Administration, is greatly appreciated. The authors are grateful for funding provided by the U.S. Food and Drug Administration Office of Women's Health. Gratitude is also expressed to Dr. A. J. Drake III, MC, U.S.N. Endocrinology Clinic, National Naval Medical Center, for serving as a sponsor for this research and authorizing the use of the Hologic DEXA scanner located in the Endocrinology Clinic at the National Naval Medical Center, Bethesda, MD. The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the Food and Drug Administration. The views expressed in this article are those of the authors and do not

reflect the official policy or position of the Department of the Navy, Department of Defense, or the U.S. Government. The Chief, Navy Bureau of Medicine and Surgery, Washington, DC, Clinical Investigation Program was a sponsor of this study, B99-044. This research was supported by the Office of Naval Research and the National Naval Medical Center.

[1] C. M. Langton, S. B. Palmer, and R. W. Porter, ''The measurement of broadband ultrasonic attenuation in cancellous bone,'' Eng. Med. **13**, 89–91 (1984).
[2] P. Rossman, J. Zagzebski, C. Mesina, J. Sorenson, and R. Mazess, ''Comparison of speed of sound and ultrasound attenuation in the os calcis to bone density of the radius, femur, and lumbar spine,'' Clin. Phys. Physiol. Meas. **10**, 353–360 (1989).
[3] M. B. Tavakoli and J. A. Evans, ''Dependence of the velocity and attenuation of ultrasound in bone on the mineral content,'' Phys. Med. Biol. **36**, 1529–1537 (1991).
[4] C. M. Langton, C. F. Njeh, R. Hodgskinson, and J. D. Currey, ''Prediction of mechanical properties of the human calcaneus by broadband ultrasonic attenuation,'' Bone (N.Y.) **18**, 495–503 (1996).
[5] J. A. Zagzebski, P. J. Rossman, C. Mesina, R. B. Mazess, and E. L. Madsen, ''Ultrasound transmission measurements through the os calcis,'' Calcif. Tissue Int. **49**, 107–111 (1991).
[6] C. F. Njeh, R. Hodgskinson, J. D. Currey, and C. M. Langton, ''Orthogonal relationships between ultrasonic velocity and material properties of bovine cancellous bone,'' Meas. Eng. Phys. **18**, 373–381 (1996).
[7] P. Laugier, P. Droin, A. M. Laval-Jeantet, and G. Berger, ''*In vitro* assessment of the relationship between acoustic properties and bone mass density of the calcaneus by comparison of ultrasound parametric imaging and quantitative computed tomography,'' Bone (N.Y.) **20**, 157–165 (1997).
[8] M. L. Bouxsein and S. E. Radloff, ''Quantitative ultrasound of the calcaneus reflects the mechanical properties of calcaneal trabecular bone,'' J. Bone Miner. Res. **12**, 839–846 (1997).
[9] P. H. F. Nicholson, R. Muller, G. Lowet, X. G. Cheng, T. Hildebrand, P. Ruegsegger, G. Van Der Perre, J. Dequeker, and S. Boonen, ''Do quantitative ultrasound measurements reflect structure independently of density in human vertebral cancellous bone?'' Bone (N.Y.) **23**, 425–431 (1998).
[10] D. Hans, C. Wu, C. F. Njeh, S. Zhao, P. Augat, D. Newitt, T. Link, Y. Lu, S. Majumdar, and H. K. Genant, ''Ultasound velocity of cancellous cubes reflects mainly bone density and elasticity,'' Calcif. Tissue Int. **64**, 18–23 (1999).
[11] H. Trebacz and A. Natali, ''Ultrasound velocity and attenuation in cancellous bone samples from lumbar vertebra and calcaneus,'' Osteoporosis Int. **9**, 99–105 (1999).
[12] S. R. Cummings *et al.*, ''The study of osteoporotic fractures research group: Bone density at various sites for the prediction of hip fractures,'' Lancet **341**, 72–75 (1993).
[13] D. Hans, P. Dargent-Molina, A. M. Schott, J. L. Sebert, C. Cormier, P. O. Kotzki, P. D. Delmas, J. M. Pouilles, G. Breart, and P. J. Meunier, ''Ultrasonographic heel measurements to predict hip fracture in elderly women: The EPIDOS prospective study,'' Lancet **348**, 511–514 (1996).
[14] D. C. Bauer, C. C. Gluer, J. A. Cauley, T. M. Vogt, K. E. Ensrud, H. K. Genant, and D. M. Black, ''Broadband ultrasound attenuation predicts fractures strongly and independently of densitometry in older women,'' Arch. Intern. Med. **157**, 629–634 (1997).
[15] M. Schott, S. Weill-Engerer, D. Hans, F. Duboeuf, P. D. Delmas, and P. J. Meunier, ''Ultrasound discriminates patients with hip fracture equally well as dual-energy x-ray absorptiometry and independently of bone mineral density,'' J. Bone Min. Res. **10**, 243–249 (1995).
[16] H. Turner, M. Peacock, L. Timmerman, J. M. Neal, and C. C. Johnston, Jr., ''Calcaneal ultrasonic measurements discriminate hip fracture independently of bone mass,'' Osteoporosis Int. **5**, 130–135 (1995).
[17] C. Glüer, S. R. Cummings, D. C. Bauer, K. Stone, A. Pressman, A. Mathur, and H. K. Genant, ''Osteoporosis: Association of recent fractures with quantitative U.S. findings,'' Radiology **199**, 725–732 (1996).
[18] P. Thompson, J. Taylor, A. Fisher, and R. Oliver, ''Quantitative heel ultrasound in 3180 women between 45 and 75 years of age: Compliance, normal ranges, and relationship to fracture history,'' Osteoporosis Int. **8**, 211–214 (1998).
[19] J. J. Faran, ''Sound scattering by solid cylinders and spheres,'' J. Acoust. Soc. Am. **23**, 405–418 (1951).
[20] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University, Princeton, NJ, 1986).
[21] L. Madsen, M. F. Insana, and J. A. Zagzebski, ''Method of data reduction for accurate determination of acoustic backscatter coefficients,'' J. Acoust. Soc. Am. **76**, 913–923 (1984).
[22] M. O'Donnell and J. G. Miller, ''Quantitative broadband ultrasonic backscatter: An approach to nondestructive evaluation in acoustically inhomogeneous materials,'' J. Appl. Phys. **52**, 1056–1065 (1981).
[23] K. A. Wear, M. R. Milunski, S. A. Wickline, J. E. Perez, B. E. Sobel, and J. G. Miller, ''Differentiation between acutely ischemic myocardium and zones of completed infarction in dogs on the basis of frequency-dependent backscatter,'' J. Acoust. Soc. Am. **85**, 2634–2641 (1989).
[24] G. Luo, J. J. Kaufman, A. Chabrera *et al.*, ''Computational methods for ultrasonic bone assessment,'' Ultrasound Med. Biol. **25**, 823–830 (1999).
[25] K. A. Wear, ''Frequency dependence of ultrasonic backscatter from human trabecular bone: Theory and experiment,'' J. Acoust. Soc. Am. **106**, 3659–3664 (1999).
[26] K. A. Wear, ''Anisotropy of ultrasonic backscatter and attenuation from human calcaneus: Implications for relative roles of absorption and scattering in determining attenuation,'' J. Acoust. Soc. Am. **107**, 3474–3479 (2000).
[27] S. Chaffai, V. Roberjot, F. Peyrin, G. Berger, and P. Laugier, ''Frequency dependence of ultrasonic backscattering in cancellous bone: Autocorrelation model and experimental results,'' J. Acoust. Soc. Am. **108**, 2403–2411 (2000).
[28] V. Roberjot, P. Laugier, P. Droin, P. Giat, and G. Berger, ''Measurement of integrated backscatter coefficient of trabecular bone,'' Proc.-IEEE Ultrason. Symp. **2**, 1123–1126 (1996).
[29] K. A. Wear and B. S. Garra, ''Assessment of bone density using broadband ultrasonic backscatter,'' Proceedings of the 22nd International Symposium on Ultrason. Imag. and Tissue Char., Washington, DC (1997), p. 14 (unpublished).
[30] P. Giat, C. Chappard, C. Roux, P. Laugier, and G. Berger, ''Preliminary clinical assessment of the backscatter coefficient in osteoporosis,'' Proceedings of the 22nd International Symposium on Ultrason. Imag. and Tissue Char., Washington, DC (1997), p. 16 (unpublished).
[31] K. A. Wear and B. S. Garra, ''Assessment of bone density using ultrasonic backscatter,'' Ultrasound Med. Biol. **24**, 689–695 (1998).
[32] K. A. Wear and D. W. Armstrong, ''The relationship between ultrasonic backscatter and bone mineral density in human calcaneus,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **47**, 777–780 (2000).
[33] K. A. Wear, A. P. Stuber, and J. C. Reynolds, ''Relationships of ultrasonic backscatter with ultrasonic attenuation, sound speed, and bone mineral density in human calcaneus,'' Ultrasound Med. Biol. **26**, 1311–1316 (2000).
[34] B. K. Hoffmeister, S. A. Whitten, and J. Y. Rho, ''Low-megahertz ultrasonic properties of bovine cancellous bone,'' Bone (N.Y.) **26**, 632–642 (2000).
[35] B. K. Hoffmeister, S. A. Whitten, S. C. Kaste, and J. Y. Rho, ''Effect of collagen and mineral content on the high-frequency ultrasonic properties of human cancellous bone,'' Bone (N.Y.) (in press).
[36] J. A. Zagzebski, L. X. Yao, E. J. Boote, and Z. F. Lu, in *Quantitative Backscatter Imaging in Ultrasonic Scattering in Biological Tissues*, edited by K. K. Shung and G. A. Thieme (CRC, Boca Raton, FL, 1993).
[37] K. A. Wear, B. S. Garra, and T. J. Hall, ''Measurements of ultrasonic backscatter coefficients in human liver and kidney *in vivo*,'' J. Acoust. Soc. Am. **98**, 1852–1857 (1995).
[38] K. A. Wear, ''The effects of frequency-dependent attenuation and dispersion on sound speed measurements: Applications in human trabecular bone,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **47**, 265–273 (2000).
[39] K. A. Wear, ''Measurements of phase velocity and group velocity in human calcaneus,'' Ultrasound Med. Biol. **26**, 641–646 (2000).
[40] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics*, 3rd ed. (Wiley, Toronto, 1982), Chap. 1.
[41] C. F. Njeh, D. Hans, J. Li, B. Fan, T. Fuerst, Y. Q. He, E. Tsuda-futami, Y. Lu, C. Y. Wu, and H. K. Genant, ''Comparison of six calcaneal quantitative ultrasound devices: Precision and hip fracture discrimination,'' Osteoporosis Int. **11**, 1051–1062 (2000).
[42] K. A. Wear, ''Ultrasonic attenuation in human calcaneus from 0.2 to 1.7 MHz,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **48**, 602–608 (2001).

[43] S. Cheng, B. Fan, L. Wang, T. Fuerst, M. Lian, C. Njeh, Y. He, M. Kern, M. Lappin, F. Tylavsky, D. Casal, S. Harris, and H. K. Genant, ''Factors affecting broadband ultrasound attenuation results of the calcaneus using a gel-coupled quantitative ultrasound scanning system,'' Osteoporosis Int. **10**, 495–504 (1999).
[44] M. L. Frost, G. M. Blake, and J. Fogelman, ''Contact quantitative ultrasound: An evaluation of precision, fracture discrimination, age-related bone loss, and applicability of the WHO criteria,'' Osteoporosis Int. **10**, 441–449 (1999).
[45] P. A. Hoel, *Introduction to Mathematical Statistics* (Wiley, New York, 1962), p. 167.
[46] J. M. Alves, W. Xu, D. Lin, R. S. Siffert, J. T. Ryaby, and J. J. Kaufmann, ''Ultrasonic assessment of human and bovine cancellous bone: A comparison study,'' IEEE Trans. Biomed. Eng. **43**, 249–258 (1996).

A bending wave simulator for investigating directional vibration sensing in insects

Ronald N. Miles,[a] Reginald B. Cocroft,[b] Colum Gibbons, and Daniel Batt
Department of Mechanical Engineering, State University of New York, Binghamton, New York 13902

(Received 12 September 2000; revised 2 March 2001; accepted 7 March 2001)

Substrate vibrations are important in social and ecological interactions for many insects and other arthropods. Localization cues include time and amplitude differences among an array of vibration detectors. However, for small species these cues are greatly reduced, and localization mechanisms remain unclear. Here we describe a method of simulating the vibrational environment that facilitates investigation of localization mechanisms in small species. Our model species was the treehopper *Umbonia crassicornis* (Membracidae; length 1 cm), which communicates using bending waves that propagate along plant stems. We designed a simulator consisting of a length of dowel and two actuators. The actuators were driven with two time signals that created the relationship between slope and displacement characteristic of steady-state bending wave motion. Because the surface of the dowel does not bend, as would a natural stem, close approximation of bending wave motion was limited to a region in the center of the dowel. An example of measurements of the dynamic response of an insect on the simulator is provided to illustrate its utility in the study of directional vibration sensing in insects. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1369106]

PACS numbers: 43.80.Ev, 43.80.Gx, 43.40.Yq, 43.40.At [WA]

I. INTRODUCTION

Substrate-borne vibrations are important in social and ecological interactions for many insects and other arthropods (reviewed in Markl, 1983; Gogala, 1985; Henry, 1994; Stewart, 1997; Barth, 1998). In contexts such as mate searching or prey detection, it is often necessary to locate the vibration source. In principle, the direction of propagation of substrate vibrations can be determined using differences in arrival time or amplitude between detectors that are sufficiently widely spaced. Vibration receptors in arthropods are typically in the legs (e.g., Kalmring, 1985), which form an array of detectors in contact with the substrate. For relatively large species (with legs spanning 5 to 10 cm), individuals are able to determine source direction using either arrival time differences (Brownell and Farley, 1979; Hergenroder and Barth, 1983; Aicher and Tautz, 1990) or an amplitude gradient (Brownell and Farley, 1979; Latimer and Schatral, 1983; Hergenroder and Barth, 1983; Cokl *et al.*, 1985). However, the size range of species using substrate vibration spans over two orders of magnitude. Smaller species are faced with a more difficult localization task. Stimulus wavelengths are many times longer than the insects' dimensions, and the time and amplitude differences among inputs are extremely small. It is not clear that the mechanisms used by large species can function at these smaller scales (see Brownell, 1977; Cokl *et al.*, 1985). If not, then any localization by small species must employ as-yet unknown directional mechanisms.

The scaling problem encountered in locating a vibration source is paralleled in acoustic sensing. For airborne sound, however, there is a clearer understanding of the mechanisms used at different size scales. The solutions provided by very small species in locating a sound source might therefore suggest mechanisms that can be applied across modalities. In sound localization, very small species can use a form of mechanical pre-processing to enhance directional cues. This principle is illustrated by the fly *Ormia ochracea* (Tachinidae), which locates singing crickets using sound wavelengths over 200 times greater than the distance between its two tympanal hearing organs. The biomechanical response of the tympana, which are mechanically coupled, converts an interaural time difference of 1.5 microseconds into large, directional amplitude differences that can be easily detected by the nervous system (Miles *et al.*, 1995).

In general, the form of mechanical directionality seen in the fly ear can be produced by the interaction of two resonant modes of vibration in a structure that responds to propagating waves (Miles *et al.*, 1995). One mode must respond to the phase, or the spatial gradient of the quantity being detected, while the other mode must respond to the spatial average of the quantity over the region sampled. Given these response properties, the two modes can combine to convert a small phase difference across the region into a large amplitude difference at different points on the structure (Gerzon, 1994). An analogous form of mechanical directionality could in principle be produced by a system responding to vibrational waves propagating across a substrate. In insects, however, the design of the system is likely to differ substantially from those used in sound reception. While sound receptors in insects are located within specialized exoskeletal structures that respond to pressure waves, vibration receptors are located in more generalized structures such as the legs. Mechanical directionality in the response to substrate vibrations, if it exists, is thus likely to involve motion of the body as a whole rather than of a localized sensory structure. If the primary receptors are in the legs, these are likely to respond to relative motion between the body and the substrate. Detect-

[a]Electronic mail: miles@binghamton.edu
[b]Present address: Division of Biological Sciences, University of Missouri, Columbia, MO 65211.

ing the propagation direction of waves on the substrate would then involve processing differences in the responses of receptors on different legs.

The properties of natural vibrational environments pose challenges for investigating the influence of direction on an insect's mechanical response to substrate vibration. Foremost is the heterogeneity and severe filtering properties of natural substrates (Michelsen *et al.*, 1982; Keuper and Kuhne, 1983). Because of these unpredictable influences on vibration transmission, it is difficult to deliver stimuli that are identical in every respect but propagation direction. For example, using a natural substrate, with two drivers some distance apart and an insect in the middle, the substrate will introduce changes in stimulus phase and frequency spectrum that can be difficult to compensate for. For these reasons, a careful investigation of mechanical directionality requires greater control over the vibrational environment experienced by the insect.

In this report we describe a new method for investigating the role of direction in the mechanical response to substrate vibrations, one that provides precise control over the vibrational environment. Our study system is a small insect that communicates using bending waves traveling through the stems of plants. We designed a simulator that provides accurate control of the frequency, intensity, and direction of propagation of a vibration stimulus. We first describe the study species and its natural vibrational environment. We then present the theoretical background used to simulate this environment. Finally, we describe the performance of the simulator relative to ideal bending waves propagating along a stem. We also illustrate the use of the simulator with an example of data that suggest directionality in the dynamic response of an insect to substrate vibrations.

II. STUDY SYSTEM

The stems and leaves of plants constitute one of the most widespread substrates used by arthropods in communication and predator-prey interactions (Henry, 1994; Stewart, 1997). In theory, a number of wave types could propagate along a rodlike structure such as a plant stem (Markl, 1983; Gogala, 1985; Cremer and Heckl, 1988). However, the most important wave type for biological interactions is bending waves (Cocroft *et al.*, 2000 and references therein). Because of the potential in natural substrates for reflected waves and unpredictable amplitude attenuation patterns, vibration transmission in nature will not always resemble a pure bending wave propagating in one direction along an ideal beam (e.g., Michelsen *et al.*, 1982). However, in spite of these constraints, bending waves are known to be used by insects and spiders that communicate using plant-borne vibrations (Michelsen *et al.*, 1982; Keuper and Kuhne, 1983; Barth, 1993, 1998).

The thornbug treehopper *Umbonia crassicornis* (Membracidae) is a sap-feeding insect about 1 cm long. Thornbugs spend most of their lives on the surface of their hostplants, and for much of this time they are sedentary (Wood, 1983). Accordingly, all of our vibration measurements have been made using freestanding individuals. The ability to do this is important, because restraining the insects would alter the mechanical response of the body. This species is also an appropriate study system because substrate-borne vibrations are central to their biology. Both adults and juveniles use communication signals that propagate through the plant in the form of bending waves (Cocroft *et al.*, 2000). These signals function in parent-offspring interactions (Cocroft, 1996, 1999) and in the mate-locating behavior of adults (Cocroft personal observations). Most of the energy in these signals is below 4000 Hz (Cocroft, 1999).

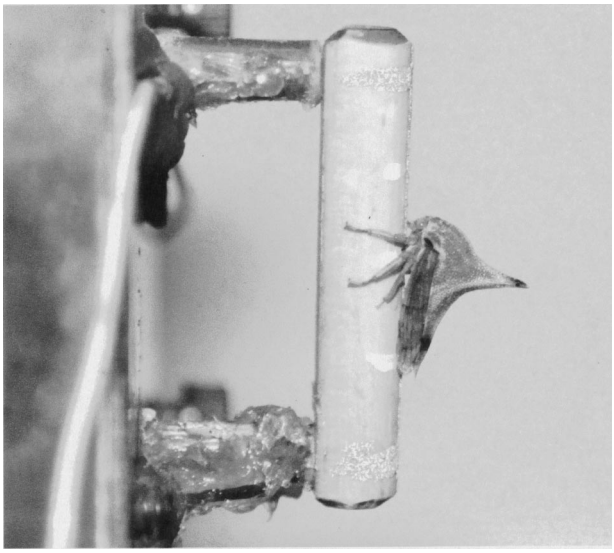


FIG. 1. *Umbonia crassicornis* (total length 10 mm) on the bending wave simulator. The insect's feet occupy a portion of the dowel (length 30 mm, diameter 6 mm) extending less than 1/8th the distance between the actuators in either direction from the middle. It is necessary to simulate a bending wave only over this narrow portion. From Cocroft *et al.* (2000), reprinted with permission.

III. THEORY

Our goal was to create a simulator that produced a vibration environment at the insect's feet that closely resembled that due to steady-state bending wave motion. We then wished to examine the mechanical response of the insect's body to the direction of substrate vibration; i.e., how the body moves relative to the substrate when the stimulus propagates from in front of or behind the insect. To measure the influence of direction on the mechanical response of the body, we needed to be able to electronically switch the direction of propagation between measurements, without altering the position or posture of the insect.

The simulator, as shown in Fig. 1, consists of a short segment of dowel (30 mm long, 6 mm in diameter) with piezoceramic actuators on each end. The actuators were bonded to the dowel with epoxy parallel to each other and perpendicular to the long axis of the dowel. The other ends of the actuators were bonded to a rigid metal fixture. Signals input into the actuators created motions at each end of the dowel so that the motion at its center mimicked that of bending waves on a typical host plant stem. The diameter of the dowel is similar to that of typical host plant stems, and the insect rests on the stem in a natural position.

In order to simulate bending wave motion at the center of the dowel, the two actuators must create a carefully con-

trolled combination of displacement and slope at the center. To examine the type of motion the simulator needs to create, consider the equation of motion for a beam in bending,

$$EI\frac{\partial^4 w(x,t)}{\partial x^4}+\rho A\ddot{w}(x,t)=0, \tag{1}$$

where E is Young's modulus of elasticity, I is the area moment if inertia, $w(x,t)$ is the beam displacement, ρ is the mass density, and A is the cross sectional area. x is the position along the beam and t is time.

If we assume that a steady-state bending wave travels in the positive direction the response can be written as

$$w(x,t)=We^{\hat{\imath}(\omega t-kx)}, \tag{2}$$

where W is the amplitude, ω is the frequency in radians/second, $\hat{\imath}=\sqrt{-1}$, $k=\omega/c$ is the wave number, and c is the propagation speed of the wave. Substitution of Eq. (2) into Eq. (1) gives the phase speed,

$$c=\sqrt{\omega}\left(\frac{B}{\rho A}\right)^{1/4}. \tag{3}$$

The propagation velocity of bending waves will vary in relation to the characteristics of the particular stem through which they are traveling. This velocity can be expressed as

$$c\approx c_c\sqrt{f}, \tag{4}$$

where c_c is a constant and f is the frequency in hertz. Measurements of the phase propagation velocity of bending waves in plant systems yield estimates of c_c ranging from 1.3 to 2.5 (calculated from data in Michelsen *et al.*, 1982). Phase propagation velocities measured in stems of *U. crassicornis* host plants (*Calliandra haematocephala*) yield estimates of c_c of 2.2 and 2.4 (Cocroft, unpublished data). For this study we chose a value of 2.4 (this value can easily be changed if a different velocity is desired). The wave number may then be expressed as

$$k=\frac{\omega}{c}=\frac{2\pi f}{c_c\sqrt{f}}=2\pi\sqrt{f}/c_c, \tag{5}$$

with units of radians/meter.

Because the simulator consists of a rigid dowel supported on two actuators (Fig. 1) it is possible to control only the slope and displacement of the motion at a given point along its length. The rigidity of the dowel does not allow actual bending motion which requires curvature of the substrate. This somewhat simplified apparatus does, however, permit us to closely approximate the displacement and rotation due to true bending wave motion, at least for a region at the center of the dowel. To do this, we need to determine the required motion at each end of the dowel such that the middle of the dowel has the desired relationship between slope and displacement.

For a displacement given by Eq. (2), the slope of the substrate is

$$w_x(x,t)=-\hat{\imath}kWe^{\hat{\imath}(\omega t-kx)}. \tag{6}$$

From Eqs. (2) and (6) it can be seen that the transfer function between the displacement and slope at any point will be

$$H_{ww_x}(\omega)=-\hat{\imath}k. \tag{7}$$

Our goal is to impose motion at the ends of the dowel such that the response in the vicinity of the insect follows Eq. (7). Let the displacements imposed at the ends of the dowel be $y_1(t)$ and $y_2(t)$. If we assume that the dowel remains rigid, the displacement, $y_m(t)$, and the slope, $\theta_m(t)$ at the middle will be

$$y_m(t)=\frac{y_1(t)+y_2(t)}{2},\quad \theta_m(t)=\frac{y_2(t)-y_1(t)}{d}, \tag{8}$$

where d is the length of the dowel. Equations (8) enable us to relate the transfer function between the motions at the ends of the dowel, $H_{y_1y_2}(\omega)$, to transfer function between the displacement and slope at the middle, $H_{y_m\theta_m}(\omega)$,

$$H_{y_1y_2}(\omega)=\frac{1+(d/2)H_{y_m\theta_m}(\omega)}{1-(d/2)H_{y_m\theta_m}(\omega)}=\frac{1-\hat{\imath}kd/2}{1+\hat{\imath}kd/2}, \tag{9}$$

where we have assumed that the relationship between displacement and slope at the middle of the dowel should be the same as given in Eq. (7),

$$H_{ww_x}(\omega)=-\hat{\imath}k=H_{y_m\theta_m}(\omega). \tag{10}$$

If it is possible to create two random displacements at the ends of the dowel, $y_1(t)$ and $y_2(t)$ that are related to each other through the transfer function given in Eq. (9), then the motion at the middle of the dowel will provide a reasonable simulation of a traveling bending wave.

Of course, since the dowel is assumed to remain straight, it clearly will not accurately represent bending motion over its entire length. If one wishes to design a simulator for bending waves over a specified range of frequencies and over a known portion of the dowel, it is helpful to be able to estimate the expected discrepancy between the motion of the dowel at each location along its length and the desired bending wave motion. Given the displacement and slope at the middle of the dowel, $y_m(t)$ and $\theta_m(t)$, the displacement at any point is given by

$$y(x,t)=y_m(t)+x\theta_m(t), \tag{11}$$

where x denotes the position along the dowel relative to the middle point. By using Eq. (11) it can be shown that the transfer function, $H_{y\theta_m}(\omega)$, between the displacement at any point, $y(x,t)$, and the slope $\theta_m(t)$, is

$$H_{y\theta_m}(\omega)=\frac{H_{y_m\theta_m}(\omega)}{(1-\hat{\imath}kx)}. \tag{12}$$

The ratio of the simulated and desired transfer functions as a function of position along the dowel is then

$$\frac{H_{y\theta_m}(\omega)}{H_{y_m\theta_m}(\omega)}=\frac{1}{(1-\hat{\imath}kx)}. \tag{13}$$

Equation (13) may be used to determine the error in the bending wave simulation as a function of position. The value of k may be determined as in the discussion preceding Eq. (5).

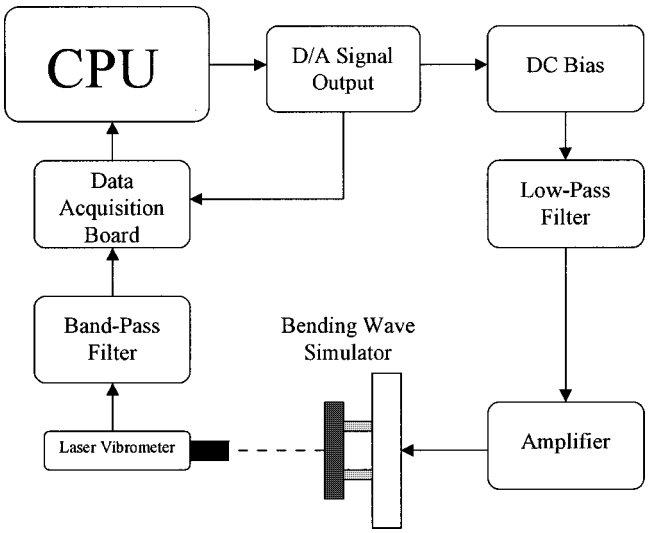


FIG. 2. Experimental setup for the bending wave simulator.

IV. METHODS

A. Experimental setup

A schematic of the experimental setup is shown in Fig. 2. A personal computer was used to generate the random signals sent to the actuators on the dowel and to acquire and analyze the vibration data. Random signals were used because they are very convenient for characterizing the dynamic response, which was the primary interest in this study. If the simulator is to be used for behavioral studies, the system will be modified to simulate insect calls. A Microstar DAP 2400/6 data acquisition PC board was used to output the two analog signals to drive the actuators. These two signals had the relative phase and amplitude needed to simulate bending wave motion in the central portion of the dowel. The details of the two-channel signal generation algorithm are given in the appendix. In order to drive the actuators, it was necessary to add a DC bias to the signals. The analog outputs of the DAP were low-pass filtered at 6 kHz by a Frequency Devices 90002 filter (48 dB/octave) and amplified by a DC coupled, Techron 5515 power amplifier before being sent to the pair of electrostrictive actuators (Xinetics XIRP0410L).

Motion of the simulator was measured with laser vibrometry, using a Polytec OFV 3000 controller and OFV 302 sensor head. The laser was positioned about 20 cm from the dowel, in the plane of motion of the actuators. The laser output was high-pass filtered at 100 Hz and low-pass filtered at 6 kHz using a Krohn–Hite 3550 filter (24 dB/octave). This signal, along with the two (AC coupled) output signals, was acquired using an Analogic FAST16 data acquisition board.

A compensation procedure was first run to account for differences in the output of the two actuators or in the amplitude of the two channels of the power amplifier. Identical signals were sent to the two actuators; the dowel motion produced at the site of each actuator was measured (with the other actuator silent) to obtain a calibration transfer function. This transfer function was used to correct the signal input into the lower of the two actuators so that it responded with the same amplitude and phase as the upper actuator. Absolute amplitude of the actuator output was adjusted to be similar to that of the insects' communication signals. As reported by Cocroft *et al.* (2000), RMS values from measurements of signals created by insects on typical plant stems ranged from 0.2 mm/s for females to 1.3 mm/s for males. RMS values for simulated random noise signals on the dowel near the insect's tarsus ranged from 0.1 mm/s to 0.3 mm/s.

Each measurement was stored in the form of a transfer function relative to the random signal sent to the upper actuator (acquired by the FAST16 board). Using the output signal as a standard allowed us to then compare nonsimultaneous measurements made along the dowel and thus characterize motion of the dowel as a whole.

The laser vibrometer and the simulator were mounted on a Newport RS 6000 $3' \times 6' \times 8''$ Optical Table to ensure that environmental vibrations did not influence the measurements.

B. Simulator performance

To examine the accuracy of the simulated bending waves, response measurements were made at nine equally spaced points along the entire length of the dowel. As mentioned above, all data were stored as transfer functions between the displacement response and the input into the upper actuator. Let $H_i(\omega)$ denote this measured transfer function corresponding to location i, where $i=1,...,9$. If we again assume that the dowel moves as a rigid body, the slope, or rotation of the dowel can be approximated by

$$\Theta(\omega)=\frac{H_9(\omega)-H_1(\omega)}{d}, \tag{14}$$

where, again, d is the total length of the dowel. The relationship between slope and displacement at location i may then be estimated from

$$H_{y_i\Theta}(\omega)=\frac{H_9(\omega)-H_1(\omega)}{dH_i(\omega)}. \tag{15}$$

To compare the measured transfer function between slope and displacement at each point with the desired results, let $G_i(\omega)$ denote the ratio of the results of Eqs. (10) and (15) corresponding to location i,

$$G_i(\omega)=\frac{H_{y_i\Theta}(\omega)}{H_{y_m\theta_m}(\omega)}=\frac{H_9(\omega)-H_1(\omega)}{-\hat{i}kdH_i(\omega)}. \tag{16}$$

One would expect that $G_i(\omega)=1.0$ when i corresponds to the middle of the dowel, $i=5$. In our case, according to Eq. (13), its magnitude would be less than 1 elsewhere.

C. Dynamic response of the insect's body

Characterization of the dynamic response of the insect's body when driven with substrate vibration is presented in Cocroft *et al.* (2000). However, we wished to provide here an example of an additional analysis that illustrates the nature of the data provided by use of the simulator. Details of the methods are provided in that paper, but, in brief, we placed healthy individual adult female *U. crassicornis* on the dowel so that the middle leg rested at the center of the dowel (see Fig. 1). Each animal took on a natural posture as when

standing on a host plant stem. Each insect was marked with small dots of reflective paint at four locations along the midline of the body from front to back. Laser measurements were made at each point along the body. For each point, the data acquisition program first took 10 samples with the stimulus propagating in one direction, then 10 samples with the stimulus propagating in the opposite direction. The laser was then moved to the next point. A reference measurement was made on the dowel surface, near the insect's middle leg, in order to calculate the motion of the insect relative to the substrate.

V. RESULTS

A. Simulator performance

Figure 3 shows measurements of the magnitude and phase of $G_i(\omega)$ for the dowel shown in Fig. 1. Again, $G_i(\omega)$ is a measure of simulator performance that compares the transfer function between slope and displacement expected in an ideal beam with the transfer function between slope and displacement measured on the dowel. The signals provided to the actuators are designed to accurately create the desired relation between slope and displacement at the center of the dowel. As shown in Eq. (7), the relation between slope and displacement in an ideal traveling bending wave is independent of position. Since the dowel moves with essentially the same slope all along its length, points that are away from the center must move with greater than the ideal displacement in order to create an accurate simulation at the dowel's center, $x=d/2$. As expected, the measurements shown in Fig. 3 indicate that the quality of the simulation degrades at points further from the middle. As shown in Fig. 1, the legs of the insect occupy a region smaller than $3d/8<x<5d/8$ (dowel length between the actuators is 2.4 cm, insects legs span approximately 5 mm).

Over the region occupied by the insect, the amplitude deviation is minimal (<2 dB). The phase deviation is larger (<40 degrees), reflecting the difference between an ideal beam, in which the surface curves as bending waves propagate, and the simulator, in which the dowel surface remains essentially straight. That is, the difference between the simulator and an ideal beam reduces to a difference in slope at a given point, increasing with distance from the center and with stimulus frequency. Because the insect's leg receptors probably do not measure slope at a single point, this deviation may not be directly applicable to vibration perception. Instead, the insect must detect the slope by comparing inputs between receptors at three separate points (corresponding to the attachment points of the legs) along the surface. The situation faced by the insect is thus more clearly reflected in the following figure.

To compare the motion of the dowel with that of an ideal bending wave in the time domain, Fig. 4 shows the motion versus time relative to one period at frequencies of 500 Hz and 2500 Hz, respectively. At 500 Hz the wavelength on an actual branch is quite long relative to the size of the dowel, and thus there is very little difference in the measured and ideal bending wave curves. At 2500 Hz, however, the dowel must undergo greater motion at the ends in order for the central region to provide an adequate simulation. In each case, the motion at the center, where the insect would be located, closely approximates the ideal case. That is, the overall slope as measured between the front and back legs of an insect on the simulator is very similar to that in an ideal beam, even though the slope at a given point [see Fig. 3(b)] differs due to the curvature of the ideal beam. The measured results shown in Fig. 4 also support our assumption that the dowel remains essentially straight at the frequencies of interest.

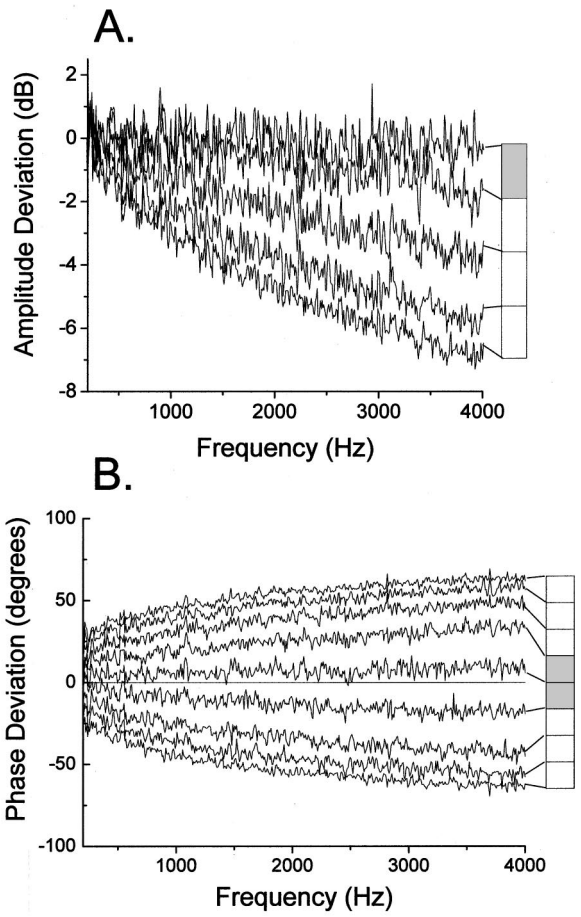


FIG. 3. Comparison of the magnitude (A) and phase (B) of the transfer function between slope and displacement measured on the dowel with that expected for an ideal bending wave [see Eq. (14)]. Note that for points away from $x=0$ (the center of the dowel), the amplitude of $G_i(\omega)$ decreases as predicted by Eq. (11). This decrease occurs because the ends of the dowel move with greater amplitude than the center. Because the slope across the dowel remains constant, the ratio of slope to displacement decreases at points away from the center. The relation between slope and displacement in the ideal bending wave is independent of position as in Eq. (5). The error in amplitude is less than 3 dB and the error in phase is less than 40 deg in the region of the dowel occupied by the insect (shaded area). In (A) only the points from the center to one end of the dowel are plotted, but the response is approximately symmetrical.

B. Vibration of the insect

To illustrate the use of the simulator in studying directional mechanisms in insects, we complement the extensive analysis presented in Cocroft *et al.* (2000) with an example of an additional kind of analysis using data measured on an adult female *U. crassicornis*. Figure 5 shows the operating

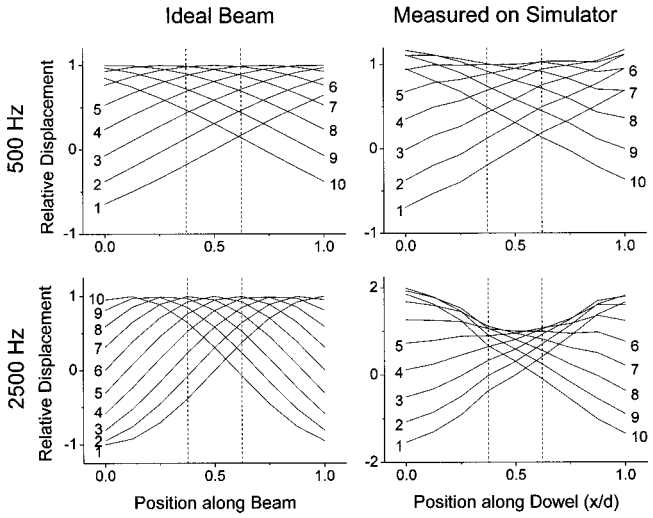


FIG. 4. Surface motion versus time during one period for an ideal beam and for the simulator. The dotted lines enclose the portion of the dowel spanned by the insect's legs. At 500 Hz the dowel length is only a fraction of a wavelength, and thus there is little difference between ideal and simulated motion. At 2500 Hz, achieving the necessary slope in the center requires greater motion of the ends of the dowel. Note the change in vertical scale for the measured results at 2500 Hz.

mode shapes measured on the insect's body for simulated bending waves traveling in opposite directions. The motion of the insect's thorax is similar to that of a rigid body on flexible legs. At lower frequencies, both ends move in phase. At higher frequencies, the back moves in the opposite direction from the front. Furthermore, the relative amplitude of motion between front and back differs depending on the direction of the stimulus. This is especially true of frequencies below 2000 Hz, which include much of the energy in the insects' communication signals (Cocroft, 1999). Note that the data were normalized with respect to motion measured on the dowel near the insect's middle leg. These data, along with those reported in Cocroft *et al.* (2000), suggest that the dynamic response of the insect's body provides a source of directional information.

VI. DISCUSSION

In this paper we describe a simulator that reproduces the vibrational environment of an insect that communicates using plant-borne vibrations. These vibrations, which travel along stems and leaves, constitute one of the most important sources of information about the environment for insects and other arthropods (Markl, 1983; Gogala, 1985; Henry, 1994; Stewart, 1997; Barth, 1998). Furthermore, for many of the social and ecological interactions mediated by these plant-borne vibrations, localization of the vibration source is an important task (Cokl *et al.*, 1985; Pfannenstiel *et al.*, 1995). For large species (5–10 cm), arrival time delays (and perhaps amplitude differences) between an array of vibration receptors in the legs provide sufficient directional cues (Hergenroder and Barth, 1983; Latimer and Schatral, 1983; Cokl *et al.*, 1985). However, for small species (<1 cm), time and amplitude differences are much smaller, raising the question of how these species might localize a vibration source. The

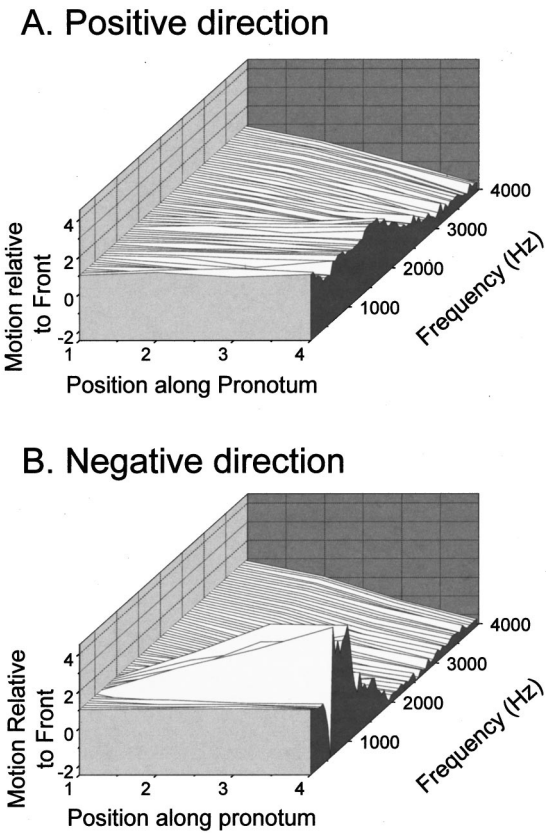


FIG. 5. An example from one insect of the effect of direction of substrate vibrations on the dynamic response of the body (for full results see Cocroft *et al.*, 2000). The measured motion is shown versus frequency, measured on the insect for simulated waves traveling in positive and negative directions. In each case the insect's motion is similar to a rigid body on flexible legs. At lower frequencies, both ends of the insect move in phase. At higher frequencies (>2500 Hz), the back moves in the opposite direction from the front. For waves traveling in the negative direction, the back of the insect moves with significantly greater amplitude than the front. The frequency dependence of the operating mode shapes is strongly influenced by the direction of propagation.

simulator described here is a useful tool that facilitates the investigation of directional sensing in small arthropods.

The simulator closely approximates the relationship between slope and displacement characteristic of bending waves propagating along an ideal beam. Because the dowel is straight, it cannot reproduce bending wave motion along its entire length, but only in a region in the center of the dowel. The width of this region varies with the wavelength of the stimulus: in order to create the required surface motion in the center, the ends must be moved through a greater and greater excursion as frequency increases. For our study species, the treehopper *U. crassicornis*, the biologically relevant frequency range appears to be mostly under 4 kHz (Cocroft, 1999). For this frequency range, the region of the simulator occupied by an insect (legs spanning 5 mm at the center of the dowel) closely approximates the vibrational environment created by propagating bending waves. The simulator allowed us to investigate the influence of stimulus direction on the dynamic response of an insect's body to substrate vibration. Detailed results presented in Cocroft *et al.* (2000) reveal a marked mechanical directionality in the motion of the insects body relative to the substrate. In this report we pro-

vided an additional type of analysis, also revealing an effect of direction on the motion measured across the insect's body. The dynamic response of the body to vibration of the substrate converts the small time differences between the arrival of the stimulus at the front and back legs (55 μs at 500 Hz) into a relatively large amplitude difference across the body (about 10 dB at some frequencies). This difference may arise from the interaction of the two modes of vibration (rotational and translational) detected in the motion of the body (see Gerzon, 1994). Because motion at the front and back of the body differs depending on stimulus direction, it provides robust directional cues that could in principle be detected by the nervous system (this has not yet been investigated; see discussion in Cocroft *et al.*, 2000). These results raise the possibility of a new means of directional sensing in insect vibrational perception.

In the case of our study species, the performance of the simulator most closely approximated bending waves across a small area in the center of the dowel spanned by the insect's legs, and for lower frequencies that contain most of the energy in the species' own signals. Use of this simulator design for other species is warranted if their dimensions, and the frequencies used, fall within a similar range. For larger insects, Eq. (11) can be used to assess whether this apparatus is appropriate (or, alternatively, to estimate the dimensions of a suitable apparatus).

This experimental apparatus should facilitate investigation of other questions involving vibration perception in insects. For example, the creation of steady-state bending wave motion should facilitate investigation of the neural coding of the directional information present in the dynamic response of the body. Furthermore, use of natural time signals that evoke an orienting response (such as those used in mate location; e.g., Cokl *et al.*, 1999) will allow behavioral assays of directional ability and of the signal features essential for directional sensing.

VII. SUMMARY

We describe a bending-wave simulator capable of reproducing the vibrational environment of many small insects and other arthropods that use plant-borne vibrations in social and ecological interactions. The purpose of this simulator is to facilitate the study of directional sensing mechanisms, because it can mimic bending waves propagating in opposite directions along a stem, thus delivering stimuli that are identical in all features except for direction of propagation. This task is difficult using natural substrates.

The simulator consists of a short section of dowel driven by two actuators. In the case of bending waves propagating along an ideal beam, the relationship between slope and displacement is independent of position along the beam. This requires curvature of the stem. In contrast, the dowel segment remains essentially straight (i.e., the slope is constant across the dowel). As a result, exact reproduction of the relationship of slope and displacement characteristic of bending waves is possible only at the center of the dowel. The simulator departs from this relationship toward the ends of the dowel, because the ends must be moved with a greater amplitude in order to achieve the necessary slope in the middle.

For a small insect positioned at the center of the dowel, and using the relatively low frequency range of biologically relevant signals, the simulator provides a close approximation of propagating bending waves. Measurements of the dynamic response of the body of a tree-hopper (*Umbonia crassicornis*) reveal a mechanical directionality that may provide a means of vibration localization.

ACKNOWLEDGMENTS

R.N.M., R.B.C., and C.B. were supported by NIDCD Grant No. 1R01DC03926-01. R.N.M., C.B., and D.B. were also supported by NSF Grant No. BCS9315854. Support for R.B.C. was also provided by the Olin Foundation. We thank D. Kunkel for Fig. 1.

APPENDIX: CONSTRUCTION OF THE EXCITATION SIGNALS

In the following, we describe the procedure for creating the signals input into the actuators on each end of the dowel shown in Fig. 1. The approach is an extension of that originally due to Rice (1954) and adapted by Shinozuka (1972) and Miles (1992).

Given the power spectral density for a complex system one can construct a time series that will have the desired power spectrum. If $S_{y_1y_1}(\omega)$ is the double sided power spectrum of the desired signal, $y_1(t)$, the $y_1(t)$ may be approximated by

$$y_1(t) = 2\sum_{n=0}^{N-1} [S_{y_1y_1}(\omega_n)\Delta\omega]^{1/2} \cos(\omega_n t - \phi_n), \qquad \text{(A1)}$$

where ϕ_n are uniformly distributed random numbers on the interval from 0 to 2π and

$$\omega_n = n\Delta\omega, \qquad \text{(A2)}$$

with

$$\Delta\omega = \omega_{\max}/N, \qquad \text{(A3)}$$

where $\omega_{\max}$ is the maximum frequency in the power spectrum $S_{y_1y_1}(\omega)$, and N is the total number of terms in the summation in Eq. (A1).

Equation (A1) simulates the time series as a distribution of sinusoidal signals having random phases. Unfortunately, this expression requires the computation of a large number of cosine functions at each desired value of the time, *t*. A considerable improvement in computational efficiency can be obtained by recasting Eq. (A1) to allow the use of the fast-Fourier transform. To accomplish this, note that Eq. (A1) may be written as

$$y_1(t) = 2\,\mathrm{Re}\left[\sum_{n=0}^{N-1} [S_{y_1y_1}(\omega_n)\Delta\omega]^{1/2} e^{\hat{\imath}\omega_n t} e^{-\hat{\imath}\phi_n}\right], \qquad \text{(A4)}$$

where Re[·] denotes the real part and, as before, $\hat{\imath}$ is $\sqrt{-1}$.

If the simulated time series, $y_1(t)$ is needed only at discrete values of time, *t*, then let

$$y_j = y_1(t_j) = y_1(j\Delta t), \qquad \text{(A5)}$$

where the time duration between the equally spaced sample times is Δt. Evaluating Eq. (A4) at $t=t_j$ gives

$$y_1(j\Delta t)=2\,\mathrm{Re}\left[\sum_{n=0}^{N-1}[S_{y_1y_1}(\omega_n)\Delta\omega]^{1/2}e^{\hat{\imath}\omega_n j\Delta t}e^{-\hat{\imath}\phi_n}\right]. \tag{A6}$$

To satisfy the Nyquist sampling criterion, the time series, $y_1(t)$ must be sampled at a high enough rate to obtain two samples during one period of the highest frequency component in the original input power spectrum, $S_{y_1y_1}(\omega)$. This gives

$$\Delta t=\frac{2\pi}{\omega_{\max}}\frac{1}{2}. \tag{A7}$$

Substituting Eqs. (A3) and (A7) into Eq. (A6) gives

$$y_1(j\Delta t)=2\,\mathrm{Re}\left[\sum_{n=0}^{N-1}[S_{y_1y_1}(\omega_n)\Delta\omega]^{1/2}e^{-\hat{\imath}\phi_n}e^{\hat{\imath}nj2\pi/2N}\right]. \tag{A8}$$

Equation (A8) may be evaluated using a fast-Fourier transform (FFT) algorithm by noting that given a discrete sequence, a_n, the FFT provides an efficient means of computing A_j, where

$$A_j=\sum_{n=0}^{N-1}a_n e^{-\hat{\imath}2\pi jn/N},\qquad \text{for } j=0,1,2,...,N-1. \tag{A9}$$

Equation (A8) may be evaluated using a FFT by defining a sequence,

$$\begin{aligned}a_n&=2[S_{y_1y_1}(\omega_n)\Delta\omega]^{1/2}e^{-\hat{\imath}\phi_n},\quad \text{for } n\leqslant N-1\\&=0,\quad \text{for } n\leqslant N.\end{aligned} \tag{A10}$$

Equation (A8) may then be written as

$$y_{1_j}=\mathrm{Re}\left[\sum_{n=0}^{2N-1}a_n e^{\hat{\imath}nj2\pi/2N}\right],\quad \text{for } j=0,1,2,...,2N-1, \tag{A11}$$

where $y_{1_j}=y_1(j\Delta t)$. Because we are taking the real part of the result of the summation, taking the complex conjugate of the right-hand side of Eq. (A11) gives

$$y_{1_j}=\mathrm{Re}\left[\sum_{n=0}^{2N-1}a_n e^{-\hat{\imath}nj2\pi/2N}\right],\quad \text{for } j=0,1,2,...,2N-1. \tag{A12}$$

This is equivalent to

$$y_{1_j}=\mathrm{Re}[\mathrm{FFT}(a_n)], \tag{A13}$$

where FFT$[\cdot]$ denotes the fast-Fourier transform. Note that the length of the sequence, a_n is $2N$.

The construction of another time series, $y_2(t)$ which is related to $y_1(t)$ according to the transfer function in Eq. (7) may be accomplished by defining a sequence that is analogous to Eq. (A10),

$$\begin{aligned}b_n&=2[S_{y_1y_1}(\omega_n)\Delta\omega]^{1/2}e^{-\hat{\imath}\phi_n}H_{y_1y_2}(\omega_n),\quad \text{for } n\leqslant N-1\\&=0,\quad \text{for } n\geqslant N,\end{aligned} \tag{A14}$$

where the transfer function $H_{y_1y_2}(\omega_n)$ is computed as in Eq. (7),

$$H_{y_1y_2}(\omega_n)=\frac{1-\hat{\imath}k_nd/2}{1+\hat{\imath}k_nd/2} \tag{A15}$$

and

$$k_n=\frac{\omega_n}{c}=\frac{2\pi f_n}{c_c\sqrt{f_n}}=2\pi\sqrt{f_n}/c_c, \tag{A16}$$

with $\omega_n=2\pi f_n$. The time series is then computed from

$$y_{2_j}=\mathrm{Re}[\mathrm{FFT}(b_n)], \tag{A17}$$

with $y_2(j\Delta t)=y_{2_j}$ for $j=0,1,2,...,2N-1$.

Aicher, B., and Tautz, J. (**1990**). ''Vibrational communication in the fiddler crab *Uca pugilator*. I. Signal transmission through the substratum,'' J. Comp. Physiol. **166**, 345–353.

Barth, F. G. (**1993**). ''Sensory guidance in spider pre-copulatory behaviour,'' Comp. Biochem. Physiol. **104A**, 717–733.

Barth, F. G. (**1998**). ''The vibrational sense of spiders,'' in *Comparative Hearing: Insects*, edited by R. R. Hoy, A. N. Popper, and R. R. Fay, Springer Handbook of Auditory Research, series edited by R. R. Fay and A. N. Popper (Springer, New York), pp. 228–278.

Brownell, P. H. (**1977**). ''Compressional and surface waves in sand: used by desert scorpions to locate prey,'' Science **197**, 479–482.

Brownell, P. H., and Farley, R. D. (**1979**). ''Orientation to vibrations in sand by the nocturnal scorpion *Paruroctonus mesaensis*: mechanism of target localization,'' J. Comp. Physiol. **131**, 31–38.

Cocroft, R. B. (**1996**). ''Insect vibrational defence signals,'' Nature (London) **382**, 679–680.

Cocroft, R. B. (**1999**). ''Offspring-parent communication in a subsocial treehopper (Hemiptera: Membracidae: *Umbonia crassicornis*),'' Behaviour **136**, 1–21.

Cocroft, R. B., Tieu, T., Hoy, R. R., and Miles, R. N. (**2000**). ''Directionality in the mechanical response to substrate vibration in a treehopper (Hemiptera: Membracidae: *Umbonia crassicornis*),'' J. Comp. Physiol. B **186**, 695–705.

Cokl, A., Otto, C., and Kalmring, K. (**1985**). ''The processing of directional vibratory signals in the ventral nerve cord of *Locusta migratoria*,'' J. Comp. Physiol. A **156**, 45–52.

Cokl, A., Virant-Doberlet, M., and McDowell, A. (**1999**). ''Vibrational directionality in the southern green stink bug, *Nezara viridula* (L.), is mediated by female song.'' Anim. Behav. **58**, 1277–1283.

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

Gerzon, M. A. (**1994**). ''Applications of Blumlein shuffling to stereo microphone techniques,'' J. Audio Eng. Soc. **42**, 435–453.

Gogala, M. (**1985**). ''Vibrational communication in insects (biophysical and behavioural aspects),'' in *Acoustic and Vibrational Communication in Insects*, edited by K. Kalmring and N. Elsner (Verlag Paul Parey, Berlin), pp. 117–126.

Henry, C. S. (**1994**). ''Singing and cryptic speciation in insects,'' Trends Ecol. Evol. **9**, 388–392.

Hergenroder, R., and Barth, F. G. (**1983**). ''Vibratory signals and spider behavior: how do the sensory inputs from the eight legs interact in orientation?,'' J. Comp. Physiol. **152**, 361–371.

Kalmring, K. (**1985**). ''Vibrational communication in insects (reception and integration of vibratory information),'' in *Acoustic and Vibrational Communication in Insects*, edited by K. Kalmring and N. Elsner (Verlag Paul Parey, Berlin).

Keuper, A., and Kuhne, R. (**1983**). ''The acoustic behaviour of the bushcricket *Tettigonia cantans*. II. Transmission of airborne-sound and vibration signals in the biotope,'' Behav. Processes **8**, 125–145.

Latimer, W., and Schatral, A. (**1983**). ''The acoustic behaviour of the bushcricket *Tettigonia cantans*. I. Behavioural responses to sound and vibration,'' Behav. Processes **8**, 113–124.

Markl, H. (**1983**). ''Vibrational communication,'' in *Neurobiology and Behavioral Physiology*, edited by F. Huber and H. Markl (Springer-Verlag, Berlin), pp. 332–353.
Michelsen, A., Fink, F., Gogala, M., and Traue, D. (**1982**). ''Plants as transmission channels for insect vibrational songs,'' Behav. Ecol. Sociobiol. **11**, 269–281.
Miles, R. N. (**1992**). ''Effect of spectral shape on acoustic fatigue life estimates.'' J. Sound Vib. **153**, 376–386.
Miles, R. N., Robert, D., and Hoy, R. R. (**1995**). ''Mechanically coupled ears for directional hearing in the parasitoid fly *O. ochracea*,'' J. Acoust. Soc. Am. **98**, 2059–2070.
Pfannenstiel, R. S., Hunt, R. E., and Yeargan, K. V. (**1995**). ''Orientation of a hemipteran predator to vibrations produced by feeding caterpillars,'' J. Insect Behav. **8**, 1–9.
Rice S. O. (**1954**). *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover, New York).
Shinozuka, M. (**1972**). ''Monte Carlo solution of structural dynamics,'' Comput. Struct. **2**, 855–874.
Stewart, K. W. (**1997**). ''Vibrational communication in insects: epitome in the language of stoneflies?,'' Am. Entomol. **1997**, 81–91.
Wood, T. K. (**1983**). ''Brooding and aggregating behavior of the treehopper, *Umbonia crassicornis*,'' Natl. Geogr. Soc. Res. Rep. **15**, 753–758.

Ultrasound-mediated disruption of cell membranes. I. Quantification of molecular uptake and cell viability

Héctor R. Guzmán, Daniel X. Nguyen, Sohail Khan, and Mark R. Prausnitz[a)]
School of Chemical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0100

(Received 18 August 2000; accepted for publication 4 April 2001)

Ultrasound-mediated drug delivery is a nonchemical, nonviral, and noninvasive method for targeted transport of drugs and genes into cells. Molecules can be delivered into cells when ultrasound disrupts the cell membrane by a mechanism believed to involve cavitation. This study examined molecular uptake and cell viability in cell suspensions (DU145 prostate cancer and aortic smooth muscle cells) exposed to varying peak negative acoustic pressures (0.6–3.0 MPa), exposure times (120–2000 ms), and pulse lengths (0.02–60 ms) in the presence of Optison (1.7% v/v) contrast agent. With increasing pressure and exposure time, molecular uptake of a marker compound, a calcein, increased and approached equilibrium with the extra cellular solution, while cell viability decreased. Varying pulse length produced no significant effect. All viability and molecular uptake measurements collected over the broad range of ultrasound conditions studied correlated with acoustic energy exposure. This suggests that acoustic energy exposure may be predictive of ultrasound's nonthermal bioeffects. © *2001 Acoustical Society of America.*
[DOI: 10.1121/1.1376131]

PACS numbers: 43.80.Gx, 43.80.Sh [FD]

I. INTRODUCTION

Over a range of conditions, ultrasound can transport molecules into viable cells by a mechanism believed to involve the transient disruption of cell membranes. Ultrasound has been demonstrated to deliver fluorescent dextran molecules (Fechheimer *et al.*, 1986; Miller *et al.*, 1999), genetic material (Bao *et al.*, 1997; Greenleaf *et al.*, 1998), and chemotherapeutic compounds (Saad and Hahn, 1987; Harrison *et al.*, 1996) into viable cells. Involving a related mechanism, molecular uptake has also been observed following exposure of cells to lithotripsy shock waves (Holmes *et al.*, 1992; Prat *et al.*, 1993). Although these and other studies highlight the potential to use ultrasound for drug delivery, the quantitative dependence of bioeffects on acoustic parameters are insufficiently understood. To gain a better understanding of ultrasound and its biological effects, this study measured the dependence of bioeffects on acoustic pressure, exposure time, and pulse length by quantifying molecular uptake and cell viability on a cell-by-cell basis in two cell lines. Our hypothesis is that bioeffects correlate with acoustic energy exposure (J/cm^2). Energy exposure may in turn relate to the strength of cavitation produced by ultrasound.

Ultrasound-mediated bioeffects are generally believed to be caused by cavitation, in the absence of ultrasonic heating (Carstensen *et al.*, 1993). Cavitation is typically generated through activation of small dissolved gas nuclei by an acoustic pressure field. These nuclei, which grow through rectified diffusion, may oscillate and implode violently, thereby releasing a burst of energy that may be sufficient to disrupt cell membranes. Under some conditions, cavitation can cause irreversible cell damage resulting in cell death (Miller *et al.*, 1996). Under other conditions, however, cavitation may reversibly disrupt cell membranes and thereby permit the entry of molecules into cells. It is the later phenomenon that we seek to achieve and control for drug delivery and other applications.

Other mechanical methods, such as cell scraping (McNeil, 1989) and syringe loading (Clarke and McNeil, 1992), are believed to apply shear forces that transiently disrupt cell membranes, which permit large molecules and genes to enter into cells. Electrical methods, such as electroporation (Chang *et al.*, 1992), have also been employed to transport molecules across reversibly-disrupted cell membranes. These techniques are believed to physically disrupt cell membranes by a process independent of drug chemistry and, thus, may be employed to deliver a wide variety of drugs or genes. However, these mechanical and electrical techniques are generally invasive if applied *in vivo*. In contrast, ultrasound can be focused noninvasively from outside the body through the use of focused transducers. Local introduction of contrast agents may improve targeting further by lowering the pressure threshold required for cavitation in specified regions of the body (Holland and Apfel, 1990; Miller and Thomas, 1995).

II. MATERIALS AND METHODS

A. Maintenance and preparation of cells

DU145 human prostate cancer cells (DU145; American Type Culture Collection, item no. HTB-81, lot no. 1145858) were cultured as monolayers in a humidified atmosphere of 95% air and 5% CO_2 at 37 °C in RPMI-1640 media, supplemented with 10% (v/v) heat inactivated fetal bovine serum and 100-μg/ml penicillin-streptomycin (Cellgro, Mediatech, Herndon, VA). Human aortic smooth muscle cells (AoSMC; Clonetics, catalog no. CC-2571, lot no. 7F0787) were initi-

a)Electronic mail: mark.prausnitz@che.gatech.edu

 0001-4966/2001/110(1)/588/9/$18.00

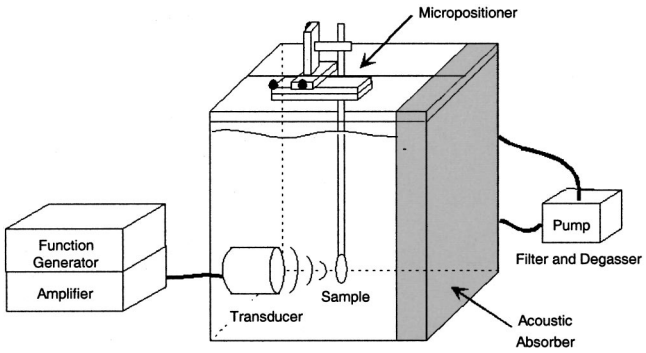


FIG. 1. The ultrasound apparatus consisted of a Plexiglas tank with a 500-kHz transducer to apply ultrasound and an acoustic absorber to minimize wave reflection. A function generator and amplifier powered the transducer. The filter and degassing unit removed cavitation nuclei from water in the tank. A micropositioner positioned the sample in the ultrasound focus.

ated from frozen stock and harvested at passage 7 prior to each experiment. They were cultured as monolayers in a humidified atmosphere of 95% air and 5% CO_2 at 37 °C in MCDB-131 media supplemented with 10% (v/v) heat inactivated fetal bovine serum, 100-μg/ml penicillin-streptomycin, and 2-mM L-glutamine (Cellgro).

Both cell types were harvested prior to each experiment by trypsin/EDTA (Cellgro) digestion at 80%–90% monolayer confluence and centrifuged at $1000\times g$ (Beckman Coulter, Fullerton, CA, model GS-15R) for 6 min. The supernatant was discarded and cell pellets were resuspended in cell media to a concentration of 1×10^6 cells/ml, as determined by a haemocytometer (model 3200, Hauser Scientific, Horsham, PA) or Coulter Counter (Coulter Multisizer II, Beckman Coulter). Cell radii were also determined using a Coulter Counter. Cell volumes were then calculated assuming the cells were spherical while in suspension.

Calcein (623 Da, radius = 0.6 nm; Molecular Probes, Eugene, OR, catalog No. C-481), a green-fluorescent molecule that cannot cross intact cell membranes, was used to quantitatively monitor the transport of molecules across the membranes of viable cells. Calcein was added to the cell suspension to attain a final concentration of 10 μM. Albumin-stabilized gas bubbles (Optison, Mallinckrodt, St. Louis, MO, catalog no. 2707-03) stored in a locking, airtight syringe (SampleLock Syringe Hamilton, VWR, catalog no. 60373-612) were slowly added to the cell suspension using a 22-gauge flat needle (400 μm i.d.) to promote acoustic cavitation (17 ± 0.05 μl/ml, approximately 1.1×10^7 bubbles/ml; bubble diameter = 2.0–4.5 μm).

B. Experimental equipment

Ultrasound was generated by a submersible focused piezoceramic transducer (Techno Scientific, Woodbridge, Ontario, Canada) with a 3-in. diameter. The transducer had a 3-$\frac{1}{2}$-in. focal length and a (−6-dB intensity area) beam width of 3 mm at the focal beam point. A 500-kHz sinusoid generated by a programmable wave form generator (Stanford Research Instruments, Sunnyvale, CA, model no. DS345) and amplified by a custom tone burst amplifier (Techno Scientific) powered and controlled the response of the transducer. As displayed in Fig. 1, the transducer was housed in a polycarbonate tank ($12\times11\frac{1}{2}\times14\frac{1}{2}$ in.) containing approximately 26 l of de-ionized, distilled, and partially degassed water at room temperature, 23–24 °C. Mounted opposite the transducer, a 2-in.-thick acoustic absorber (SC-501 Acoustic Rubber, Sonic Concepts, WA) minimized standing wave formation. A degassing unit (Kaiser *et al.*, 1996) and a 0.2-μm filtration unit (Fin-L-Filter, Cole-Parmer, Vernon Hills, IL) were operated for 3 h prior to each experiment to remove potential cavitation nucleation sites within the water bath. For calibration of the apparatus, spatial-peak–temporal-peak (SPTP) negative pressure ($\bar{P}_{\mathrm{SPTP}}$) was measured at the focal beam point using a 0.2-mm aperture PVDF membrane hydrophone (NTR Systems, Seattle, WA, model no. MHA200A) in the absence of a sample container.

The spatial-peak (SP) acoustic energy exposure (E_{SP}) measured in this study was approximated by the product of the spatial-peak-pulse-average (SPPA) acoustic intensity (I_{SPPA}) of a plane traveling wave and the total ultrasound exposure time, t,

$$I_{\mathrm{SPPA}}=\frac{P^2}{\rho c}, \tag{1}$$

$$E_{\mathrm{SP}}=I_{\mathrm{SPPA}}\,t, \tag{2}$$

where P is rms pressure, ρ the density of water, and c the speed of sound in water. The following assumptions were made: (1) Ultrasound impinging on the cell sample was composed only of plane traveling waves, (2) the cell sample container was acoustically transparent, and (3) all of the acoustic energy measured at the focal beam point was transferred to the cell sample for all of the ultrasound conditions tested. In reality, however, a large fraction of the measured acoustic energy passes through the sample and is collected in the absorber. Thus, the energy actually delivered to cells is interpreted as being less than, but approximately proportional to, the reported energy exposure values.

C. Ultrasound protocol

Prior to ultrasound exposure, cell samples were slowly aliquoted via a 3-ml syringe (Becton Dickinson, Franklin Lakes, NJ) with a 22-gauge needle (Perkin Elmer, Foster City, CA) into 1.2-ml polyethylene transfer pipets (8.8 mm i.d., 0.3 mm wall thickness, and 3 cm height; Samco, San Fernando, CA, catalog no. 241). A metal rod attached to a three-way micropositioner (1-mm resolution, Velmex, Bloomfield, NY) was inserted into the pipet orifice to plug it closed and to hold and position the chamber in the focal beam point of the transducer (Fig. 1).

Ultrasound was delivered using pulses at a 6% duty cycle and with pulse lengths that varied from 20 μs to 60 ms. Exposure time was varied from 120 to 2000 ms. Since a 6% duty cycle was used, the actual duration of each experiment ranged from 2 to 34 s. $\bar{P}_{\mathrm{SPTP}}$ was varied between 0.6 and 3.0 MPa. ''Sham'' control exposures were conducted using the same protocol, but no ultrasound was applied. Postexposure cell samples were immediately transferred to 1.5-ml microcentrifuge tubes and left to incubate for 5 min at room tem-

perature to permit the cells to ''recover'' (Keyhani *et al.*, 1998). The samples were then placed on ice until all of the samples were exposed (10–30 min).

Cell samples were washed and centrifuged ($800\times g$, 4 min, Eppendorf 5415C, Brinkman, Westbury, NY) three times to remove calcein present in the extracellular fluid (i.e., supernatant). Cell pellets were resuspended to a final volume of 0.5 ml in Dulbecco's phosphate buffered saline (Cellgro) and incubated for at least 10 min with red-fluorescent propidium iodide (PI) solution (Molecular Probes, catalog no. P-1304) to stain nonviable cells. Fluorescent calibration beads (Molecular Probes, catalog no. L-14821) were added at a concentration of 2.4×10^4 beads/ml to facilitate cell viability analysis, as described previously (Prausnitz *et al.*, 1993).

To verify that ultrasound exposure did not produce thermal effects, the temperature rise in a sample was measured at the largest peak negative pressure and longest exposure duration used in this study (i.e., 3.0 MPa and 2000 ms). Using a digital thermometer (HI 98501, Hanna Instruments, Woonsocket, RI) placed in the sample immediately after ultrasound exposure, the temperature rise in the sample was found to be less than 1 °C. To determine if a small temperature rise might induce bioeffects, samples were quickly heated (without ultrasound exposure) in a 37 °C water bath to a temperature increase of 5 °C (i.e., from 24 to 29 °C). Analysis of the samples showed no difference between the heated samples and unheated controls (Student's t-test $p>0.5$), and thus any bioeffects observed in this study were considered to be nonthermal.

D. Quantification of bioeffects

Optical properties of the cell suspensions were measured using a FacsVantage SE flow cytometer with Cell Quest software (Becton Dickinson). A minimum of 20 000 viable cells, collected at a rate of 1 700 cells/s, was analyzed per sample to ensure that a statistically significant population was collected. Cell samples were excited with a 488-nm laser (Enterprise II, Innova, Coherent, Palo Alto, CA). Light scatter, collected by two photodiodes (forward scatter and side scatter), was used to determine the size and shape of particles (e.g., cells, debris, microspheres) in the sample. Fluorescence measurements, collected by two photomultiplier tubes, were used to distinguish viable from nonviable cells (PI fluorescence, 665- to 685-nm bandpass filter) and to measure calcein uptake (calcein fluorescence, 515- to 545-nm bandpass filter).

Figure 2 is a typical forward scatter—PI fluorescence plot of the collected raw data. The plot distinguishes viable cells (a), dead cells (b), and fluorescent beads (c). Viable cells show weak fluorescence in the red channel due to autofluorescence, optical and electrical noise, and possible low-level staining by propidium iodide. By comparing the ratio of cells to fluorescent beads in each sample to that of the control samples, it can be determined if cells were destroyed or otherwise ''lost'' during the protocol (Prausnitz *et al.*, 1993). Quantitative calibration beads (Flow Cytometry Standards Corporation, Fishers, IN, catalog No. 825) were used to convert calcein fluorescence into an average number of

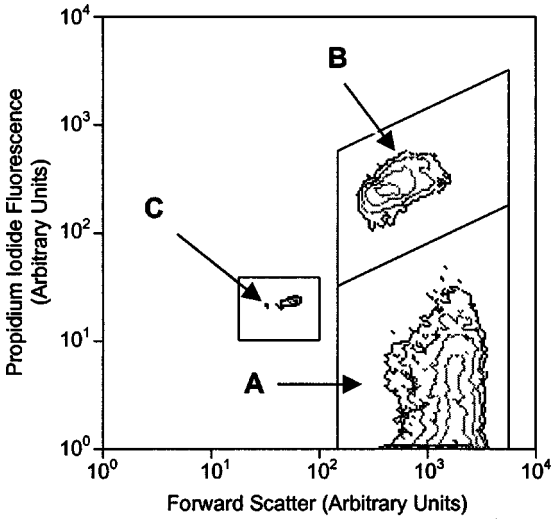


FIG. 2. Flow cytometry forward scatter versus propidium iodide (PI) fluorescence plot. (A) Viable cells display low fluorescence in the PI channel. (B) Dead cells are characterized by high PI fluorescence. (C) Fluorescent beads have low forward-scatter light signals and low fluorescence in the PI channel.

molecules delivered per cell (Prausnitz *et al.*, 1993). The raw data collected with the flow cytometer were analyzed using Windows Multiple Document Interface (WINMDI; TSRI Flow Cytometry, San Diego, CA) flow cytometry software (Canatella *et al.*, 2001).

E. Statistical analysis

At each condition tested, a minimum of three replicate data points was collected. Replicates were utilized to calculate experimental means and standard errors. One-way analysis of variance (ANOVA, $\alpha=0.05$) was performed when comparing three or more experimental conditions to a single factor. When two factors were compared, a two-way analysis of variance was used. A value of $p<0.05$ was considered statistically significant.

The dependence of experimental data on acoustic energy exposure was described mathematically by fitting to a sigmoidal curve. The fit was optimized using the Microsoft EXCEL solver function (Microsoft, Redmond, WA). The sigmoidal dependence was selected as the appropriate functionality based on the observation that the extent of ultrasound's effects on cells increased monotonically with acoustic energy above a low-energy threshold and below a high-energy plateau or saturation. A coefficient of determination (R^2) was used to quantify ''goodness'' of fit.

To identify trends in experimental data where no functionality was known or assumed, regression models based on restricted cubic splines (S-Plus, MATHSOFT, Seattle, WA) were used. ''Goodness'' of fit for each trend was determined using the multiple R^2 statistic, which represents the amount of variability in the response variable (e.g., uptake) that is explained by the fitted variable (e.g., pressure). A multiple R^2 of 1 indicates a perfect relationship between the fit and response variables, while a multiple R^2 of 0 indicates no relationship.

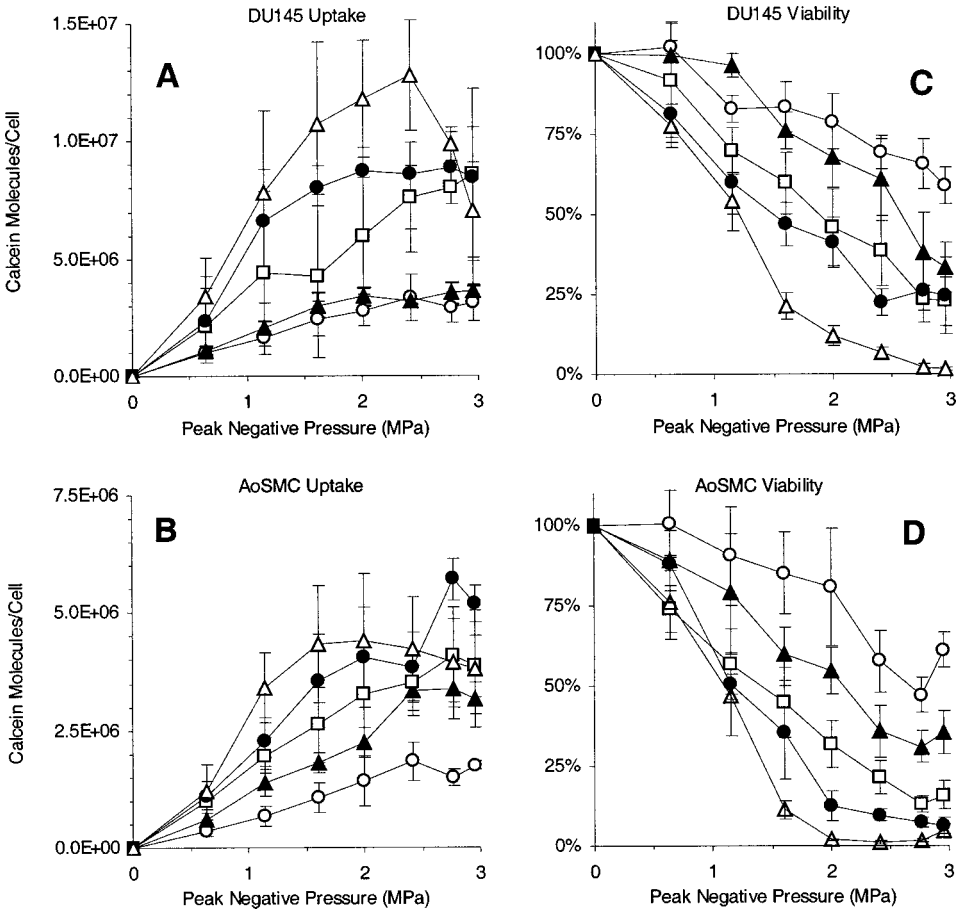


FIG. 3. Calcein uptake increased in (A) DU145 and (B) AoSMC cells with increasing pressure. Cell viability decreased in (C) DU145 and (D) AoSMC cells with increasing pressure. Total ultrasound exposure durations were: ○=120, ▲=240, □=540, ●=1000, and △=2000 ms. Pulse length was 60 ms. Data expressed as mean±SEM.

III. EXPERIMENTAL RESULTS

This study characterized the effects of acoustic pressure, total exposure time, and pulse length on human prostate cancer (DU145) and human aortic smooth muscle (AoSMC) cell suspensions. Quantitative analysis of the bioeffects (i.e., molecular uptake and cell viability) was performed using flow cytometry.

A. Acoustic pressure dependence

The effects of peak negative pressure over the range of 0.6–3.0 MPa on DU145 and AoSMC cells were studied using ultrasound exposures of different durations. As shown in Figs. 3(a) and (b), the number of calcein molecules delivered per viable cell generally increased with increasing pressure (Table I). Over the same range of acoustic conditions, cell viability generally decreased with increasing pressure [Figs. 3(c) and (d), Table I]. Together, these results show that large numbers of molecules (e.g., millions of molecules per cell) can be delivered into cells, but there can also be significant loss of cell viability.

Closer inspection of the data shows that a large (>50%) decrease in viability occurred between 1.6 and 3.0 MPa at 1- and 2-s exposure times, which suggests that high negative pressures (>1.6 MPa) and long exposure times (⩾1 s) may be less desirable for transiently disrupting viable cells. However, these conditions may be useful for applications where

TABLE I. Multiple R^2 values for restricted cubic spline fits (data from Figs. 3 and 4).

Multiple R^2 values for dependence on pressure					Multiple R^2 values for dependence on exposure time				
Exposure time (ms)[a]	Uptake		Viability		Pressure (MPa)[b]	Uptake		Viability	
	DU145	AoSMC	DU145	AoSMC		DU145	AoSMC	DU145	AoSMC
120	0.32	0.53	0.56	0.53	0.6	0.21	0.27	0.32	0.30
240	0.68	0.68	0.70	0.59	1.2	0.47	0.49	0.54	0.34
540	0.53	0.45	0.62	0.67	1.6	0.59	0.39	0.67	0.65
1000	0.44	0.62	0.53	0.93	2.0	0.73	0.26	0.63	0.80
2000	0.36	0.12	0.93	0.85	2.4	0.79	0.34	0.75	0.79
					2.8	0.64	0.44	0.66	0.90
					3.0	0.65	0.41	0.63	0.87
Average	0.47	0.48	0.67	0.71	Average	0.65	0.41	0.63	0.87

[a]The dependence of uptake and viability on pressure is determined at five different exposure times (data from Fig. 3).
[b]The dependence of uptake and viability on exposure time is determined at seven different pressures (data from Fig. 4).

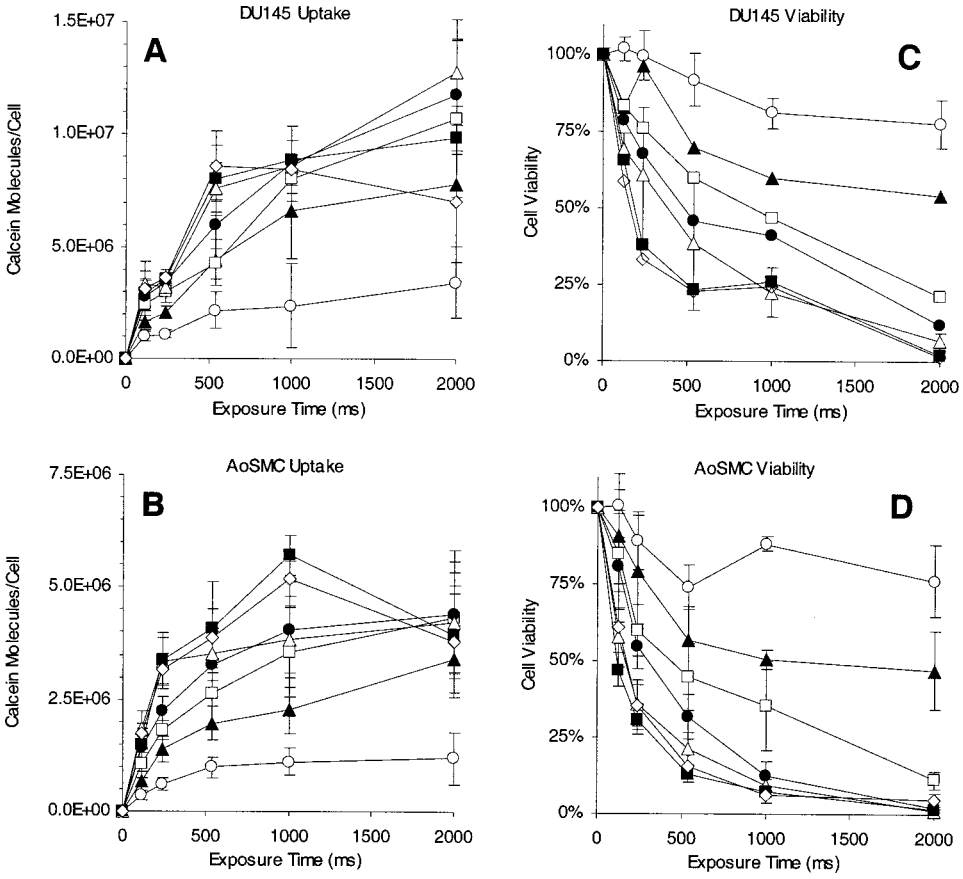


FIG. 4. Calcein uptake increased in (A) DU145 and (B) AoSMC cells with increasing ultrasound exposure duration. Cell viability decreased in (C) DU145 and (D) AoSMC cells with increasing exposure time. Peak negative acoustic pressures were: ○=0.6, ▲=1.2, □=1.6, ●=2.0, △=2.4, ■=2.8, and ◇=3.0 MPa. Pulse length was 60 ms. Data expressed as mean±SEM.

high molecular uptake accompanied by significant cell death is beneficial, such as in targeted cancer chemotherapy.

B. Exposure time dependence

To better show the effect of ultrasound exposure time on cells, the data presented in Fig. 3 are replotted as shown in Fig. 4. In this set of experiments, exposure time was varied between 120 and 2000 ms. Figures 4(a) and (b) show that calcein uptake generally increased with exposure time (Table I). Similarly, cell viability generally decreased with exposure time [Figs. 4(c) and (d), Table I].

C. Pulse length dependence

Because pulse length was held constant at 60 ms in the experiments described previously, we conducted additional experiments in which the pulse length was varied between 0.02 ms (10 cycles/pulse) and 60 ms (30 000 cycles/pulse). As shown in Fig. 5, pulse length did not have a significant effect on molecule uptake or cell viability at the conditions studied (one-way ANOVA: $p=0.97$ for uptake and $p=0.42$ for viability).

D. Acoustic energy correlation

The families of curves shown in Figs. 3–5 indicate that bioeffects exhibit a strong dependence on both exposure time and acoustic pressure, but no dependence on pulse length. This suggests that energy exposure, which scales with exposure time and the square of pressure, but is independent of pulse length [see Eq. (2)], might also correlate with bioeffects and possibly serve as a single unifying parameter that could collapse all of the data into a single curve.

To test this hypothesis, we replotted all of the uptake and viability measurements shown in Figs. 3–5 versus applied acoustic energy exposure. As shown in Fig. 6, calcein uptake and cell viability correlated with energy, where each set of data collapsed into a single curve (restricted cubic spline multiple $R^2=0.76$ and 0.78 for viability of DU145 and AoSMC cells, respectively; multiple $R^2=0.61$ and 0.39 for uptake by DU145 and AoSMC cells, respectively). The observation that bioeffects generated over a broad range of ultrasound conditions all correlate with acoustic energy gives insight into possible mechanisms and can guide development of applications, as discussed in the following sections.

The above-stated statistical analysis could be enhanced by assigning a functionality to the dependence of bioeffects on energy exposure. We noted that the shape of each curve in Fig. 6 appears to be sigmoidal: At low energy both uptake and loss of viability show weaker dependence on energy; at moderate energies the bioeffects increase as strong functions of energy; and at high energy the bioeffects level off. For this reason, we used a sigmoidal function to mathematically describe the data in Fig. 6, which yielded the following:

$$N=N_{\max}(1-e^{\alpha E^{2/3}}), \tag{3}$$

$$V=100\%(e^{\alpha E^{2/3}}), \tag{4}$$

where N is the number of calcein molecules delivered per cell, V is cell viability, E is acoustic energy exposure [J/cm^2, as defined in Eq. (2)], $N_{\max}$ is an empirically determined

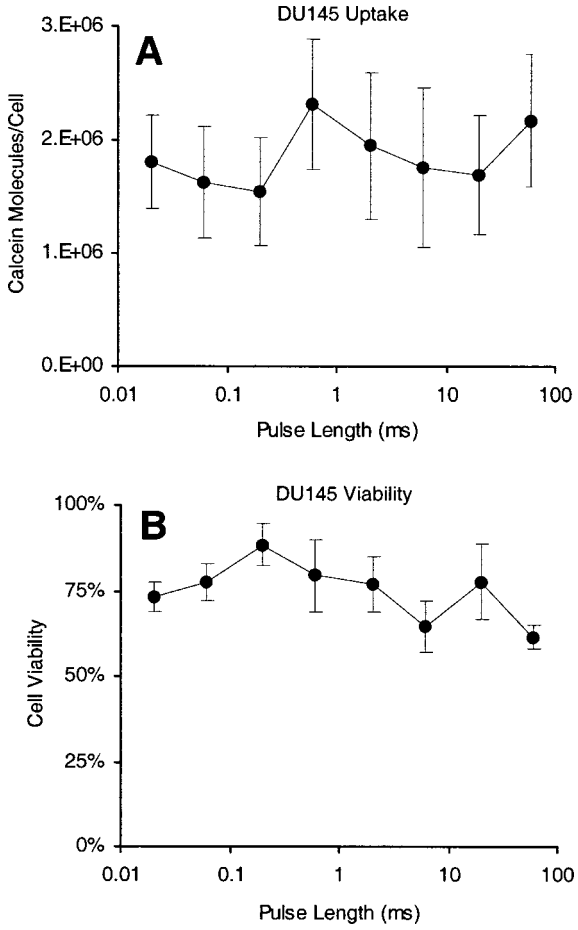


FIG. 5. Varying pulse length did not affect (A) calcein uptake or (B) cell viability. Ultrasound peak negative pressure and exposure duration were held constant at 1.6 MPa and 240 ms, respectively. Data expressed as mean±SEM.

upper limit for the number of molecules that can be delivered into each cell type, and the adjustable parameter α empirically accounts for differences in cell type ($N_{\max}$ equals 1.2×10^7 and 4.6×10^6, and α equals -0.036 and -0.057 for DU145 and AoSMC cells, respectively).

When compared to the fits generated previously using the multiple functionalities permitted by restricted cubic splines, the goodness of fit using the sigmoidal relationships provided equally robust correlations for uptake [Eq. (3), $R^2=0.59$ and 0.34 for DU145 and AoSMC, respectively] and for viability [Eq. (4), $R^2=0.76$ and 0.79 for DU145 and AoSMC, respectively], indicating that sigmoidal fits are appropriate. Interestingly, the correlations shown in Figs. 4 and 5 have the same exponent for the energy dependence for both cell types ($E^{2/3}$), suggesting similarity in mechanism. A two-way ANOVA shows significantly different levels of uptake between DU145 and AoSMC as a function of increasing energy exposure ($p<0.01$). However, differences in viability were not significant between cell types as a function of increasing energy exposure ($p=0.98$). Future studies are needed to establish if the correlation with energy observed here also exists at other frequencies, contrast agent concentrations, experimental conditions, and in other cell types.

E. Optimal exposure conditions

For practical applications such as gene transfection, optimal ultrasound conditions may be those that maximize the product of the number of molecules delivered per viable cell (N) and cell viability (V). Conditions that yield the greatest NV would produce a cell population with the most molecules in the most viable cells. Previous work has suggested that NV correlates with degree of gene transfection for many cell types (P. Canatella, personal communication).

As shown in Fig. 7, NV generally increased with energy, reached a maximum, and then decreased with further increasing energy (Multiple $R^2=0.17$ and 0.43 for DU145 and AoSMC, respectively). This suggests that NV may be optimized within a defined energy window. At low energy levels, little uptake occurred and, at high energy levels, significant cell death occurred, each resulting in low NV values. The region of interest lies in between (i.e., combinations of low pressure-long exposures, high pressure-short exposures, and moderate pressure-medium duration exposures). Figure 7 shows that optimal NV was at approximately the same energy exposure for both cell lines (i.e., $\sim$50 J/cm^2).

IV. DISCUSSION

A. Correlation with energy exposure

In this study, simultaneous measurements of molecular uptake and cell viability were performed on DU145 and AoSMC cells over a wide range of acoustic parameters. In support of our proposed hypothesis, bioeffects generally correlated with acoustic energy exposure regardless of the acoustic pressure, exposure time, and pulse length used. A similar finding was made by Mitragotri *et al.* (2000), who observed that acoustic energy exposure correlated with increased skin conductivity induced by exposure to 20-kHz ultrasound.

The dependence of bioffects on energy exposure may be of use to suggest or validate mechanistic understanding, but is also relevant to practical applications. A dependence on energy gives an experimentalist or clinician considerable freedom to use different ultrasound conditions that produce the same energy. For example, one could use long, low-pressure exposures which require a less expensive transducer and power supply or similarly use short, high-pressure pulses of the same energy if extremely rapid (i.e., subsecond) effects are desirable. In addition, the correlation of acoustic energy with NV, which is believed to be a marker of transfection, can guide optimization of gene delivery protocols. Further study is needed to determine if these results can be extended to tissue and the *in vivo* environment.

B. Maximum levels of uptake

Over the full range of acoustic conditions studied, average uptake never exceeded 1.2×10^7 molecules per DU145 cell or 4.6×10^6 molecules per AoSMC cell. Based on average cell volumes of 2200 and 2400 μm^3 for DU145 and AoSMC cells, respectively (determined using a Coulter Counter), this corresponds to intracellular concentrations of approximately 9.1 μM in DU145 and 3.2 μM in AoSMC

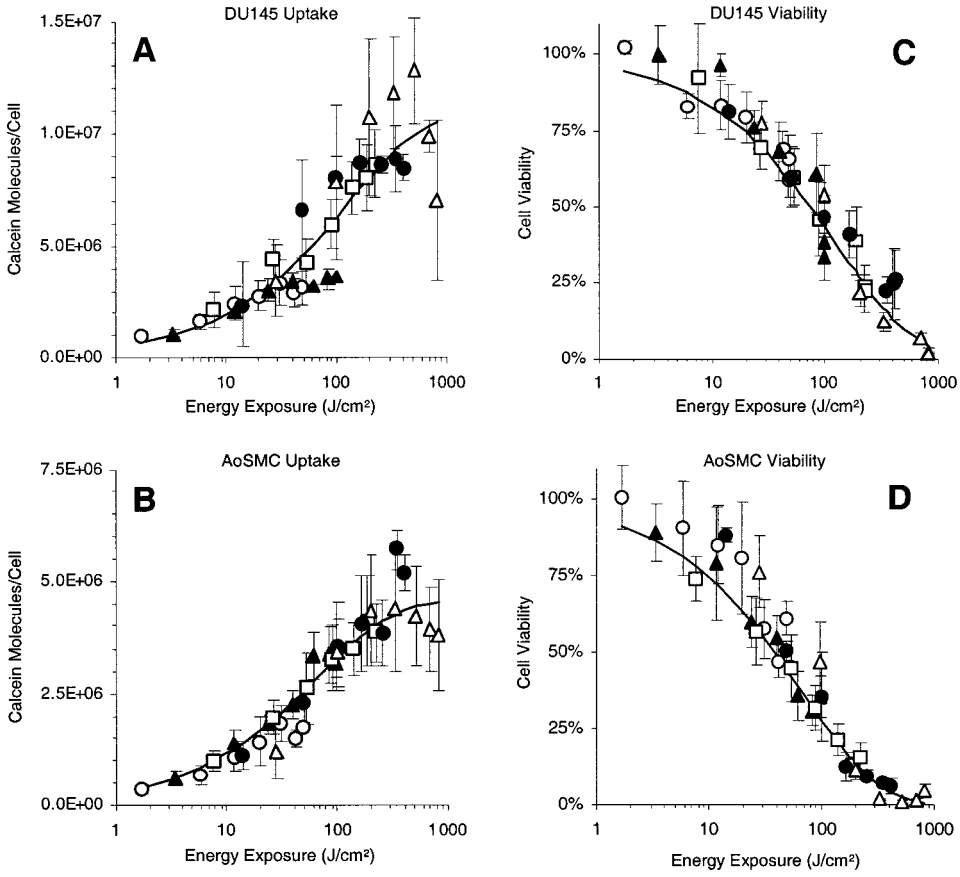


FIG. 6. Calcein uptake increased in (A) DU145 and (B) AoSMC cells with increasing acoustic energy exposure. Cell viability decreased in (C) DU145 and (D) AoSMC cells with increasing acoustic energy exposure. Ultrasound exposure durations were: ○ = 120, ▲ = 240, □ = 540, ● = 1000, and △ = 2000 ms (data are from Figs. 3 and 4). Pulse length data (i.e., from Fig. 5) are represented by ◇. Sigmoidal data fits are given by the solid lines [Eqs. (3) and (4)]. Data expressed as mean ± SEM.

cells. Since the extracellular concentration was 10 μM, this indicates that approximate thermodynamic equilibrium with the extracellular solution was achieved in DU145 cells, although associated with large loss in viability. This finding contrasts with results for molecular uptake by electroporation, where ''subequilibrium'' uptake was always observed (Prausnitz *et al.*, 1993; Canatella *et al.*, 2001).

C. Cavitation-based mechanism

The observed dependence of bioeffects on energy suggests the possibility of a thermal mechanism of action. However, our data in combination with insight from the literature indicate that cavitation, rather than heat, is responsible for the observed effects. In this study, all experiments were performed with contrast agent (CA) present in the cell sample. CA is known to nucleate cavitation and thereby lower the pressure required to generate and sustain bubble activity (Holland and Apfel, 1990; Miller and Thomas, 1995). CA is not expected to influence thermal effects. As shown in Fig. 8, cell samples lacking CA experienced minimal loss of cell viability (one-way ANOVA $p = 0.97$) (and insignificant uptake of molecules; data not shown) as opposed to cells exposed to ultrasound in the presence of CA (one-way ANOVA $p < 0.001$).

This observation verified our *a priori* assumption about cavitation-induced bioeffects and is in agreement with others who have shown that CA can significantly enhance bioeffects produced by ultrasound (Miller *et al.*, 1997; Greenleaf *et al.*, 1998; Poliachik *et al.*, 1999). As described in Sec. II, the possible role of thermal effects was further investigated by rapidly increasing cell sample temperature by 5 °C. This rise in temperature produced no statistically significant bioeffect on cells. When exposed to ultrasound, the temperature rise in the cell samples was always less than 1 °C. This lack of significant bulk heating and the need for CA indicates that cavitation mediated the observed bioeffects.

D. Lack of dependence on pulse length

Figure 4 showed that bioeffects were independent of pulse length over the range of conditions tested (10–30 000 cycles/pulse), in agreement with previous work (Brayman and Miller, 1999). However, this observation differs from other studies in which the degree of red blood cell hemolysis was shown to depend on pulse length (Ciavarino *et al.*, 1981; Kober and Ellwant, 1989; Liu *et al.*, 1998). Recent work performed in our lab using DU145 cells also demonstrated a dependence of bioeffects of pulse length (Cochran and Prausnitz, 2001). A possible explanation for this discrepancy may involve the presence of CA microbubbles; those studies that observed a pulse length dependence did not have CA present during insonification, while those that observed no dependence had CA in their samples. In the absence of CA, it may take a minimum amount of time or number of acoustic cycles within a pulse to nucleate and grow bubbles, suggesting that a minimum pulse length is needed. In contrast, Flynn and Church (1988) noted that cavitation could occur after just one acoustic cycle if appropriately sized nuclei are supplied, such as those provided by CA. Thus, if cavitation

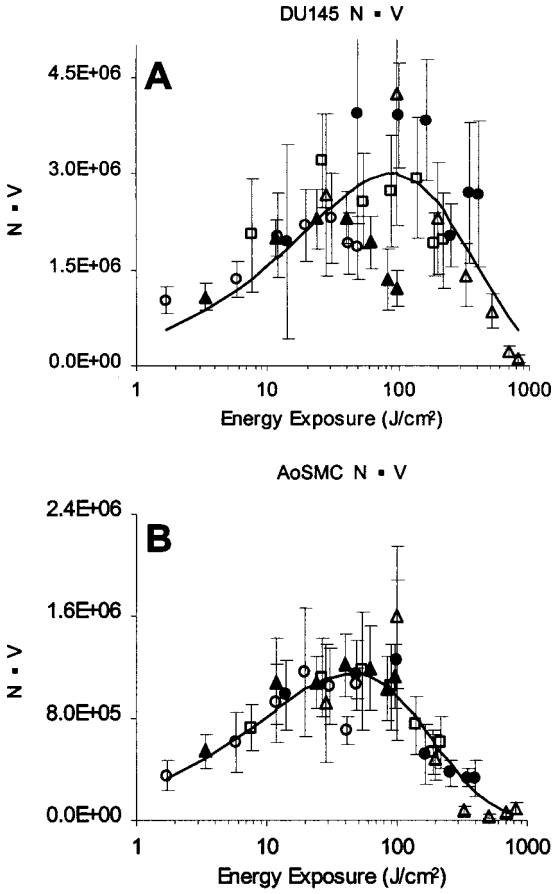


FIG. 7. The $N\ V$ product (N = number of molecules delivered per cell, V = cell viability) increased, reached a maximum, and then decreased with increasing energy exposure in both (a) DU145 and (b) AoSMC cells. Ultrasound exposure durations were: ○ = 120, ▲ = 240, □ = 540, ● = 1000, and △ = 2000 ms. Data fits (solid line) were obtained from the product of Eqs. (3) and (4). Data expressed as mean ± SEM.

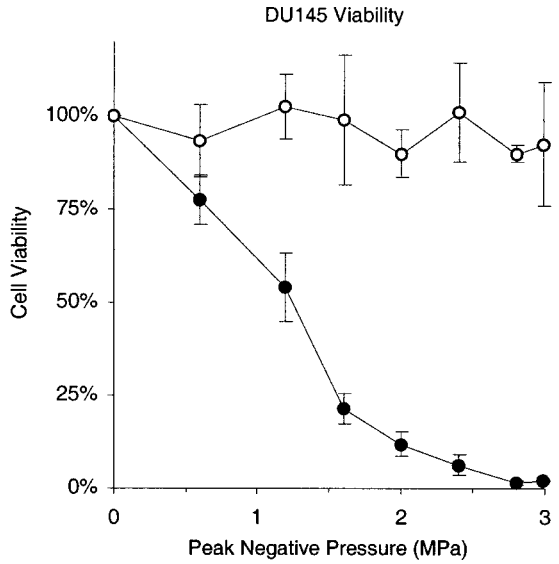


FIG. 8. The presence of 1.7% (v/v) contrast agent (●) during insonation caused viability to decrease with increasing pressure at 2-s total exposure time and 60-ms pulse length. No effect was observed under the same ultrasound conditions in the absence of contrast agent (○). Data expressed as mean ± SEM.

can be induced at the onset of ultrasound exposure through the use of CA, then pulse length appears not to be an important parameter.

V. CONCLUSIONS

By exposing DU145 and AoSMC cells to a broad range of ultrasound conditions, we determined that ultrasound can deliver millions of calcein molecules into viable cells at an intracellular concentration that approaches the extracellular concentration. In both cell types, uptake increased and viability decreased with increasing ultrasound pressure and exposure time; pulse length, however, did not influence bioeffects in the samples, which were preseeded with cavitation nuclei. All of the data collapsed into a single curve when plotted as a function of acoustic energy exposure, suggesting that this may be predictive of ultrasound's nonthermal bioeffects. Based on a criterion of maximizing the product of molecular uptake and cell viability, optimal ultrasound conditions were identified on the basis of an optimal acoustic energy exposure. Because contrast agent was needed to observe bioeffects and heating was negligible, we propose that cavitation mediated by acoustic energy exposure is the mechanism of ultrasound's effects on cells observed in this study.

ACKNOWLEDGMENTS

We would like to thank Dr. Thomas N. Lewis, Dr. Paul J. Canatella, Dr. Keyvan Keyhani, Steve I. Woodard, and Dr. Russel Heikes for their generous help and critical comments. This work was supported in part by The Whitaker Foundation, National Science Foundation, and National Institutes of Health.

Bao, S., Thrall, B. D., and Miller, D. L. (**1997**). ''Transfection of a reporter plasmid into cultured cells by sonoporation *in vitro*,'' Ultrasound Med. Biol. **23**, 953–959.

Brayman, A. A., and Miller, M. W. (**1999**). ''Sonolysis of Albunex-supplemented, 40% hematocrit human erythrocytes by pulsed 1-MHz ultrasound: Pulse number, pulse duration and exposure vessel rotation dependence,'' Ultrasound Med. Biol. **25**, 307–314.

Canatella, P. J., Karr, J. F., Petros, J. A., and Prausnitz, M. R. (**2001**). ''Quantitative study of electroporation-mediated molecular uptake and cell viability,'' Biophys. J. **80**, 755–764.

Canatella, P. J. (personal communication).

Carstensen, E. L., Kelly, P., Church, C. C., Brayman, A. A., Child, S. Z., Raeman, C. H., and Schery, L. (**1993**). ''Lysis of ertythrocytes by exposure to cw ultrasound,'' Ultrasound Med. Biol. **19**, 147–165.

Chang, D. C., Chassy, B. M., Saunders, J. A., and Sowers, A. E. (**1992**). *Guide to Electroporation and Electrofusion* (Academic Press, New York).

Ciaravino, V., Flynn, H. G., and Miller, M. W. (eds.) (**1981**). ''Pulsed enhancement of acoustic cavitation: A postulated model,'' Ultrasound Med. Biol. **7**, 159–166.

Clarke, M. S., and McNeil, P. L. (**1992**). ''Syringe loading introduces macromolecules into living mammalian cell cytosol,'' J. Cell. Sci. **102**, 533–41.

Cochran, S., and Prausnitz, M. R. (2001). ''Sonoluminescence as an indicator of cell membrane disruption by acoustic cavitation,'' Ultrasound Med. Biol. **27**, 841–850.

Fechheimer, M., Denny, C., Murphy, R. F., and Taylor, D. L. (**1986**). ''Measurement of cytoplasmic *p*H in dictyostelium discoideum by using a new method for introducing macromolecules into living cells,'' Eur. J. Cell Biol. **40**, 242–247.

Flynn, H. G., and Church, C. C. (**1988**). ''Transient pulsations of small gas bubbles in water,'' J. Acoust. Soc. Am. **84**, 1863–1876.

Greenleaf, W. J., Bolander, M. E., Sarkar, G., Goldring, M. B., and Greenleaf, J. F. (**1998**). ''Artificial cavitation nuclei significantly enhance acoustically induced cell transfection,'' Ultrasound Med. Biol. **24**, 587–595.

Harrison, G. H., Balcer-Kubiczek, E. K., and Gutierrez, P. L. (**1996**). ''In vitro mechanisms of chemopotentiation by tone-burst ultrasound,'' Ultrasound Med. Biol. **22**, 355–362.

Holland, C. K., and Apfel, R. E. (**1990**). ''Threshold for transient cavitation produced by pulsed ultrasound in a controlled nuclei environment,'' J. Acoust. Soc. Am. **88**, 2059–2069.

Holmes, R. P., Yeaman, L. F., Taylor, R. G., and McCullough, D. L. (**1992**). ''Altered neutrophil permeability following shock wave exposure *in vitro*,'' J. Urol. (Baltimore) **147**, 733–737.

Kaiser, A. R., Cain, C. A., Hwang, E. Y., Fowlkes, J. B., and Jeffers, R. J. (**1996**). ''A cost effective degassing system for use in ultrasonic measurements: The multiple pinhole degassing system,'' J. Acoust. Soc. Am. **99**, 3857–3859.

Keyhani, K., Parsons, A., Siddeeq, D., Lewis, T. N., and Prausnitz, M. R. (**1998**). ''Size, lifetime, and permeability of ultrasound-mediated cell membrane disruptions,'' Proceedings of the Internationl Symposium on Controlled Release of Bioactive Materials **25**, 695–696.

Kober, L. O., and Ellwart, J. W. (**1989**). ''Effect of the pulse length of ultrasound on cell membrane damage *in vitro*,'' J. Acoust. Soc. Am. **86**, 6–7.

Liu, J., Lewis, T. N., and Prausnitz, M. R. (**1998**). ''Non-invasive assessment and control of ultrasound-mediated membrane permeabilization,'' Pharm. Res. **15**, 918–923.

McNeil, P. L. (**1989**). ''Incorporation of macromolecules into living cells,'' Methods Cell Biol. **29**, 153–173.

Miller, D. L., Bao, S., and Morris, J. E. (**1999**). ''Sonoporation of cultured cells in the rotating tube exposure tube,'' Ultrasound Med. Biol. **25**, 143–149.

Miller, D. L., Gies, R. A., and Chrisler, W. B. (**1997**). ''Ultrasonically induced hemolysis at high cell and gas body concentrations in a thin-disc exposure chamber,'' Ultrasound Med. Biol. **23**, 625–633.

Miller, D. L., and Thomas, R. M. (**1995**). ''Ultrasound contrast agents nucleate inertial cavitation *in vitro*,'' Ultrasound Med. Biol. **21**, 1059–1065.

Miller, M. W., Miller, D. L., and Brayman, A. A. (**1996**). ''A review of *in vitro* bioeffects of inertial ultrasonic cavitation from a mechanistic perspective,'' Ultrasound Med. Biol. **22**, 1131–54.

Mitragotri, S., Farrell, J., Tang, H., Terahara, T., Kost, J., and Langer, R. (**2000**). ''Determination of threshold energy dose for ultrasound-induced transdermal drug transport,'' J. Controlled Release **63**, 41–52.

Poliachik, S. L., Chandler, W. L., Mourad, P. D., Bailey, M. R., Bloch, S., Cleveland, R. O., Kaczkowski, P., Keilman, G., Porter, T., and Crum, L. A. (**1999**). ''Effect of high-intensity focused ultrasound on whole blood with and without microbubble contrast agent,'' Ultrasound Med. Biol. **25**, 991–998.

Prat, F., Chapelon, J. Y., El Fadil, F. A., Theillere, Y., Ponchon, T., and Cathignol, D. (**1993**). ''*In vivo* effects of cavitation alone or in combination with chemotherapy in a perotoneal carcinomatosis in the rat,'' Br. J. Cancer **68**, 13–17.

Prausnitz, M. R., Lau, B. S., Milano, C. D., Conner, S., Langer, R., and Weaver, J. C. (**1993**). ''A quantitative study of electroporation showing a plateau in net molecular transport,'' Biophys. J. **65**, 414–422.

Saad, A. H., and Hahn, G. M. (**1987**). ''Ultrasound enhances Adriamycin toxicity *in vitro*,'' in *Heat Transfer in Bioengineering and Medicine*, edited by J. C. Chato, T. E. Diller, K. R. Killer, and R. B. Roemer (American Society of Mechanical Engineers Press, New York).

Ultrasound-mediated disruption of cell membranes. II. Heterogeneous effects on cells

Héctor R. Guzmán, Daniel X. Nguyen, Sohail Khan, and Mark R. Prausnitz[a)]
School of Chemical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0100

(Received 18 August 2000; accepted for publication 4 April 2001)

Ultrasound has been shown to reversibly and irreversibly disrupt membranes of viable cells through a mechanism believed to involve cavitation. Because cavitation is both temporally and spatially heterogeneous, flow cytometry was used to identify and quantify heterogeneity in the effects of ultrasound on molecular uptake and cell viability on a cell-by-cell basis for suspensions of DU145 prostate cancer and aortic smooth muscle cells exposed to varying peak negative acoustic pressures (0.6–3.0 MPa), exposure times (120–2000 ms), and pulse lengths (0.02–60 ms) in the presence of Optison (1.7% v/v) contrast agent. Cell-to-cell heterogeneity was observed at all conditions studied and was classified into three subpopulations: nominal uptake (NUP), low uptake (LUP), and high uptake (HUP) populations. The average number of molecules within each subpopulation was generally constant: 10^4-10^5 molecules/cell in NUP, $\sim 10^6$ molecules/cell in LUP, and $\sim 10^7$ molecules/cell in HUP. However, the fraction of cells within each subpopulation showed a strong dependence on both acoustic pressure and exposure time. Varying pulse length produced no significant effect. The distribution of cells among the three subpopulations correlated with acoustic energy exposure, which suggests that energy exposure may govern the ability of ultrasound to induce bioeffects by a nonthermal mechanism. [DOI: 10.1121/1.1376130]

PACS numbers: 43.80.Gx, 43.80.Sh [FD]

I. INTRODUCTION

Drug delivery and gene therapy are limited by the need to deliver large numbers of molecules into living cells. As a possible solution, ultrasound has been shown to enhance molecular transport across cell membranes through a mechanism believed to involve acoustic cavitation. The ability to reversibly increase cell membrane permeability has been observed in studies using small molecules, macromolecules, and genetic material (Fechheimer *et al.*, 1986; Saad and Hahn, 1987; Harrison *et al.*, 1996; Bao *et al.*, 1997; Greenleaf *et al.*, 1998; Miller *et al.*, 1999; Guzmán *et al.*, 2001). However, these and other studies have generally not quantified molecular uptake within affected cells on a cell-by-cell basis. This type of analysis would provide absolute numbers of molecules within cells, which is important for drug and gene delivery applications and to support quantitative modeling efforts. It would also identify and quantify heterogeneity in ultrasound's effects, which is important for applications where uniform responses among a population of cells may be desirable.

In this study, molecular uptake and cell viability are quantified on a cell-by-cell basis for large numbers of individual cells (e.g., 20 000 cells per sample) using flow cytometry. We hypothesize that within a population of cells sonicated under the same conditions, the number of molecules delivered into each cell will be highly variable. To test this hypothesis, data from a companion study (Guzmán *et al.*, 2001) are reanalyzed to quantify levels of molecular uptake on a cell-by-cell basis. This re-analysis permits identification of distributions in uptake and provides a means to quantify any observed heterogeneity.

Our companion study (Guzmán *et al.*, 2001), which measured the effects of acoustic pressure, exposure time, and pulse length on cell viability and average molecular uptake within cells, concluded that these bioeffects correlated with acoustic energy exposure over the conditions tested. We therefore propose as a second hypothesis for the present study that cell-to-cell heterogeneity in molecular uptake resulting from sonication will correlate with acoustic energy exposure.

It is not clear *a priori* that cell-to-cell heterogeneity should be expected, which in part motivates this study. For example, molecular uptake induced by an electrical method to increase cell membrane permeability—electroporation—has been shown to be homogeneous over a broad range of experimental conditions in mammalian cells (Prausnitz *et al.*, 1993; Canatella *et al.*, 2001) and can show bimodal distributions in yeast (Gift and Weaver, 1995). The observed homogeneity can be attributed to the uniform electric field experienced by cells during electroporation, whereas the heterogeneity may be due to the nonspherical shape of yeast cells that results in different transmembrane voltages as a function of cell orientation. In contrast, ultrasound's bioeffects are generally attributed to cavitation, which is heterogeneous in both time and space (Leighton, 1994; Miller *et al.*, 1996; Barnett, 1998). For this reason, we have hypothesized that each cell will experience different interactions with cavitation bubbles, which thereby result in different levels of bioeffects on a cell-to-cell basis.

a)Electronic mail: mark.prausnitz@chc.gatech.edu

II. MATERIALS AND METHODS

To measure possible heterogeneity in the effects of ultrasound on cells, we used data collected in a companion study (Guzmán *et al.*, 2001) which quantified molecular uptake and cell viability over a range of peak negative acoustic pressures (0.6–3.0 MPa), exposure times (120–2000 ms), and pulse lengths (0.02–60 ms). Detailed descriptions of cell culture, experimental procedures, ultrasound application, and data analysis are presented in that study. The following is a summary of those experimental methods, as well as a detailed description of additional analysis used in this study.

A. Experimental methods

DU145 prostate cancer cells (DU145) or human aortic smooth muscle cells (AoSMC) grown as monolayers were harvested and resuspended to a concentration of 1×10^6 cells/ml in Dulbecco's phosphate buffered saline. Calcein (623 Da, radius$=0.6$ nm), a molecule simulating a small drug, was added at a concentration of 10 μM and Optison contrast agent, which provides nuclei for cavitation, was added to achieve a final concentration of 1.7% v/v. Cell samples were then exposed to ultrasound at the conditions described previously using focused 500-kHz ultrasound in the apparatus described previously (Guzmán *et al.*, 2001). Using flow cytometry, molecular uptake caused by ultrasound was determined by measuring the intensity of green fluorescence emitted by calcein taken up by cells. Cell viability was determined by measuring the intensity of red fluorescence emitted by propidium iodide added as a viability stain to cell samples after sonication (Guzmán *et al.*, 2001).

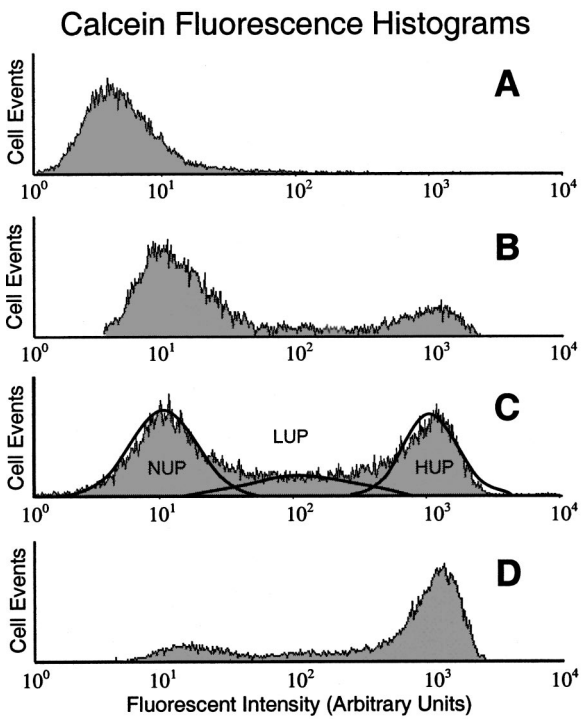


FIG. 1. Fluorescence histograms of cell samples showing uptake of calcein (20 000 cells/sample). (A) Fluorescence of a control sample shows Gaussian-distributed background fluorescence in the first decade of the histogram. (C) Fluorescence signal in the exposed sample is heterogeneous and can be divided into regions termed nominal (NUP), low (LUP), and high (HUP) uptake populations (60-ms pulse length, 540-ms exposure time, and 3.0 MPa peak negative pressure). (B, D) Histograms showing that at other ultrasound conditions there still exist three subpopulations, but with different relative numbers of cells in each sub-population (60-ms pulse length, 2-s exposure time, and 1.2 (B) and 2.4 (D) MPa peak negative pressure).

B. Flow cytometry analysis of heterogeneity

Intracellular calcein fluorescence was measured using flow cytometry; 20 000 cell measurements were collected per sample to ensure that a statistically significant cell population was analyzed. As shown in Fig. 1, the fluorescence intensity of each sample of cells can be presented as a histogram. In a representative control sample of cells unexposed to ultrasound [Fig. 1(a)], a single population of cells is present with a distribution of fluorescence most likely due to autofluorescence, optical and electrical noise, and low-level surface binding of calcein. In representative samples of cells exposed to ultrasound [Figs. 1(b)–(d)], the histograms show broad, heterogeneous distributions of cell fluorescence, which appear to contain multiple subpopulations. The shape observed in these histograms, i.e., two peaks and a wide valley in between, was common to almost all cell samples exposed to ultrasound. For this reason, the observed heterogeneity was analyzed using three subpopulations of cells, as shown pictorially in Fig. 1(c).

To analyze the distributions of viable cells contained within each sample, raw data from histograms generated in WINMDI (TSRI Flow Cytometry, San Diego, CA), were exported into EXCEL (Microsoft, Redmond, WA) as ASCII files using the utility LDATA (Robinson and Kelly, 1998). In EXCEL, the data were formatted and then exported into MIX Software 3.1 (Ichthus Data Systems, Hamilton, Ontario, Canada). MIX is a statistical software package used to analyze populations containing a mixture of subpopulations. Using MIX, we found the best fit for a set of three normal distributions to describe the uptake histogram for each sample exposed to ultrasound [Fig. 1(c)]. MIX then calculated the fraction of cells, the fluorescence mean, and the standard deviation of the fluorescence mean for each of the three cell subpopulations. Molecular uptake in each subpopulation was measured by subtracting the mean fluorescence of control samples [e.g., Fig. 1(a)] from the mean fluorescence of each subpopulation in each of the exposed samples. The number of calcein molecules delivered per cell was then determined from these mean fluorescence measurements, as described previously (Guzmán *et al.*, 2001).

C. Visual verification of heterogeneity

To visualize molecular uptake into cells, a Zeiss LSM510 confocal microscope (Carl Zeiss, Thornwood, NY) was used to image the fluorescence emitted from cells exposed to 488-nm argon UV lasers (optical section at ~8-μm penetration depth, which is near the center of each cell). Green fluorescence (calcein) indicated molecular uptake, red fluorescence (propidium iodide) stained the nuclei of dead cells, and blue fluorescence (Hoechst 33342; Molecular Probes, catalog no. H-1399) nonspecifically identified the

nuclei of all cells. Confocal images were used to visually corroborate the heterogeneity observed in flow cytometry data.

D. Statistical analysis

At each condition tested, a minimum of three replicate cell samples was measured, from which the mean and standard error were calculated. A Student's t-test was used when comparing two experimental conditions and a one-way analysis of variance (ANOVA $\alpha = 0.05$) was performed when comparing one factor with three or more experimental conditions. A p value <0.05 was considered statistically significant. Data are expressed as mean $\pm$ SEM in the Figures.

To identify trends in experimental data, regression models based on restricted cubic splines (S-Plus, MATHSOFT, Seattle, WA) were used. ''Goodness'' of fit for each trend was determined using the multiple R^2 statistic, which represents the amount of variability in the response variable (e.g., uptake) that is explained by the fitted variable (e.g., pressure). A multiple R^2 of 1 indicates a perfect relationship between the fit and response variables, while a multiple R^2 of 0 indicates no relationship.

III. EXPERIMENTAL RESULTS

In a companion study (Guzmán *et al.*, 2001), we measured the effects of acoustic pressure, exposure time, and pulse length on the uptake of a model compound, calcein, and the loss of cell viability in DU145 prostate cancer and aortic smooth muscle cell suspensions in the presence of Optison contrast agent. As is commonly done, each cell sample was treated as a single homogeneous population, and uptake and viability were expressed as overall average values representative of each cell sample. However, closer examination of the data shows that cell samples are heterogeneous and therefore should be described with multiple subpopulations (Fig. 1). Therefore, in this study the data were reanalyzed to account for the observed heterogeneity.

A. Cell heterogeneity

Heterogeneous bioeffects were observed in both DU145 and AoSMC cell samples at almost all of the ultrasound conditions tested. Figure 1 contains histograms of calcein fluorescence associated with viable cells (cells rendered nonviable by ultrasound are discussed further in the following sections). As shown in Fig. 1(a), the fluorescence emitted from a representative control sample of viable cells is distributed across a single population having low fluorescence. Figures 1(b)–(d) show a representative set of samples exposed to ultrasound. The cells in these samples have fluorescence (i.e., uptake) corresponding to one of three regions: (1) a low-fluorescence peak, (2) a high-fluorescence peak, or (3) a wide valley in between. All samples exposed to ultrasound showed this distribution containing two peaks and a wide valley. Only the relative heights of the peaks varied among samples [Figs. 1(b)–(d)].

The low-fluorescence peak of Figs. 1(b)–(d) contains a subpopulation of cells with fluorescence similar to that observed in control samples and is termed the nominal uptake population (NUP). Although NUP cells have fluorescence somewhat higher than that of control cells, this level of fluorescence is just above the detection limit of the fluorescence measurement and may correspond to low-level binding of calcein to cell membranes following ultrasound exposure. Because fluorescence emitted by these cells was so dim, it was not possible to visualize by microscopy whether the fluorescence was associated with the cell membrane or cytosol.

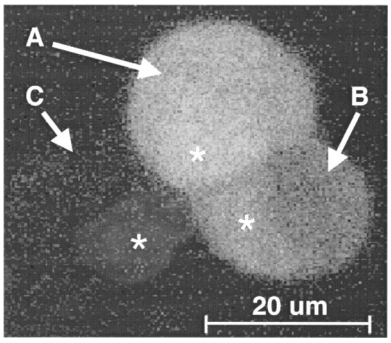


FIG. 2. A confocal microscopy image of three adjacent cells shows the simultaneous presence of three calcein uptake subpopulations. (A) The brightly fluorescent cell is indicative of high uptake of green-fluorescent calcein (HUP), (B) the dimmer fluorescent cell is indicative of low uptake (LUP), and (C) the dark cell is indicative of nominal uptake (NUP). Cell nuclei are identified by an asterisk. In this image, fluorescence is due to the combined signals from calcein uptake and Hoescht 33342 nuclear stain. The NUP cell nucleus is observed by the faint glow given off by the Hoescht dye. In the LUP and HUP cells, the nuclei are indicated by the densely bright signal that results from combined calcein and Hoescht fluorescence.

The second and third subpopulations in Figs. 1(b)–(d) contain cells with higher levels of uptake. Those cells in the broad valley are defined as the low uptake population (LUP), while those in the highly fluorescent peak are defined as the high uptake population (HUP). Although, the histograms in Fig. 1 might also be described using just two populations, for example, with non-Gaussian distributions, we felt that three populations represent the data better, since the broad range of fluorescence found in the ''valleys'' warrants its own descriptor (i.e., LUP). The three-subpopulation classification provides a means to describe data concisely, that would otherwise be difficult to describe, using the simplest fit that does not leave out important information. This categorization into three subpopulations is based solely on phenomenological observation and is not based on theoretical or mechanistic considerations.

As further evidence that cells exposed to the same ultrasound conditions can respond with highly heterogeneous amounts of uptake, confocal microscopy was used to image calcein fluorescence within viable cells. Figure 2 shows three adjacent cells that experienced the same ultrasound exposure, but have very different fluorescence intensities. The cell with almost no visible fluorescence is representative of NUP, the cell with brighter fluorescence represents LUP, and the cell with the brightest fluorescence represents HUP.

B. Molecular uptake within each subpopulation

The average number of calcein molecules taken up by viable cells within each of the three subpopulations of in-

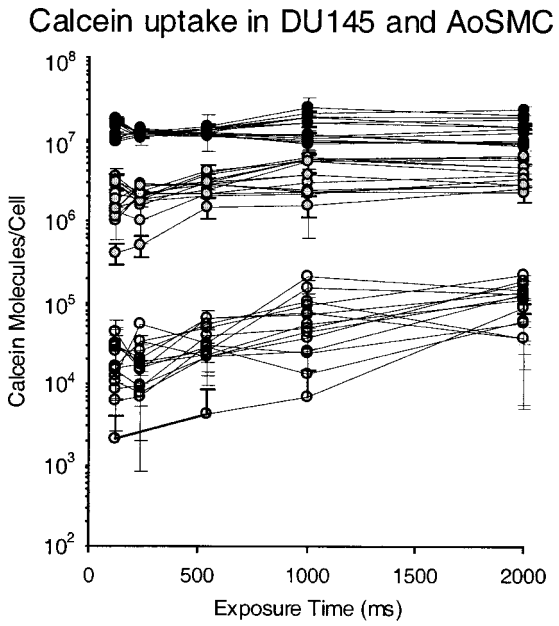


FIG. 3. The number of calcein molecules delivered per cell in each subpopulation: NUP (white), LUP (gray), and HUP (black). Ultrasound peak negative pressures were: ○=0.6, □=1.2, △=1.6, ◇=2.0, ×=2.4, ⋆=2.8, and +=3.0 MPa. Points are connected with solid lines (DU145) or dashed lines (AoSMC).

sonated cells was determined using flow cytometry. As shown in Fig. 3, calcein uptake was significantly different for each of the three subpopulations (Student's t-test, $p<0.001$). On average, NUP cells took up $4.8(\pm 5.9)\times 10^4$ molecules/DU145 cell and $5.7(\pm 7.8)\times 10^4$ molecules/AoSMC cell, LUP cells took up $3.0(\pm 1.9)\times 10^6$ molecules/DU145 cell and $2.2(\pm 1.2)\times 10^6$ molecules/AoSMC cell, and HUP cells took up $1.36(\pm 0.5)\times 10^7$ molecules/DU145 cell and $1.1(\pm 0.3)\times 10^7$ molecules/AoSMC cell (mean±standard error). Because the fraction of cells within each subpopulation varied substantially (as discussed in the following section), overall uptake for a total population of cells (i.e., the sum of uptake from all three subpopulations weighted by the fraction of viable cells in each subpopulation) exhibits a strong dependence on ultrasound exposure conditions, as shown in our companion study (Guzmán *et al.*, 2001).

C. Pressure dependence of subpopulation distribution among viable cells

Having established that within each subpopulation the number of molecules per cell showed small variation (Fig. 3), we wanted to determine the effects of ultrasound exposure conditions on the distribution of cells between the three subpopulations. Figure 4 shows the effect of pressure on this distribution among the viable cells. The NUP fraction of viable cells (NUP_{viable}; white bars) generally decreased with increasing pressure (statistical analysis provided in Fig. 4). In contrast, the LUP fraction of viable cells (LUP_{viable}; gray bars) remained relatively constant, and the HUP fraction of viable cells (HUP_{viable}; black bars) generally increased with increasing acoustic pressure, indicating that the composition of viable cells became richer in HUP at higher pressures.

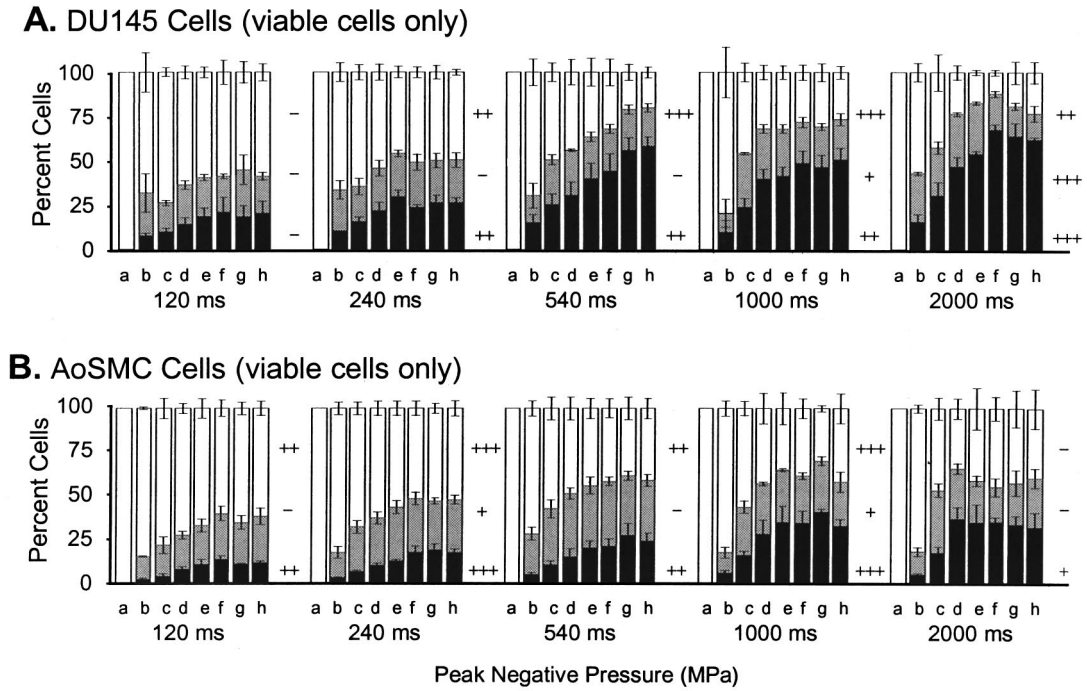


FIG. 4. Subpopulation distribution among viable cells as a function of pressure at different exposure times. The white bars indicate cells in NUP, gray bars indicate LUP, and black bars indicate HUP. The fraction of cells in NUP decreased, the fraction of cells in LUP varied little, and the fraction of cells in HUP increased with increasing peak negative pressure in both (A) DU145 and (B) AoSMC cell samples. Ultrasound peak negative pressures were: $a=0$, $b=0.6$, $c=1.2$, $d=1.6$, $e=2.0$, $f=2.4$, $g=2.8$, and $h=3.0$ MPa. Pulse length was held constant at 60 ms. One-way ANOVA indicates statistical significance of dependence on acoustic pressure for the fraction of cells in each sub-population: $-$ ($p>0.05$), + ($p<0.05$), ++ ($p<0.01$), +++ ($p<0.001$), for NUP (upper), LUP (middle), and HUP (lower) subpopulations.

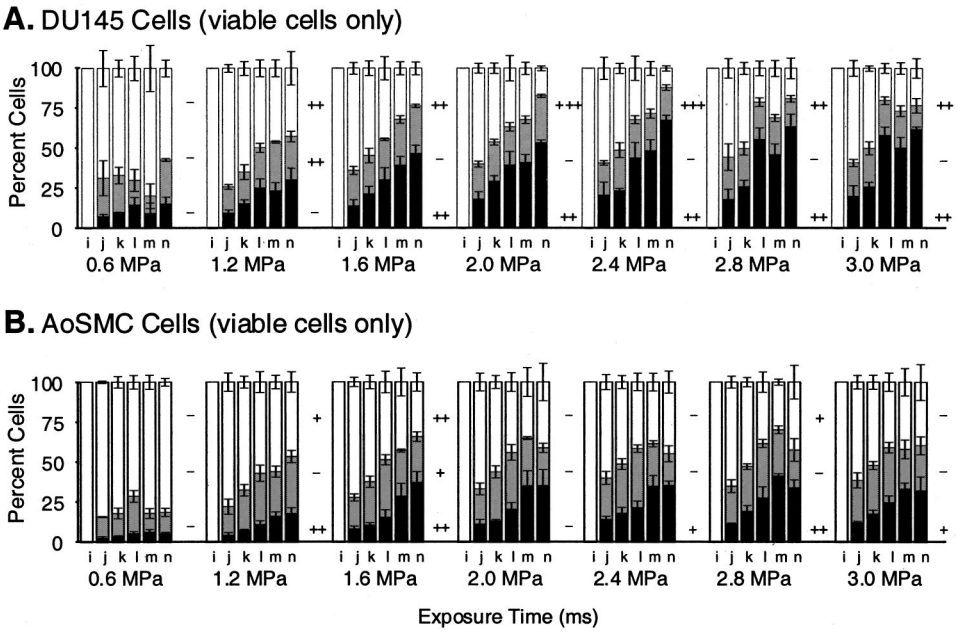


FIG. 5. Subpopulation distribution among viable cells as a function of exposure time at different pressures. The white bars indicate cells in NUP, gray bars indicate LUP, and black bars indicate HUP. The fraction of cells in NUP decreased, the fraction of cells in LUP varied little, and the fraction of cells in HUP increased with increasing exposure time in both (A) DU145 and (B) AoSMC cell samples. Ultrasound exposure times were $i=0$, $j=120$, $k=240$, $l=540$, $m=1000$, $n=2000$ ms. Pulse length was held constant at 60 ms. One-way ANOVA indicates statistical significance of dependence on exposure time for the fraction of cells in each subpopulation: $-$ ($p>0.05$), $+$ ($p<0.05$), $++$ ($p<0.01$), $+++$ ($p<0.001$), for NUP (upper), LUP (middle), and HUP (lower) subpopulations.

D. Exposure time dependence of subpopulation distribution among viable cells

A similar trend was observed when studying the effects of ultrasound exposure time at constant pressure. To demonstrate this more clearly, the data in Fig. 4 were replotted as a function of exposure time in Fig. 5, which shows that NUP_{viable} generally decreased, LUP_{viable} remained relatively constant, and HUP_{viable} generally increased with increasing exposure time. As with increasing pressure, the composition of viable cells became richer in HUP at longer exposure times.

E. Subpopulation distribution among all cells

The analysis presented so far has addressed only the cells that remain viable after exposure to ultrasound and divided them into three subpopulations. However, large numbers of cells can be made nonviable by ultrasound. A more complete analysis should consider that a cell exposed to ultrasound could have one of four fates: either it remains viable and falls into the NUP, LUP, or HUP subpopulation or it is rendered nonviable.

To account for nonviable cells, we recalculated the dis-

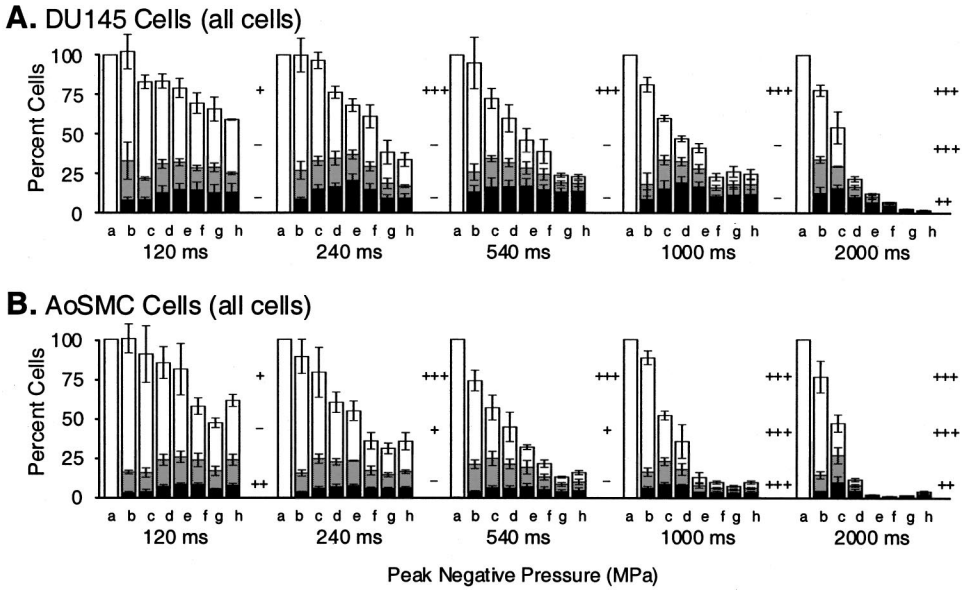


FIG. 6. Subpopulation distribution among all cells as a function of pressure at different exposure times. The white bars indicate cells in NUP, gray bars indicate LUP, and black bars indicate HUP. Overall DU145 (A) and AoSMC (B) cell viability decreased with increasing pressure at constant exposure time. Percent cells in NUP decreased with increasing pressure. Percent cells in LUP and HUP changed little as pressure was varied except when viability was very low. Ultrasound peak negative pressures were: $a=0$, $b=0.6$, $c=1.2$, $d=1.6$, $e=2.0$, $f=2.4$, $g=2.8$, and $h=3.0$ MPa. Pulse length was held constant at 60 ms. One-way ANOVA indicates statistical significance of dependence on acoustic pressure for the fraction of cells in each subpopulation: $-$ ($p>0.05$), $+$ ($p<0.05$), $++$ ($p<0.01$), $+++$ ($p<0.001$), for NUP (upper), LUP (middle), and HUP (lower) subpopulations.

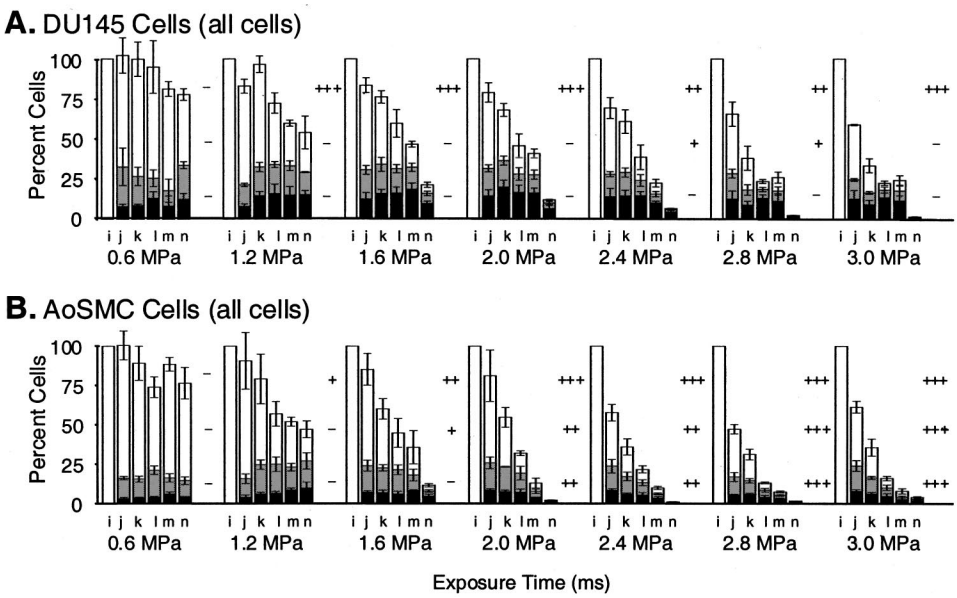


FIG. 7. Subpopulation distribution among all cells as a function of exposure time at different pressures. The white bars indicate cells in NUP, gray bars indicate LUP, and black bars indicate HUP. Overall DU145 (A) and AoSMC (B) cell viability decreased with increasing exposure time at constant peak negative pressure. Percent cells in NUP decreased with increasing exposure time. Percent cells in LUP and HUP changed little as exposure time was varied except when viability was very low. Ultrasound exposure times were: $i=0$, $j=120$, $k=240$, $l=540$, $m=1000$, and $n=2000$ ms. Pulse length was held constant at 60 ms. One-way ANOVA indicates statistical significance of dependence on exposure time for the fraction of cells in each subpopulation: $-$ ($p>0.05$), $+$ ($p<0.05$), $++$ ($p<0.01$), $+++$ ($p<0.001$), for NUP (upper), LUP (middle), and HUP (lower) subpopulations.

tribution of cells among the three subpopulations as a fraction of all cells present during ultrasound exposure. Figures 6 and 7 present the distribution of all cells as functions of pressure and exposure time. The height of each bar represents the fraction of cells that remained viable and the white, gray, and black bars represent the three subpopulations of viable cells. These figures show that, overall, cell viability generally decreased with increasing pressure and exposure time, as discussed previously (Guzmán *et al.*, 2001). Figures 6 and 7 also demonstrate that increasing pressure and exposure time generally caused a decrease in $\mathrm{NUP}_{\mathrm{all\ cells}}$ (i.e., the NUP subpopulation expressed as a fraction of all cells exposed to ultrasound), but generally did not affect $\mathrm{LUP}_{\mathrm{all\ cells}}$ or $\mathrm{HUP}_{\mathrm{all\ cells}}$. At long exposure times and high pressures, viability was extremely low and therefore the LUP and HUP subpopulations were decreased.

These trends contrast with those observed in Figs. 4 and 5, which did not account for the loss of viability associated with increasing pressure and exposure time. The analysis shown in Figs. 4 and 5 on the basis of only viable cells may be useful for scenarios where cell death is not of primary concern, but delivery of molecules is critical. For example, laboratory scientists may be more concerned with efficiency of gene or protein uptake among viable cells, since the surviving cell population can be rapidly grown in culture to yield more cells. Alternatively, analysis on the basis of all cells shown in Figs. 6 and 7 may be useful for medical researchers and clinicians interested in ultrasound conditions that deliver large amounts of therapeutic material without excessive cell death (e.g., useful for targeted drug delivery or gene therapy) or with extensive cell death (e.g., useful for cancer treatment).

F. Pulse length dependence of subpopulation distribution

When pulse length was varied between 20 μs (10 cycles/pulse) and 60 ms (30 000 cycles/pulse), subpopulation distributions did not change significantly (Fig. 8). This conclusion is similar to the result presented previously (Guzmán *et al.*, 2001), where overall viability was not affected by varying pulse length over the range of conditions tested.

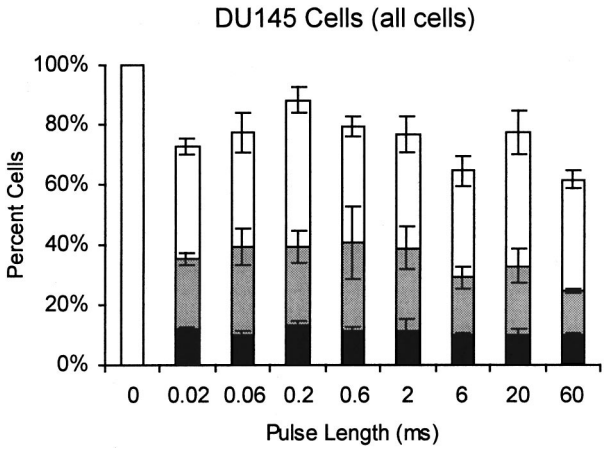


FIG. 8. Subpopulation distribution among all cells as a function of pulse length. The distribution of subpopulations does not vary significantly with pulse length over the range of 0.02–60 ms at 240-ms exposure time and 1.6-MPa peak negative acoustic pressure. Black bars indicate cells in HUP, gray bars indicate LUP, and white bars indicate NUP. One-way ANOVA indicates statistical significance of dependence on pulse length for the fraction of cells in each subpopulation: p value$=0.53$, 0.87, and 0.79 for NUP, LUP, and HUP cells, respectively.

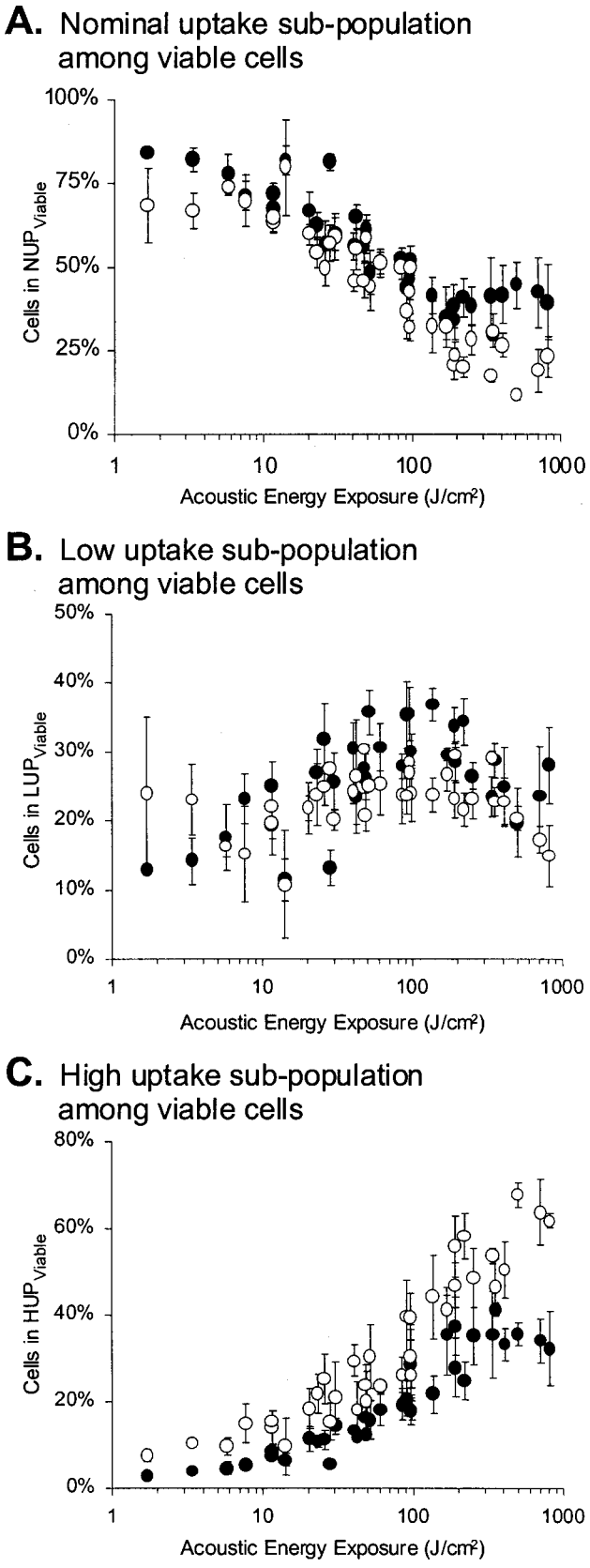


FIG. 9. Subpopulation distribution among viable cells as a function of acoustic energy exposure. (A) The NUP subpopulation decreased (multiple $R^2 = 0.68$ and 0.59 for DU145 and AoSMC cells, respectively), (B) the LUP subpopulation varied little (multiple $R^2 = 0.13$ and 0.29 for DU145 and AoSMC cells, respectively), and (C) the HUP subpopulation increased with increasing energy (multiple $R^2 = 0.66$ for both DU145 and AoSMC cells) (○ = DU145, ● = AoSMC).

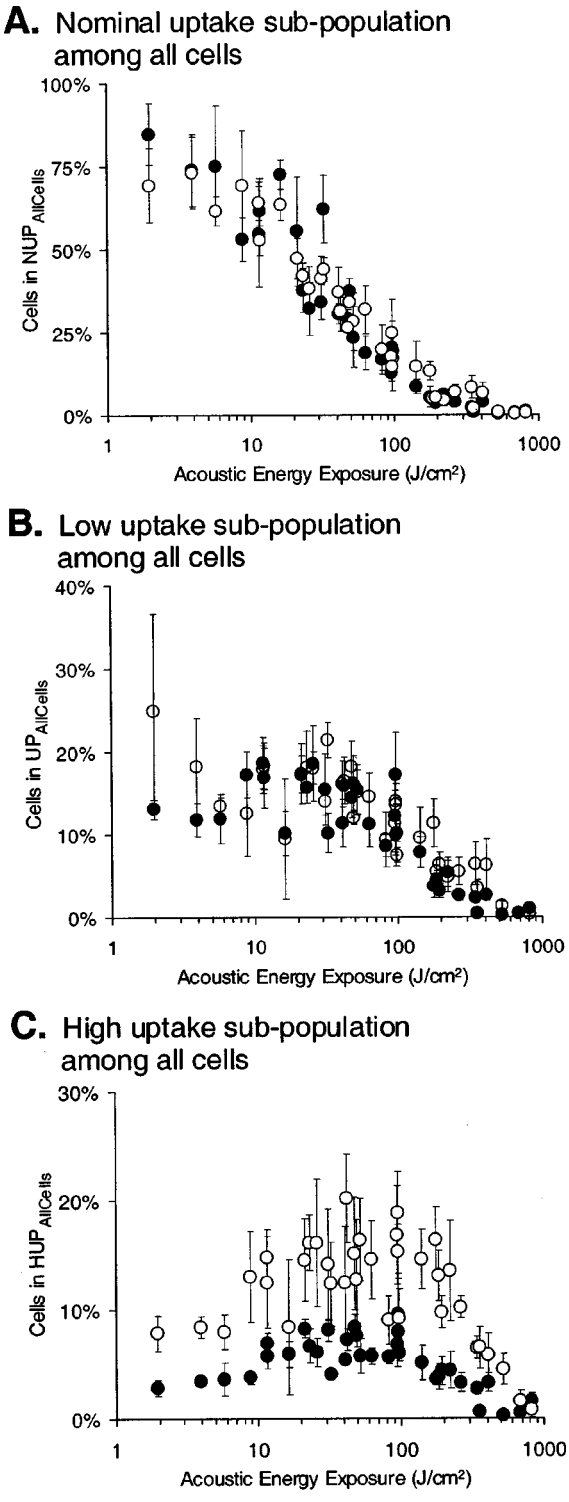


FIG. 10. Subpopulation distribution among all cells as a function of acoustic energy exposure. (A) The NUP subpopulation decreased (multiple $R^2 = 0.79$ and 0.76 for DU145 and AoSMC cells, respectively) and (B) LUP and (C) HUP subpopulations increased, reached a maximum, and decreased with increasing energy (LUP: multiple $R^2 = 0.41$ and 0.56 for DU145 and AoSMC cells; HUP: multiple $R^2 = 0.27$ and 0.42 for DU145 and AoSMC cells) (○ = DU145, ● = AoSMC).

G. Acoustic energy correlation with subpopulation distribution

Because the distribution of cell subpopulations depends strongly on exposure time and even more strongly on pres-

sure, we hypothesized that the families of distributions shown in Figs. 4–8 could be collapsed down into single curves when plotted as a function of acoustic energy exposure (E). In Fig. 9, the subpopulation distributions among viable cells (Figs. 4 and 5) were replotted as a function of acoustic energy exposure. Figure 9 shows that for both cell types the data collapsed into single curves, where NUP_{viable} decreased, LUP_{viable} was scattered, and HUP_{viable} increased with increasing acoustic energy exposure.

The subpopulation distributions calculated on the basis of all cells were also correlated with acoustic energy exposure. For this, the data from Figs. 6 and 7 were replotted versus energy in Fig. 10. The figure shows that $NUP_{all\ cells}$ decreased (ANOVA $p<0.05$); $LUP_{all\ cells}$ and $HUP_{all\ cells}$ probably initially increased (since no LUP or HUP cells were present in controls), leveled out ($E<50\,J/cm^2$, ANOVA $p>0.05$), and then decreased ($E>50\,J/cm^2$, ANOVA $p<0.05$) with increasing acoustic energy exposure.

H. Acoustic energy correlation with molecular uptake

Figure 3 indicates that uptake within each subpopulation (i.e., NUP, LUP, and HUP) was not significantly different from each other. This is shown again in Fig. 11, in which these data are replotted versus acoustic energy exposure. However, the scatter within each subpopulation can be partially explained by an increasing trend with increasing energy for NUP and LUP cells (one-way ANOVA $p<0.01$). Uptake within the HUP subpopulation was independent of energy exposure (one-way ANOVA $p>0.10$).

The average uptake in the HUP subpopulation (i.e., average of all HUP data points in Fig. 11) was $1.36\ (\pm 0.5)\times 10^7$ and $1.10(\pm 0.3)\times 10^7$ molecules per DU145 and AoSMC cell, respectively. Based on average cell volumes of 2200 μm^3 for DU145 cells and 2400 μm^3 for AoSMC cells (Guzmán *et al.*, 2001), this corresponds to intracellular concentrations of 10.3 ± 3.7 and $7.6\pm 2.1\,\mu M$, respectively. Because calcein was supplied extracellularly at 10 μM, average uptake by HUP cells approached the maximum possible based on chemical equilibrium in the absence of binding.

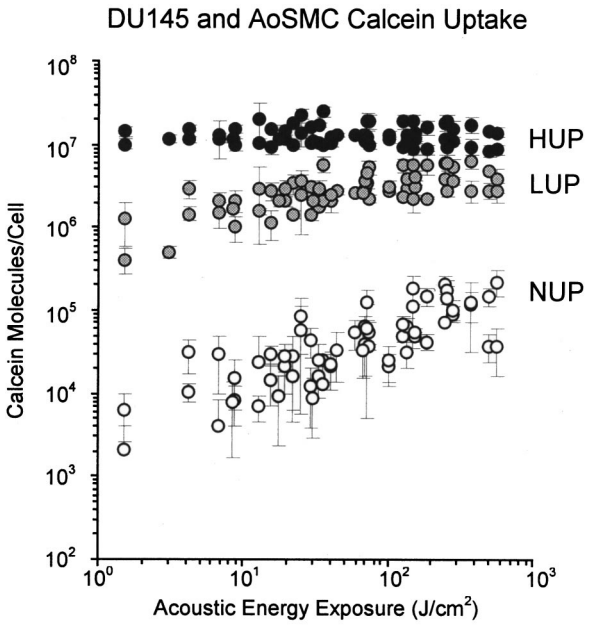


FIG. 11. The average number of calcein molecules delivered per cell (combined plot of DU145 and AoSMC cells) for each uptake subpopulation as a function of acoustic energy exposure. Calcein uptake per cell increased as a function of energy exposure for NUP and LUP cells, but showed no dependence for HUP cells.

IV. DISCUSSION

A. Cavitation-based versus cell-based heterogeneity

The most notable finding in this study is that cells exposed to the same acoustic environment exhibited cell-to-cell heterogeneity in their response to ultrasound, which is in direct support of our first hypothesis. This observation raises an interesting question: Are these heterogeneous effects due to (i) spatial and temporal heterogeneity in cavitation generated by ultrasound or (ii) heterogeneity based on differences between cells and their responses to acoustic cavitation?

In support of the first explanation, cavitation is known to be a stochastic phenomenon controlled by the location and availability of nucleation sites, which makes cavitation extremely heterogeneous in both time and space (Leighton, 1994; Miller *et al.*, 1996; Barnett, 1998). As a result, a cell's fate may be determined by the degree to which that cell came in contact with one or more of the finite number of stable or inertial cavitation bubbles. The mechanism of cell disruption could involve just a single cell-with-bubble interaction or the accumulated effect of multiple interactions. In this way, cells that do not come into sufficiently close contact with a cavitation bubble experience no effect (i.e., NUP); cells that have progressively closer interactions with bubbles fall among the broad LUP and HUP subpopulations; and cells that locally experience too much cavitation are made nonviable. Based on the calculation method described by Ward *et al.* (2000) the average initial cell-to-bubble spacing in the present study was 45 μm, which is 2–3 times the diameter of the cells used.

Heterogeneity in cavitation could also come from insufficient mixing. The literature indicates that exposure vessel rotation may be necessary to promote cell–bubble mixing and thereby enhance bioeffects (Church *et al.*, 1982). However, other studies have shown that when contrast agent nucleation sites are present, there is no difference in bioeffects caused in rotating and nonrotating exposure vessels (Miller *et al.*, 1999; Brayman and Miller, 1999). For our experiments, in which contrast agent was present, we visually observed vigorous mixing within cell samples during ultrasound exposure and therefore believe that insufficient mixing was not the source of heterogeneity. Instead the finite number and short lifetime of bubbles may limit the odds of having cell–bubble interactions.

The second possible explanation for heterogeneity is based on biological differences between individual cells that affect the cell's mechanical properties or otherwise influence how each cell interacts with cavitation bubbles. In this case, all cells could have identical cell–bubble interactions (i.e., no time-averaged heterogeneity in cavitation), but different subpopulations of cells might react to cavitation differently. Based on differences in cell cycle or other factors, some cells

might be reversibly affected by a given ultrasound exposure (i.e., LUP or HUP), others might be unaffected (i.e., NUP), and still others might be rendered nonviable. Additional experiments are needed to determine the relative roles of heterogeneity in cavitation and between cells as the cause of heterogeneity in uptake and cell viability.

B. Comparison to heterogeneity and homogeneity seen in other studies

Heterogeneity and homogeneity seen in other studies using ultrasound and a related phenomenon, electroporation, provide additional perspective on the results reported here. Previous ultrasound studies have observed heterogeneity in the form of three possible responses: cells that are unaffected, induced to take up molecules, or rendered nonviable. For example, the literature has numerous reports of ultrasound having heterogeneous effects on cell viability, where ultrasound renders only a fraction of cells nonviable (Brayman and Miller, 1999; Church *et al.*, 1982; Greenleaf *et al.*, 1998; Miller *et al.*, 1996). Other studies have reported heterogeneity in gene transfection, which is an indirect measure of the combined effects of gene uptake and cell viability (Bao *et al.*, 1997; Greenleaf *et al.*, 1998). While the present study confirms and quantifies these earlier observations, it introduces an additional aspect of heterogeneity: among cells induced to take up molecules, there is a broad distribution in levels of molecular uptake, that starts just above control values (NUP) and approaches chemical equilibrium with the extracellular solution (HUP). This result has not been reported before, probably because the cell-by-cell measurements needed to show it have not been performed.

Ultrasound's heterogeneous effects are dramatically different from the distribution of molecular uptake caused by electroporation. Electroporation is a phenomenon that similarly causes varying degrees of molecular uptake and cell viability in response to a brief (e.g., μs to ms) electric field pulse (Chang *et al.*, 1992). In contrast to results seen here, all cells subjected to a given electroporation exposure have been shown to take up molecules in a homogeneous manner for studies with DU145 cells (Canatella *et al.*, 2001) and human red blood cell ghosts (Prausnitz *et al.*, 1993). For spherical cells, electroporation histograms similar to those in Fig. 1 do not show multiple populations, but are always characterized by a single peak with an approximately Gaussian distribution. Moreover, the location of that peak can occur anywhere along the *x* axis of the histogram, depending on electroporation conditions used. This difference may be explained by the spatial and temporal uniformity of the electric field during electroporation, as opposed to the heterogeneity of cavitation.

C. Significance of correlation with acoustic energy exposure

In support of our second hypothesis, heterogeneity in molecular uptake was found to correlate with acoustic energy exposure. A similar dependence on energy exposure was shown previously for cell viability and total molecular uptake (Guzmán *et al.*, 2001) and increased skin conductivity (Mitragotri *et al.*, 2000). This dependence on energy is important because it provides a single parameter that correlates with the various bioeffects induced by extremely complex cavitational activity. Using this correlation, the observed bioeffects can be approximately described by energy exposure regardless of the applied pressure, exposure time, or pulse length. This approach may allow a researcher or clinician to identify ultrasound conditions that yield useful effects. For example, a physician performing gene therapy may want to maximize $HUP_{\text{all cells}}$ and therefore apply the optimal energy exposure (e.g., 50–350 J/cm^2 in the present study, see Fig. 10) to deliver large numbers of molecules necessary for transfection into as many cells as possible. For targeted chemotherapy, one may want to minimize $NUP_{\text{all cells}}$ and therefore operate at a much higher energy to assure that all cancer cells are killed either directly by ultrasound or indirectly due to uptake of a chemotherapeutic agent.

V. CONCLUSIONS

When measured on a cell-by-cell basis, the number of molecules taken up by cells exposed to the same ultrasound conditions was shown to be extremely heterogeneous. This cell-to-cell heterogeneity was observed in both AoSMC and DU145 cells at all of the ultrasound conditions tested. After exposure to ultrasound, viable cells could be divided into three subpopulations, for which the fraction of cells in each subpopulation depended strongly on pressure and exposure time, but not on pulse length, and was found to correlate with acoustic energy exposure. Molecular uptake within each subpopulation remained relatively constant. Because ultrasound's effects are thought to be mediated by cavitation, the existence of subpopulations suggests heterogeneity in the time and location of acoustic cavitation and/or in cellular responses to cavitation.

ACKNOWLEDGMENTS

We would like to thank Dr. Paul Canatella, Dr. Keyvan Keyhani, Dr. Thomas Lewis, Dr. Russel Heikes, and Dr. Peter MacDonald for their helpful comments and suggestions.

Bao, S., Thrall, B. D., and Miller, D. L. (**1997**). ''Transfection of a reporter plasmid into cultured cells by sonoporation *in vitro*,'' Ultrasound Med. Biol. **23**, 953–959.

Barnett, S. (**1998**). ''Nonthermal issues: cavitation—its nature, detection and measurement,'' Ultrasound Med. Biol. **24**, S11–S21.

Brayman, A. A., and Miller, M. W. (**1999**). ''Sonolysis of Albunex-supplemented, 40% hematocrit human erythrocytes by pulsed 1-MHz ultrasound: Pulse number, pulse duration, and exposure vessel rotation dependence,'' Ultrasound Med. Biol. **25**, 307–314.

Canatella, P. J., Karr, J. F., Petros, J. A., and Prausnitz, M. R. (**2001**). ''Quantitative study of electroporation-mediated molecular uptake and cell viability,'' Biophys. J. **80**, 755–764.

Chang, D. C., Chassy, B. M., Saunders, J. A., and Sowers, A. E., (**1992**). *Guide to Electroporation and Electrofusion* (Academic, New York).

Church, F. C., Flynn, H. G., Miller, M. W., and Sacks, P. G. (**1982**). ''The exposure vessel as a factor in ultrasonically-induced mammalian cell lysis. II,'' Ultrasound Med. Biol. **8**, 299–309.

Fechheimer, M., Denny, C., Murphy, R. F., and Taylor, D. L. (**1986**). ''Measurement of cytoplasmic *p*H in dictyostelium discoideum by using a new method for introducing macromolecules into living cells,'' Eur. J. Cell Biol. **40**, 242–247.

Gift, E. A., and Weaver, J. C. (**1995**). ''Observation of extremely heterogeneous electroporative molecular uptake by *Saccharomyces cerevisiae* which changes with electric field pulse amplitude,'' Biochim. Biophys. Acta **1234**, 52–62.

Greenleaf, W. J., Bolander, M. E., Sarkar, G., Goldring, M. B., and Greenleaf, J. F. (**1998**). ''Artificial cavitation nuclei significantly enhance acoustically induced cell transfection,'' Ultrasound Med. Biol. **24**, 587–595.

Guzmán, H. R., Nguyen, D. X., Sohail, K., and Prausnitz, M. R. (**2001**). ''Ultrasound-mediated disruption of cell membranes. I. Quantification of molecular uptake and cell viability,'' J. Acoust. Soc. Am. **110**, 588–596.

Harrison, G. H., Balcer-Kubiczek, E. K., and Gutierrez, P. L. (**1996**). ''*In vitro* mechanisms of chemopotentiation by tone-burst ultrasound,'' Ultrasound Med. Biol. **22**, 355–362.

Leighton T. G. (**1994**). *The Acoustic Bubble* (Academic, New York).

Miller, D. L., Bao, S., and Morris, J. E. (**1999**). ''Sonoporation of cultured cells in the rotating tube exposure system,'' Ultrasound Med. Biol. **25**, 143–149.

Miller, M. W., Miller, D. L., and Brayman, A. A. (**1996**). ''A review of *in vitro* bioeffects of inertial ultrasonic cavitation from a mechanistic perspective,'' Ultrasound Med. Biol. **22**, 1131–54.

Mitragotri, S., Farrell, J., Tang, H., Terahara, T., Kost, J., and Langer, R. (**2000**). ''Determination of threshold energy dose for ultrasound-induced transdermal drug transport,'' J. Controlled Release **63**, 41–52.

Prausnitz, M. R., Lau, B. S., Milano, C. D., Conner, S., Langer, R., and Weaver, J. C. (**1993**). ''A quantitative study of electroporation showing a plateau in net molecular transport,'' Biophys. J. **65**, 414–422.

Robinson, J. P., and Kelley, S., computer code LDATA, Purdue University, West Lafayette, IN, 1998.

Saad, A. H., and Hahn, G. M. (**1987**). ''Ultrasound enhances Adriamycin toxicity *in vitro*,'' in *Heat Transfer in Bioengineering and Medicine*, edited by J. C. Chato, T. E. Diller, K. R. Diller, and R. B. Roemer (American Society of Mechanical Engineers Press, New York).

Ward, M., Wu, J., and Chiu, J. (**2000**). ''Experimental study of the effects of Optison concentration on sonoporation *in vitro*,'' Ultrasound Med. Biol. **26**, 1169–1175.

A model of echolocation of multiple targets in 3D space from a single emission[a)]

Ikuo Matsuo[b)]
Research Institute of Electrical Communication, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan

Junji Tani
Institute of Fluid Science, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan

Masafumi Yano
Research Institute of Electrical Communication, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan

(Received 29 November 2000; revised 2 April 2001; accepted 9 April 2001)

Bats, using frequency-modulated echolocation sounds, can capture a moving target in real 3D space. The process by which they are able to accomplish this, however, is not completely understood. This work offers and analyzes a model for description of one mechanism that may play a role in the echolocation process of real bats. This mechanism allows for the localization of targets in 3D space from the echoes produced by a single emission. It is impossible to locate multiple targets in 3D space by using only the delay time between an emission and the resulting echoes received at two points (i.e., two ears). To locate multiple targets in 3D space requires directional information for each target. The frequency of the spectral notch, which is the frequency corresponding to the minimum of the external ear's transfer function, provides a crucial cue for directional localization. The spectrum of the echoes from nearly equidistant targets includes spectral components of both the interference between the echoes and the interference resulting from the physical process of reception at the external ear. Thus, in order to extract the spectral component associated with the external ear, this component must first be distinguished from the spectral components associated with the interference of echoes from nearly equidistant targets. In the model presented, a computation that consists of the deconvolution of the spectrum is used to extract the external-ear-dependent component in the time domain. This model describes one mechanism that can be used to locate multiple targets in 3D space.
[DOI: 10.1121/1.1377294]

PACS numbers: 43.80.Lb, 43.64.Bt, 43.80.Jz [WA]

I. INTRODUCTION

Bats have the special ability to form an image of the world based on acoustic information, while most animals form images of the world that are more strongly based on visual information. They emit trains of high-frequency sounds and can locate an individual target among multiple objects by using the echoes of these emissions that overlap in time (Griffin, 1958; Simmons *et al.*, 1995a). In order to recognize objects in three-dimensional (3D) space, bats need to process the information that the echoes contain to determine the distance and direction to each object.

With regard to the determination of distance, it has been found in physiological studies that delay-tuned neurons in the central auditory system are capable of determining the interval between an emitted sound and the returning echo (Dear *et al.*, 1993a, b; Kawasaki *et al.*, 1988; Schuller *et al.*, 1991; Suga, 1984; Suga *et al.*, 1978; Sullivan, 1982). In a case of interfering echoes, responses of neurons in the inferior colliculus (IC) depend on the amplitude change based on the interference between echoes (Sanderson and Simmons, 2000). There might be the mechanism that a bat uses to determine a target's distance.

There are a number of existing behavioral experiments involving bats and their determination of distances to objects. The results of these experiments regarding the resolution with which bats are capable of determining these distances vary greatly, depending on the specific conditions of the experiments. When two different targets are presented, either simultaneously or sequentially on either side of the animal, the accuracy of the determination of the difference in distance to the two objects is about 1–2 cm, corresponding to a time resolution of about 50–100 μs (Simmons, 1973; Miller, 1991; Surlykke, 1992). Contrastingly, when two phantom targets, one of which emits ''phantom echoes'' at times which follow the bat's emissions after a fixed interval and one of which emits ''phantom echoes'' at times which follow the bat's emissions after intervals whose lengths fluctuate about some average value (a ''jitter condition''), are presented, bats can distinguish objects separated by distances of well under 1 mm (Simmons *et al.*, 1990a). In such distance discrimination experiments, it has been clearly demonstrated that a bat can accurately distinguish minute variations

[a)]A portion of this work has been presented at the WESTPRAC, Kumamoto, Japan, October 2000.
[b)]Electronic mail: matsuo@riec.tohoku.ac.jp

in the arrival times of echoes from simple targets with single reflecting surfaces, although the mechanism allowing this incredible spatial resolution is not yet well understood.

Bats distinguish the arrival times of echoes from objects that are positioned at nearly equal distances (Simmons, 1989; Simmons *et al.*, 1990b). The echoes from objects whose distances from a bat are sufficiently close, termed ''nearly equidistant objects,'' will overlap and interfere with each other, so that the arrival time corresponding to the more distant target cannot be directly determined from the detection of returning echoes. It thus may be inferred from a bat's ability to discriminate two closely situated targets that they locate each target by utilizing spectral information produced by the interference of the echoes from the two targets. However, there remains the important problem of determining how the bat can determine the correspondence between echoes and targets among those received from multiple and nearly equidistant targets.

Bats are able to discriminate the direction of an echo's source to within 1.5°–3.0° (Lawrence and Simmons, 1982; Simmons *et al.*, 1983). Binaural cues used for directional localization are the interaural time difference (ITD) and interaural intensity difference (IID). In most animals, neurons in the medial superior olive (MSO) within the superior olive complex have been found to respond to the ITD (Pickles, 1988). However, neurons in the MSO of echolocating bats are neither significately excited nor inhibited by input through the ipsilateral ear (Covey and Casseday, 1991; Grothe *et al.*, 1992). This suggests that the ITD is not a crucial cue in horizontal localization for bats. Another possibility as a useful cue in horizontal localization is the IID. In fact, there are IID sensitive neurons in the lateral superior olive (LSO) of bats, which are excited by input through the ipsilateral ear and inhibited by input through the contralateral ear (Casseday *et al.*, 1988; Covey and Casseday, 1991). Moreover, neurons in the superior colliculus (SC) and auditory cortex (AC) exhibit directional sensitivity (Jen *et al.*, 1984, 1989; Valentine and Moss, 1997). However, it remains unclear how bats actually use these cues for the localization of targets.

The localization of targets in 3D space has been extensively investigated in computational studies. A binaural model was proposed by Altes (1978) to determine the distance and the direction of a single target using an idealized directivity pattern generated by emission from the mouth and reception at the ears. This model contains a set of arbitrarily chosen frequency spectral pattern templates, one corresponding to each direction (Altes, 1978). In this model, the distance and direction to a target can be determined from the return time and by comparing the composite echo with the template set, respectively. In order to avoid the necessity for comparison between an echo and the elements of the template set one by one, Kuc (1994) simplified this model by introducing an emission consisting of a fundamental frequency and two overtone components. In this case, if the dependence of the reception at two ears for the two overtone components on the direction of the emission and from the mouth is given, the azimuthal direction and the elevational direction can be determined from the IID and the interfrequency intensity difference (IFID) (Kuc, 1994). These two binaural models are useful only for a single target, because the returning echoes from multiple targets usually interfere with each other to produce a spectrum that is too complex to identify the direction of any single target.

In order to distinguish multiple targets, in particular two targets closely located on a one-dimensional range axis, Saillant *et al.* (1993) proposed a monaural model referred to as the SCAT (spectrogram correlation and transformation) model. In this model, the returning composite of an echo is processed parallel in two pathways, one of which processes temporal information and the other spectral information (Saillant *et al.*, 1993). Initially, the time interval between an emission and the arrival of the resulting echoes is calculated in the temporal information pathway. This time interval corresponds to the absolute distance to the nearest target. In the spectral information pathway, the spectral peaks and notches in the composite echo are deconvoluted to determine the difference between delays for the two targets. Then the outputs from the two pathways are integrated to yield a representation which contains information describing the distance to the two objects as a whole and the finer structure of the two objects as individual bodies. The mechanism producing the high accuracy in estimating delay times of interfering echoes and the limitation of this model are analyzed (Peremans and Hallam, 1998).

This SCAT model can be easily extended to a binaural model, and such an extended model could locate a single target in a two-dimensional plane. In this model, to locate a single target in 3D space requires sampling of echoes from at least two different aspect angles. Also, in the case of the identification of two nearly equidistant targets in two-dimensional (2D) space, there arises the problem of the ambiguity of the correspondence between perceived echoes and objects. In particular, each ear perceives two echoes, and the problem is to determine the object corresponding to each of the four perceived echoes. To remove this ambiguity and to correctly locate two nearly equidistant targets in 2D or 3D space, in this model it is necessary to sample echoes from at least two different aspect angles. However, it must be the case that bats can locate each target's position in 3D space using echoes from only a single emission and thereby be capable of following a target's changing position, because they are able to capture a flying target among multiple objects by successively emitting trains of sound during the target's pursuit (Simmons *et al.*, 1995a). Theoretically, however, it is not yet known how such a determination of the location of a single or two nearly equidistant targets in 3D space can be made from echoes produced by a single emission.

Determining the locations of multiple targets in 3D space requires independently fixing the direction as well as the distance of each target. Recently, it has been found by measuring the transfer function at the external ear that spectral notches, which are local minima in the amplitude of echoes, are produced at each external ear (Fuzessery, 1996; Wotton *et al.*, 1995). The frequency of these spectral notches is called the ''external ear dependent notch frequency (EEDNF).'' Since the EEDNF varies with the direction of a

sound source, this might be a crucial cue for the localization of targets (Fuzessery, 1996; Wotton *et al.*, 1995). In fact, it has been demonstrated in behavioral experiments that the EEDNF is an essential cue for the determination of target direction (Wotton *et al.*, 1996; Wotton and Simmons, 2000). In the case of multiple nearly equidistant targets, the echo spectrum includes contributions from the interference between echoes from these targets and from the interference created through the physical process of reception at the external ear (Simmons *et al.*, 1995b). In order to locate each object in 3D space, it is therefore necessary to separate the contributions from the interference created through the physical process of reception at the external ear and the interference between echoes from the targets. In this article, we propose a model to discriminate multiple nearly equidistant targets in 3D space by analyzing the echoes produced by a single emission and determining the difference between delay times and the EEDNF corresponding to each object.

In Sec. II, we describe the model to determine the locations of multiple targets in 3D space. The results that are obtained by numerical analysis are presented in Sec. III. The predominant features of this model are discussed in Sec. IV.

II. MODEL

A. Outline of the model

We propose a model capable of determining the positions of multiple nearly equidistant targets using only the overlapping and interfering echoes from each target resulting from a single emission. To determine the location of each of multiple closely positioned targets in 3D space requires cues regarding both the distance and direction to each target. The distances from a bat to objects in a target region, including single or multiple closely positioned targets, are determined by the delay time between the emission and the return of the group of echoes. Also, the spectrogram of echoes from multiple closely positioned targets includes contributions from both the interference between the echoes from different nearly equidistant targets and interference created through the physical process of reception at the external ear. The component of the spectral pattern dependent on the interference between the echoes from the targets is called the ''target dependent spectral pattern (S_{target}),'' and the component of the spectral pattern dependent on the interference caused by the physical process of reception at the external ear is called the ''external ear dependent spectral pattern (S_{ext_ear}).'' The S_{target} and the S_{ext_ear} provide information for determination of the targets' distances and directions, respectively, and for this reason, in order to make such determinations, it is necessary to distinguish them within the spectrogram of echoes. The position of each object in the target region is represented by information regarding the distance to the target region as a whole and information regarding the fine-scale structure within this region. The EEDNF associated with the target in question is the frequency corresponding to the minimum amplitude in the S_{ext_ear}. In this model, the necessary mechanism to localize is the method to discriminate between the S_{target} and the S_{ext_ear} from echoes and then to determine the difference between delay times of echoes from nearly equidistant targets and the EEDNF associated with each target.

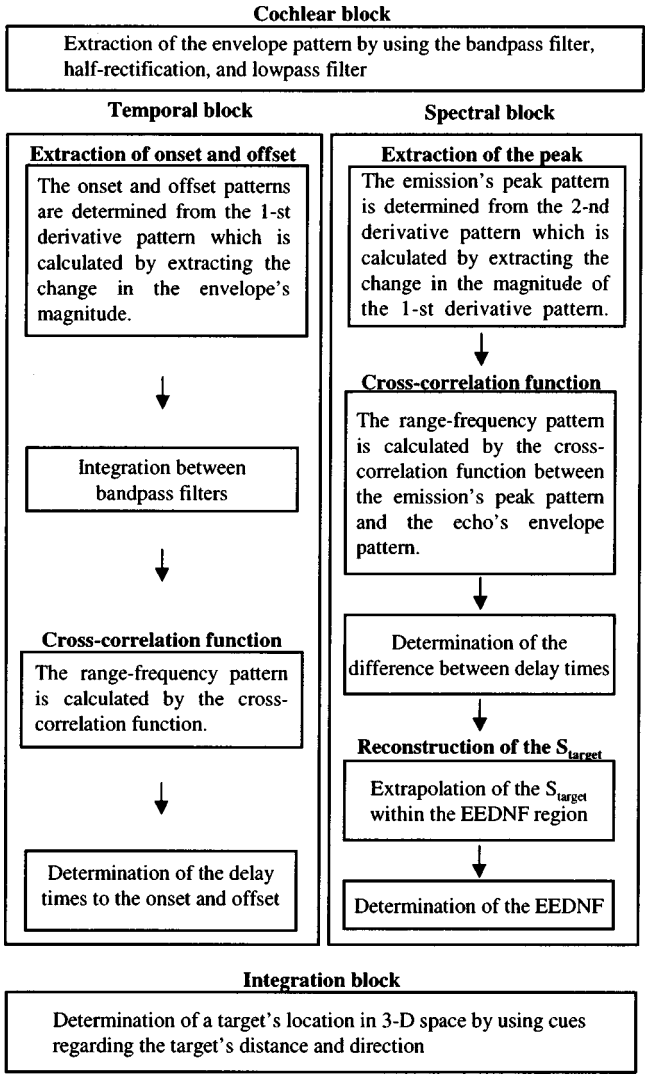


FIG. 1. Diagram of the model. The model consists of four blocks. The inputs to the cochlear block are the emission itself and the echoes reflected from targets.

As shown in Fig. 1, our model consists of four processing blocks. Inputs to the cochlear block are the emission and echoes reflected from objects. This block is followed by two parallel pathways, the temporal and spectral blocks. The function of the temporal block is to determine the distance to the target region, and that of the spectral block is to determine both the difference between delay times of echoes from nearly equidistant targets and the EEDNF associated with each target. Finally, the function of the binaural integration block is to determine a target's location in 3D space by using cues containing information about the distance to each target and the EEDNF for each target.

B. Characteristics of emission and echo

1. Emission

Emissions produced by echolocating bats consist of constant-frequency (CF) and frequency-modulated (FM)

components. There are two types of bat emission patterns, those that consist of CF-FM emissions and those that consist of FM emissions. Bats of the species *Myotis lucifugus* and *Eptesicus fuscus* use only FM emissions to locate objects (Griffin, 1958; Simmons *et al.*, 1995a). In our model, we use FM emissions, because the purpose of the model is to identify a target's location.

A single emission of an echolocating bat consists of a principle frequency that sweeps through a range of values (usually from ~10 kHz to ~100–200 kHz), together with a single or multiple harmonics (Fenton and Bell, 1981). An emission, for example, of *Myotis lucifugus* has one type of fundamental FM emission. Its duration is from 2 to 5 ms, and its frequency begins at 100 kHz and sweeps through values down to 40 kHz (Fenton and Bell, 1981). In our model, we use an emission having a single FM sweep. The spectrum of the emission has a constant amplitude from 20 to 120 kHz. For this reason, the 3.6-ms-long FM emission begins at 130 kHz and sweeps down to 10 kHz. The rise/fall time is 50 μs. The amplitude of the different wavelength components emitted by actual bats depends on the direction of emission (Hartley and Suthers, 1989; Shimozawa *et al.*, 1974; Simmons *et al.*, 1995b). However, because each wavelength is emitted over a large range of directions, in our model we assume no directional dependence of emitted wavelengths for simplicity.

2. Echo

a. Interference of echoes from nearly equidistant targets. We assume the following conditions to simplify the simulation.

(i) Each target is a point target. An extended object is considered as consisting of a distribution of acoustically reflecting points, called ''glints'' (Altes, 1976). In behavioral experiments with simulated targets, situations in which this assumption can be considered valid are created by enabling only one loudspeaker at a time.

(ii) The reflecting target is a perfect reflector in the sense that it does not modify the phase or amplitude of the sound wave. This assumption simplifies the analysis, because an echo returning from a single reflecting point is an exact copy of the emission.

(iii) The emitting mouth is midway between two ears, and the distance between the mouth and each ear is 2 cm. This approximates the small head of an echolocating bat.

b. Interference created by physical processes at the external ear. The external ears of the bat *Eptesicus fuscus* serve as receiving antennas for sonar echoes (Wotton *et al.*, 1995). The transfer function at each external ear varies with the directions (azimuthal and elevational angles) of the targets. The EEDNF ranges from 30 kHz for elevations between 30° and 40° below the horizontal to about 50 kHz for elevations at or near the horizontal (Wotton *et al.*, 1995). Also, for Pallid bats (*Antrozous p. pallidus*), the EEDNF changes according to some fixed prescription in both azimuthal and elevational directions (Fuzessery, 1996).

To simplify our simulation, the EEDNF (in kHz) as a function of the azimuthal and elevational angles (in degrees) of a sound incident to an ear is defined by making reference to the measured transfer function of the external ear of the bats, *Eptesicus fuscus* (Wotton *et al.*, 1995) as follows:

$$\mathrm{EEDNF}(A,E) = CA/2 - E/5 + 40, \tag{1}$$

where A and E are the azimuthal and elevational angles of the incident sound. Since the EEDNF is assumed to possess bilateral symmetry, the value of C is 1 for the left ear and -1 for the right ear. Because azimuthal directions from $-8°$ to 8° and elevational directions from 0° to 20° are considered in this simulation, the EEDNF is varied from 31 to 44 kHz. This frequency region is called the ''external ear dependent notch frequency region (EEDNF region).'' The frequency region outside this region is called the ''non-EEDNF region.''

A transfer function of the external ear is modeled by a second order infinite impulse response (IIR) filter. The filter's cutoff bandwidth is 3 kHz, and the attenuation of the amplitude at the notch frequency is 20dB with respect to the maximum.

C. Example of emission and echo

The polar coordinate axes (r, θ, ϕ) are used to represent a target's location in 3D space. Throughout the paper, we consider the mouth to be at the origin of the coordinate system. Then, r, θ, and ϕ are the distance, azimuthal angle, and elevational angle, respectively, to the target as measured from the position of the mouth. Here, we consider the situation in which three targets are located at (0.730 m, 0°, 0°), (1.000 m, 0°, 0°), and (1.040 m, 0°, 0°), in order to demonstrate the output of the model's processing due to echoes from a single isolated target and two nearly equidistant targets. The delay times for the respective targets are 4295, 5883, and 6118 μs, and the distances from the mouth to each ear via the targets are 1.460, 2.000, and 2.080 m. The difference between delay times for the second and third targets is 235 μs, corresponding to 80 mm along the range axis. The EEDNF is 40 kHz for each target, as given by Eq. (1).

Figure 2 displays the waveforms of the emission and the returning echoes, which are the inputs to the cochlear block. Since the difference between delay times for the first target and the third target is shorter than the duration of each echo, echoes from the three targets overlap.

D. The cochlear block

The function of the cochlear block is to transform the waveforms of the emission and the echoes into the spectrograms in a manner that simulates this process in the mammalian auditory system. The processing of the bat's inner ear can be modeled by a bandpass filter bank followed by an envelope-smoothing process as proposed by the SCAT model (Saillant *et al.*, 1993). In this model, this bandpass filter bank is composed of the 101 constant-bandwidth, second-order IIR filters. The center frequencies of these fil-

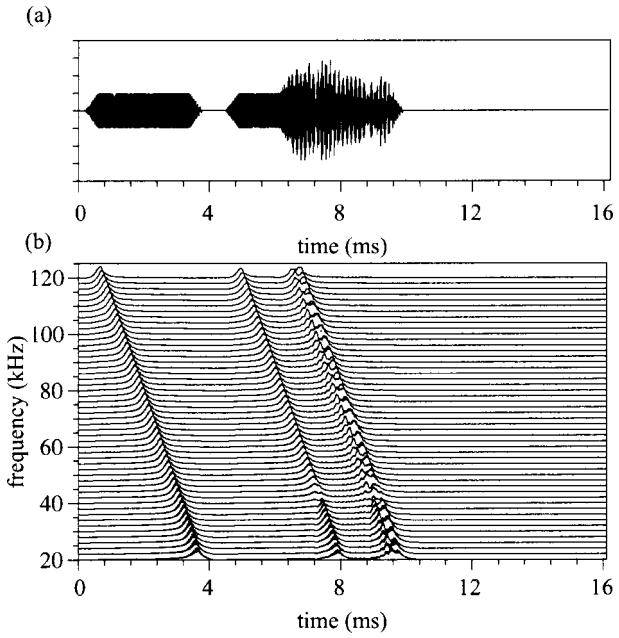


FIG. 2. Waveforms and the spectrogram of the emission and the returning echoes. In this case, three targets are located at (r,θ,ϕ) $=(0.730\ \text{m},0°,0°)$, (1.000 m, 0°, 0°), and (1.040 m, 0°, 0°). Here r is a target's distance from the mouth, θ is its azimuthal angle, and ϕ is its elevational angle. The delay times for the targets are 4295, 5883, and 6118 μs, and the distances from the mouth to each ear via the targets are 1.460, 2.000, and 2.080 m. The difference between delay times of the second and third targets is 235 μs, corresponding to 80 mm along the range axis. The EEDNF is found to be 40 kHz for each target. Echoes from the three targets overlap with one another. (b) The spectrogram of the emission and the returning echoes exhibits an array of smooth envelopes in the case of a linear frequency modulated (LFM) input signal (a FM sweep from 130 to 10 kHz in 3.6 ms).

ters are spaced from 20 to 120 kHz and are positioned with regular separations. These filter's bandwidths are the same with ones used by the SCAT model (Saillant *et al.*, 1993). The envelope-smoothing process is modeled by a half-wave rectification and a low-pass filtering of the output of each the 101 bandpass filters (Brugge *et al.*, 1969; Rose *et al.*, 1967; Russell and Palmer, 1986; Saillant *et al.*, 1993). Figure 2(b) displays the envelope pattern (B_{env}), which represents the outputs of the cochlear block for the emission and the echoes. Echoes from the second target and the third target interfere and are not discriminated from each other along the time axis. The use of a 6-dB/oct attenuating low-pass characteristic smoothes the half-wave rectified outputs of the bandpass filters without completely eliminating ripples at these center frequencies (Saillant *et al.*, 1993).

E. The temporal block

The function of the temporal block is to determine the delay times corresponding to the onset and offset in a group of echoes. Thus the distance to the nearest target can be directly determined from the delay time corresponding to onset, and the distance to the farthest target can be directly determined from the delay time corresponding to offset. The processing of the temporal block is made up of the three stages. The first stage is the transform from the envelope pattern into the onset and offset patterns at each bandpass filter channel. The second stage is the integration processing among bandpass filter channels to improve the temporal precision. The third stage is the cross-correlation function between the emission and the echoes returning from targets.

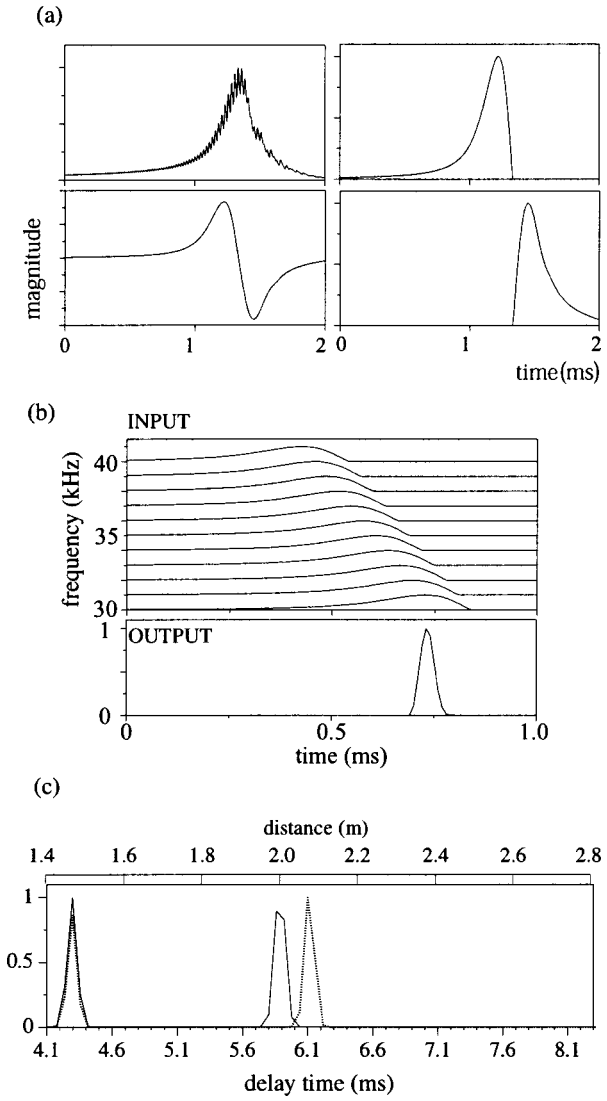


FIG. 3. Determination of the distance to the target region. (a) The extraction of the onset and offset patterns: The envelope pattern in a bandpass filter channel (center frequency= 40 kHz) (upper left), the envelope pattern transformed into the first derivative of the envelope pattern using the sine-wave asymmetric filter (lower left), the onset pattern extracted from the positive part of the first derivative (upper right), and the offset pattern extracted from the negative part of the first derivative are displayed (lower right). (b) Diagram of the integration processing among bandpass filter channels: The onset or offset pattern in each frequency channel is shown in the upper figure. The sweep rate between adjacent filter channels is called ''SR.'' The values of the onset and offset patterns in a given bandpass filter channel multiply those of the respective patterns in the next lower bandpass filter channel with the time delay SR in Eqs. (3) and (4). The integrated pattern is calculated through the integration processing among the 11 bandpass filter channels. This procedure sharpens the onset and offset peaks, as seen in the lower figure. (c) Final outputs of the temporal block: The solid and dotted curves represent the outputs calculated from the onset and offset patterns, respectively. The delay times corresponding to the peaks of the onset pattern are 4298 and 5858 μs, corresponding to 1.461 and 1.992 m along the range axis. The delays times corresponding to the peaks of the offset pattern are 4298 and 6098 μs, corresponding to 1.461 and 2.073 m, respectively. The distance to the first group of echoes is 1.461 m, and the distance to the second group of echoes ranges from 1.992 to 2.073 m along the range axis.

The onset and offset patterns are obtained from the envelope pattern, which is the output from the cochlear block [Fig. 3(a)]. The envelope pattern of the emission (B_{env}) in one bandpass filter channel (with center frequency 40 kHz) includes ripples due to the characteristic of the low-pass filter at the cochlear block. To determine the onset and offset delay times not in these ripples but in the group of echoes from a single or nearly equidistant targets along the range axis requires extracting the change in envelope's magnitude without these ripples along the time axis. Therefore, it needs both the low-pass filter attenuating enough at 20 kHz of the emission's minimum frequency and the filter to extract the change in envelope's magnitude along the time axis. For this reason, the envelope pattern is transformed into the first derivative pattern by deconvolution of a sine-waveform asymmetric filter with a period of 350 μs. This asymmetric filter is represented by

$$W_{\text{asym}}(t) = \sin(2\pi t/T), \quad 0 < t < T, \tag{2}$$

where t is time and T is the period ($=350$ μs). The onset pattern (B_{on}) and the offset pattern (B_{off}) are the positive part and the negative part of the derivative pattern, respectively. This process has no effect on the determination of the delay time because of deconvoluting both the emission's and echo's envelope patterns in the same way.

In the integration processing among bandpass filter channels, the precision of the temporal block is improved by integrating the onset and offset patterns of different bandpass filter channels. As shown in Fig. 3(b), the sweep rate between adjacent bandpass filter channels is referred to as ''SR''($=30$ μs/kHz). The values of onset or offset pattern in a bandpass filter channel multiply those of the respective pattern in the next lower bandpass filter channel with a time delay SR. A bank of integrated filters is composed of 19 integrated filter channels whose center frequencies range from 25 to 115 kHz at regular separation (channel number: 0 at 115 kHz to 18 at 25 kHz). In our model, an 18-kHz integrated filter channel bandwidth is obtained by integrating over 11 bandpass filter channels. The Q_{10dB} values range from 1.4 at 115 kHz to 6.3 at 25 kHz. This integrated filter (IF) in the nth frequency channel is described by the equations:

$$\text{IF}_{\text{on}}(f_n, t) = \prod_{i=-5}^{5} \text{B}_{\text{on}}(f_n + \text{FD} \times i, t - SR \times i), \tag{3}$$

$$\text{IF}_{\text{off}}(f_n, t) = \prod_{i=-5}^{5} \text{B}_{\text{off}}(f_n + \text{FD} \times i, t - SR \times i), \tag{4}$$

where f_n(kHz) is the center frequency for the nth integrated filter channel and t is time (in seconds). The quantity FD ($=1$kHz) is the frequency difference between bandpass filter channels.

The transform of the time axis into the range axis is carried out by the coincidence circuit, which calculates the cross-correlation function between the emission and the echoes. This cross-correlation function for the nth integrated filter channel is represented by the following equations:

$$\text{C}_{\text{on}}(f_n, \tau_i) = \int \text{IF}_{\text{on}}(f_n, t + \tau_i)\text{IF}_{\text{on}}(f_n, t)\, dt, \tag{5}$$

$$\text{C}_{\text{off}}(f_n, \tau_i) = \int \text{IF}_{\text{off}}(f_n, t + \tau_i)\text{IF}_{\text{off}}(f_n, t)\, dt, \tag{6}$$

$$\tau_i = \tau_0 + i\,d\tau, \quad i = 0,1,2,3,...,70, \tag{7}$$

where the value τ_i is the delay time corresponding to the ith coincidence-detecting cell, and the quantity $d\tau$ is the difference between delay times corresponding to consecutive coincidence-detecting cells. In this model, we set $d\tau = 60$ μs (which corresponds to a distance 20.4 mm along the range axis), because the interval between BDs of delay-tuned neurons in the CF/FM bat *Pteronotus parnellii* is about 100 μs (Suga and O'Neill, 1979).

The final outputs of the temporal block (O_{on} and O_{off}) are calculated by summing up the outputs of 19 integrated filter channels at each delay time:

$$\text{O}_{\text{on}}(\tau_i) = \sum_{f_n} \text{C}_{\text{on}}(f_n, \tau_i), \tag{8}$$

$$\text{O}_{\text{off}}(\tau_i) = \sum_{f_n} \text{C}_{\text{off}}(f_n, \tau_i). \tag{9}$$

The differences between the onset and offset delay times corresponding to the distances to the nearest and farthest targets, for a group of the echoes returning from single or multiple nearly equidistant targets, are determined from O_{on} and O_{off}.

F. Examples of outputs from the temporal block

The solid curve in Fig. 3(c) represents the output resulting from the onset pattern along the range axis. The peak delays are 4298 and 5858 μs, corresponding to 1.461 and 1.992 m along the range axis. The dotted curve in Fig. 3(c) represents the output resulting from the offset pattern along the range axis. The peak delays are 4298 and 6098 μs, corresponding to 1.461 and 2.073 m. Thus, the first group of echoes corresponds to 1.461 m and the second group from 1.992 to 2.073 m along the range axis. We thus see that important information about the target distances can be obtained from the onset and offset patterns.

G. The spectral block

The function of the spectral block is to determine the differences among delay times of echoes from nearly equidistant targets and the EEDNF associated with each target. For this reason, it needs to distinguish between the S_{target} and the S_{ext_ear} in the spectrogram resulting from echoes returning from nearly equidistant targets. From the measurement of the external ear's transfer functions of real bats, it has been found that the EEDNF varies within the somewhat restricted frequency range (Wotton *et al.*, 1995). While in the EEDNF region the spectrogram includes contributions from both the S_{target} and the S_{ext_ear}, in the non-EEDNF region, the spectrogram contains only the S_{target}. For this reason, when the deconvolution of the basis vectors (see the Appendix) is applied to the echo spectral pattern (S_{echo}) in the non-EEDNF

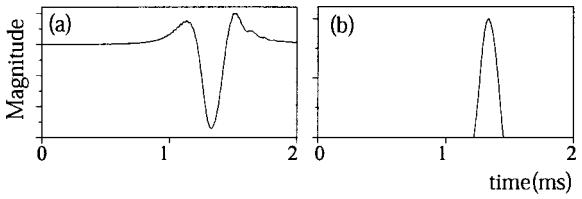


FIG. 4. Diagram describing the extraction of the peak pattern. (a) The second derivative calculated by extracting the change in the magnitude of the first derivative [see Fig. 3(b)] along the time axis. (b) The peak pattern is the negative part of the second derivative.

region, this is transformed into an impulse response resulting only from the interference between echoes from targets. The difference between delay times can be determined from the peak time in this impulse response along the time axis. To extract the S_{ext_ear} we need to determine the S_{target} in the EEDNF region. However, the S_{target} in the EEDNF region cannot be reconstructed from this impulse response which is calculated from the S_{echo} in the non-EEDNF region. Since the S_{target} is dependent on the difference between delay times, the S_{target} in the EEDNF region can be extrapolated from information on the difference between delay times calculated from the S_{echo} within the non-EEDNF region and then the S_{ext_ear} can be determined by subtracting the S_{target} from the S_{echo}.

The spectrograms of the emission and the echoes returning from targets are transformed into the range-frequency representation by using the cross-correlation function between the emission and the echoes in each bandpass filter channel. In order to discriminate echoes from nearly equidistant targets, the spectrum of the echo associated with each target must be extracted more precisely. Therefore, it needs to improve the temporal precision of the emission's pattern. For this reason, the second derivative of the spectrogram is calculated by extracting the change in the magnitude of the first derivative of the envelope pattern [see Fig. 3(a)]. As shown in Fig. 4(b), the peak pattern is the negative part of the second derivative of the spectrogram. Since the width of the peak pattern is shorter than the width of the envelope pattern, the range-frequency pattern is obtained from the cross-correlation function between the emission's peak pattern (B_{peak}) and the echo's envelope pattern (B_{env}) at the nth bandpass frequency channel as follows:

$$C(f_n,\tau_i)=\int B_{peak}(f_n,t+\tau_i)B_{env}(f_n,t)\,dt,$$
$$\tau_i=\tau_0+id\tau,\quad i=0,1,2,3,...,70. \tag{10}$$

Here $C(f_n,\tau_i)$ is the cross-correlation function between an emission and its resulting echoes. The quantity τ_i is the delay time, and $d\tau$ is the time interval between delay times to coincidence-detecting cells. These values correspond to the values for the cross-correlation function in the temporal block.

The difference between delay times of echoes from nearly equidistant targets is calculated from the S_{echo} associated with the average (τ_{ave}) of delay times corresponding to the onset and offset for a group of echoes, since the S_{echo} associated with the average delay time is strongly influenced

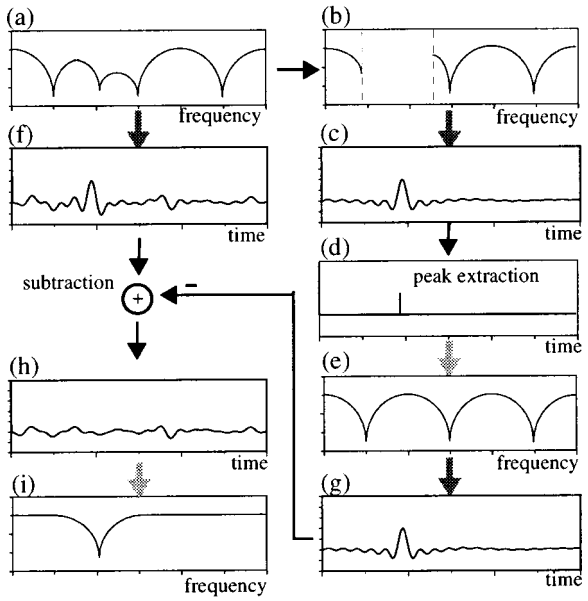


FIG. 5. Diagram of the spectral block. (a) The S_{echo} includes contributions from the interference between echoes from nearly equidistant targets and the interference created by the physical process of reception at the external ear. (b) The S_{echo} associated with the point half-way between the onset and offset delay times in non-EEDNF region includes a contribution only from the interference between echoes from nearly equidistant targets. (c) The S_{echo} associated with the point half-way between the onset and offset delay times in the non-EEDNF region is transformed into the I_{sel_pass} through the deconvolution of the selective-pass basis vectors. (d) The reduced impulse response is represented by extracting the values of the I_{sel_pass} at $t=0$ and t = peak time equal to the difference between delay times. (e) The reduced impulse response is transformed into the S_{target} through the convolution of the all-pass basis vectors. (f) The S_{echo} associated with the delay time corresponding to each target is transformed into the associated I_{echo} through the deconvolution of the all-pass basis vectors. (g) The S_{target} is transformed into the I_{target} through the deconvolution of the all-pass basis vectors. (h) The subtracted impulse response is calculated by subtracting the I_{target} from the I_{echo}. (i) The subtracted impulse response is transformed into the S_{ext_ear} through the convolution of the all-pass basis vectors. The EEDNF corresponding to each target is the frequency corresponding to the minimum of the amplitude in the S_{ext_ear}.

by the interference of the echoes returning from nearly equidistant targets. The S_{echo} associated with the average delay time is obtained as the logarithm of the cross-correlation function $\log(C(f_n,\tau_{ave}))$ using the interpolation. Since the S_{echo} of the average delay time in the non-EEDNF region only includes a contribution from the interference between echoes from nearly equidistant targets, the S_{echo} of the average delay time in this region [Fig. 5(b)] can be transformed into the impulse response [Fig. 5(c)] using the deconvolution of the selective-pass improved basis vectors (see the Appendix). This impulse response is called the ''selective-pass impulse response (I_{sel_pass}).'' The difference between delay times of echoes is determined from the peak time of the I_{sel_pass}.

To determine the S_{ext_ear} in the EEDNF region requires extrapolating the S_{target} in the EEDNF region from the S_{echo} in the non-EEDNF region. For this reason, we extract the values of the I_{sel_pass} at $t=0$ and $t=$peak time equal to the difference between delay times. The extracted impulse response [Fig. 5(d)] is called the ''reduced impulse response.''

The reduced impulse response is transformed into the S_{target} [Fig. 5(e)] in the entire frequency region using the convolution of the all-pass basis vectors (see the Appendix). If the shape of the S_{target} is similar to that of the S_{echo} at the average delay time in the non-EEDNF region, the difference between delay times of echoes from nearly equidistant targets can be considered to be the peak time.

The EEDNF corresponding to each target can be determined by comparing the S_{target} and the S_{echo} associated with the delay time for the target in question. The delay time for the nearest target is equal to that of the onset of the group of echoes. The delay time for the farthest target is determined by adding the delay time for the nearest target and the difference between delay times of the echoes from the nearly equidistant targets. The S_{echo} associated with the delay time (τ) for each target is calculated as $\log(C(f_n,\tau))$. The amplitudes of peak and notch frequencies for the S_{target} are shifted to coincide with those of peak and notch frequencies at the S_{echo} associated with the target [Figs. 5(a) and (e)]. The S_{echo} associated with each target and the S_{target} [Figs. 5(a) and (e)] are transformed into the echo and target dependent impulse responses (I_{echo} and I_{target}), respectively [Figs. 5(f) and (g)], through the deconvolution of the all-pass improved basis vectors. In order to eliminate the component relating to the difference in delays ($:=t_{peak}$), it needs to correct the I_{target}. For this reason, I_{target} is corrected in accordance with the stipulation that $I_{target}(t_{peak})$ coincide with $I_{echo}(t_{peak})$ in Eq. (11) and this corrected impulse response is called ''the corrected target dependent impulse response (I_{cor_tar}):''

$$I_{cor\_tar}(t)=\frac{I_{echo}(t_{peak})}{I_{target}(t_{peak})}I_{target}(t). \tag{11}$$

As shown in Fig. 5(h), the external ear dependent impulse response (I_{ext_ear}) can be determined by subtracting the I_{cor_tar} from the I_{echo} as follows:

$$I_{ext\_ear}(t)=I_{echo}(t)-I_{cor\_tar}(t). \tag{12}$$

The I_{ext_ear} is transformed into the S_{ext_ear} [Fig. 5(i)] through the convolution of the all-pass improved basis vectors (see the Appendix). The EEDNF associated with each target is the frequency corresponding to the minimum amplitude of the S_{ext_ear}. In this way, the difference between delay times for each target and the EEDNF associated with each target are obtained as outputs of the spectral block.

H. Examples of outputs from the spectral block

Figure 6 exhibits outputs of the selective-pass impulse response associated with the first and second group of echoes along the range axis [see Fig. 3(c)]. Figure 6(a) displays the S_{echo} of the average delay time for the first group of echoes, whose distance is 1.461 m. The spectral pattern in the non-EEDNF region has almost a constant amplitude in the non-EEDNF region, and the I_{sel_pass} is almost constant, except for the existence of a peak at $t=0$ along the time axis, as shown in Fig. 6(b). It was found that the first group of echoes results from a single target. The S_{ext_ear} is identical to the S_{echo} here, because in this case the spectral pattern includes only the

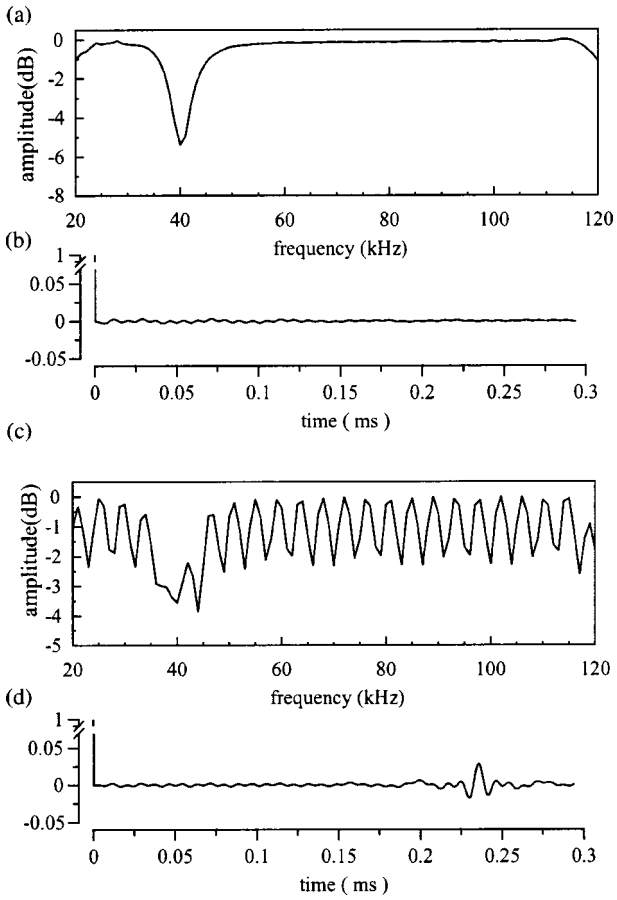


FIG. 6. Determination of the difference between delay times. (a) The S_{echo} associated with the point half-way between the onset and offset delay times in the first group of echoes. The point half-way between the onset and offset delay times corresponds to 1.461 m along the range axis. This S_{echo} has a constant amplitude in the non-EEDNF region. (b) I_{sel_pass}. The I_{sel_pass} is normalized by the value corresponding to $t=0$ along the time axis. The I_{sel_pass} is almost constant along the time axis, except for the peak at $t=0$. It is thus concluded that the first target region contains a single target. (c) The S_{echo} associated with the point half-way between the onset and offset delay times (corresponding to 2.033 m) in the second group of echoes. (d) I_{sel_pass}. The difference between delay times is 235 μs, corresponding to 80 mm along the range axis.

interference created by the physical process of the reception at the external ear. The calculated EEDNF to the target at each ear is 40 kHz. Figures 6(c) and (d) display the S_{echo} of the average delay time for the second group of echoes and the I_{sel_pass}. The difference between delay times is 235 μs, corresponding to 80 mm along the range axis, as obtained from the peak time of the I_{sel_pass}.

The solid curve in Fig. 7(a) represents the S_{echo} associated with the nearest target and the dotted curve represents the S_{target}. The S_{echo} associated with the nearest target differs slightly from the S_{target} because the peak and notch frequencies change along the range axis. Figure 7(b) displays the I_{echo} and I_{cor_tar}, and Fig. 7(c) displays the subtracted impulse response. The solid curve in Fig. 7(d) represents the S_{ext_ear}, and the calculated EEDNF is 40 kHz. The dotted curve in this figure represents the spectral pattern (S_{sub}) which is obtained by subtracting the S_{target} [Fig. 7(a) dotted line] from

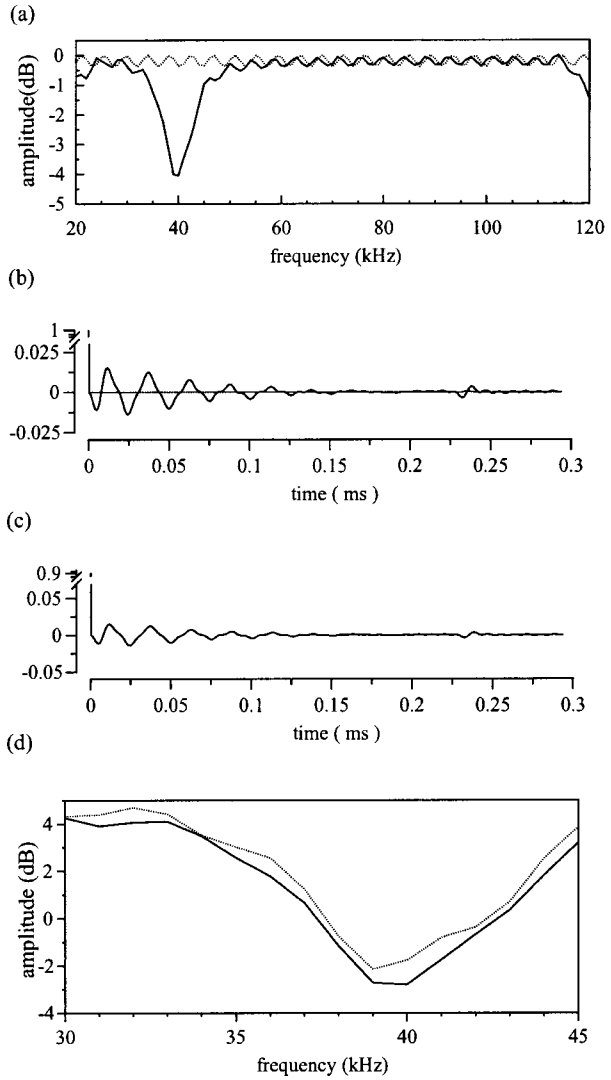


FIG. 7. Determination of the EEDNF associated with the nearest target corresponding to the second group of echoes. (a) The S_{echo} associated with the delay time of the nearest target (solid curve) and the S_{target} (dotted curve). (b) The I_{echo} (solid curve) and the I_{target} (dotted curve). (c) Subtracted impulse response. (d) S_{ext_ear} are obtained from the subtracted impulse response calculated in Eq. (12) (solid curve), and the spectral pattern that is obtained by subtracting the S_{target} from the S_{echo} associated with the nearest target along the frequency axis in Eq. (13) (dotted curve). The value of the EEDNF calculated by subtracting these impulse responses along the time axis is 40 kHz. In contrast, the value of the EEDNF calculated by subtracting along the frequency axis is 39 kHz.

the S_{echo} associated with the nearest target [Fig. 7(a) solid line] as follows:

$$S_{sub}(f) = S_{echo}(f) - S_{target}(f). \tag{13}$$

In this case, the calculated EEDNF is 39 kHz. The difference between the two values of the EEDNF indicates that its value for each target can be more accurately obtained by using the impulse responses along the time axis rather than using the S_{echo} and the S_{target} along the frequency axis.

The solid curve in Fig. 8(a) represents the S_{echo} associated with the farthest target, and the dotted curve represents the S_{target}. The S_{target} is similar to the S_{echo} associated with the farthest target in the non-EEDNF region. As shown in Fig. 8(b), the I_{echo} corresponds to the I_{cor_tar} around the time equal

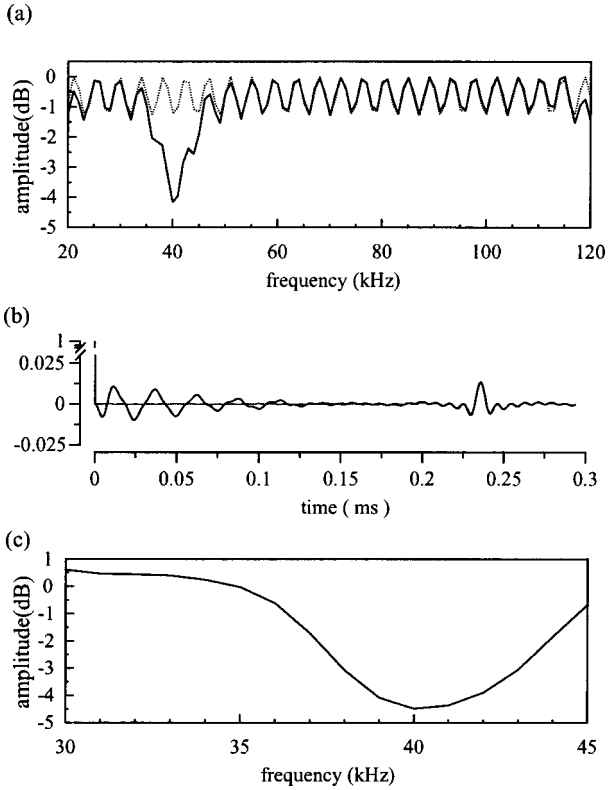


FIG. 8. Determination of the EEDNF associated with the farthest target corresponding to the second group of echoes. (a) The S_{echo} associated with the delay time of the farthest target (solid curve) and the S_{target} (dotted curve). (b) The I_{echo} (solid curve) and the I_{target} (dotted curve). (c) S_{ext_ear} calculated from the subtracted impulse response. The calculated EEDNF is 40 kHz.

to the difference between delay times, 235 μs. The calculated EEDNF is 40 kHz, as shown in Fig. 8(c).

I. The integration block

In the binaural integration block, each target in 3D space is located using two pieces of information regarding the distance to the target and the EEDNF associated with the target. The distance to the nearest target in the target region can be obtained directly from the delay time corresponding to the onset, which is given as an output of the temporal block. The distance to the farthest target is determined by adding the distance to the nearest target and the distance corresponding to the difference between delay times of the echoes from the two nearly equidistant targets. Each curve in Fig. 9(a), drawn in a horizontal plane, represents equal distances. Each curve corresponds to the distance (calculated in the manner described above) from the mouth to one of the ears, via one of the targets. Four possible locations are determined for each target's position in a horizontal plane by using the interaural range difference (IRD), which is equivalent to the ITD. However, it is difficult to precisely determine the target's azimuthal direction by using only the information on the IRD between nearest targets (IRD_{near}) because of the smallness of the bat's head. By using the directional information concerning the EEDNF associated with each target, our model can uniquely determine the location of each target in 3D space from the possible positions determined in the manner previ-

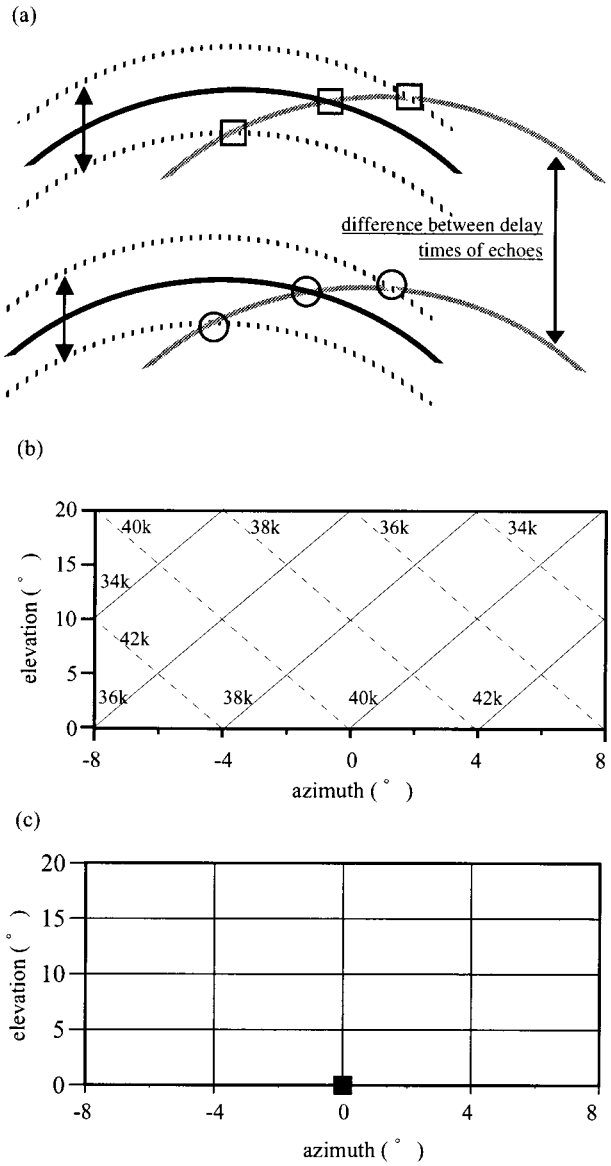


FIG. 9. Determination of the target direction. (a) The curves represent equal from mouth-to-ear distances via two targets in a horizontal plane. The black curves correspond to the left ear and the gray curves correspond to the right ear. Four possible locations are determined for each target's position in a horizontal plane by using the interaural range difference (IRD). The IRD between nearest targets (IRD_{near}) varies from -16 to 16 μs (corresponding to azimuthal angles from $-8°$ to 8°) as results from the head size. In accordance with the IRD_{near}, the IRD_{far} determines possible locations for the farthest target in a horizontal plane. (b) EEDNF as a function of sound direction. The solid and dashed lines represent the EEDNF for the left and right ears, respectively. The EEDNF changes according to some fixed prescription in both azimuthal and elevational directions. (c) Final output of the integration block. The direction of each target (block square) is determined as $(\theta,\phi)=(0°,0°)$.

ously described. The direction to a target is the direction whose notch frequency in the template set is the same as the calculated EEDNF. Thus, the target direction can be calculated by requiring that the difference between the notch frequency in the template set and the calculated EEDNF be the minimum.

Figure 9(c) displays outputs of the integration block. From this, we see that the direction to the target in the first group of echoes is $(\theta,\phi)=(0.0°,0.0°)$ as determined from the EEDNFs (=40 kHz) associated with the two ears [see Fig. 9(b)]. Also, the results indicate that the second group of echoes has two targets. From the difference between delay times (=235 μs) for the two ears, the azimuthal direction associated with the nearest target is found to be the same as the azimuthal direction of the farthest target. By using the EEDNF (=40 kHz) corresponding to each target, two target's direction in the second group of echoes is determined to be $(\theta,\phi)=(0.0°,0.0°)$. From these results, we conclude that our model can accurately locate the three targets.

TABLE I. The distances and EEDNFs corresponding to five actual targets.

Target's location (r,[a] θ,[b] ϕ[b])	Distance[a] (Delay time[c])		EEDNF[d]	
	Left ear	Right ear	Left ear	Right ear
(0.734, 6, 0)	1.470 (4325)	1.466 (4312)	43	37
(0.850, 0, 0)	1.700 (5001)	1.700 (5001)	40	40
(1.064, 0, 20)	2.129 (6260)	2.129 (6260)	36	36
(1.201, −8, 15)	2.400 (7060)	2.400 (7076)	33	41
(1.351, 2, 0)	2.702 (7949)	2.701 (7944)	41	39

[a]The distance unit is meters.
[b]The unit is degrees.
[c]The unit is microsecond.
[d]The unit is kHz.

III. RESULTS

A. Multiple targets without interference among echoes from different targets

As an example of the case of an arbitrary number of targets in 3D space, we consider that in which there exist five targets and the echoes returning from them overlap but do not interfere with each other. The locations of the five targets are $(r,\theta,\phi)=(0.734\,\text{m},6°,0°)$, (0.850 m, 0°, 0°), (1.064 m, 0°, 20°), (1.201 m, −8°, 15°), and (1.351 m, 2°, 0°). The distances from the mouth to two ears via the five targets and the EEDNF corresponding to each target at each ear are listed in Table I.

The waveforms and spectrograms of the emission and the echoes returning from the five targets are displayed in Figs. 10(a) and (b). We see from these figures that the echoes from the targets overlap but do not interfere with each other. Figures 10(c) and (d) display the outputs of the temporal block. The five clear peaks of each curve indicate the delay times for the five targets. These peaks allows us to determine the distances from the mouth to two ears via each target, as listed in Table II. Comparing Table II to Table I, it is clear that the model of the temporal block can closely determine the distances to the targets.

In the present case, the echoes from the five targets overlap but do not interfere with each other, so the differences between delay times corresponding to the different targets is on the order of several hundred to several thousand microseconds. Therefore, the echo returning from each target can be easily separated, and the two EEDNFs for each of the five targets can be calculated from the S_{echo} associated with each target. Each S_{echo} is almost constant in the non-EEDNF region, and the I_{sel_pass} corresponding to the first target for the left ear has no peaks along the time axis, except for a peak at $t=0$, as shown in Fig. 10(f). This means that only one target

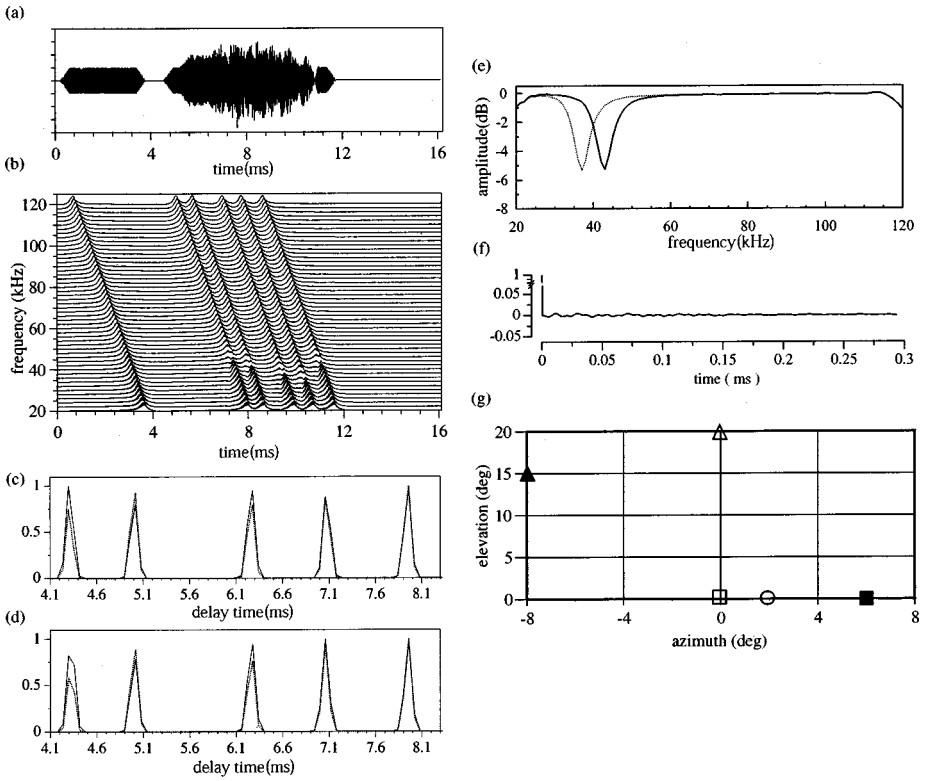


FIG. 10. Multiple targets with no interference among echoes from targets. Five targets are located at $(r,\theta,\phi)=(0.734\text{ m},6°,0°)$, (0.850 m, 0°, 0°), (1.064 m, 0°, 20°), (1.201 m, −8°, 15°), and (1.351 m, 2°, 0°). (a) Waveforms of the emission and the returning echoes. (b) Spectrogram of the emission and the echoes. The echoes from the five targets overlap but do not interfere with each other. (c) and (d) Outputs of the temporal block for left and right ears, respectively. The solid and dotted curves represent outputs obtained from the onset and offset patterns, respectively. The five clear peaks of each curve indicate the delay times to the five targets. As shown in Table II, the distances from the mouth to the two ears for the targets can be determined from this information. (e) The S_{echo} associated with the point half-way between the onset and offset delay times in the first group of echoes (corresponding to a distance of 1.461 m). (f) I_{sel_pass} for the first group of echoes. Since the I_{sel_pass} has no peaks except a peak at $t=0$ along the time axis, it indicates that the first target region contains only a single target. For this reason, the S_{echo} associated with the delay time of the target is identical to the S_{ext_ear}, and the EEDNF is 43 kHz. (g) The directions of the five targets. The calculated direction (θ, ϕ) is (6.0°, 0.0°) for the first target (solid square), (0.0°, 0.0°) for the second target (open square), (0.0°, 20.0°) for the third target (open triangle), (−8.0°, 15.0°) for the fourth target (solid triangle), and (2.0°, 0.0°) for the fifth target (open circle). Each of these calculated directions is in good agreement with the actual direction.

is contained in the first target region. The EEDNFs obtained for the first target are 43 kHz for the left ear and 37 kHz for the right ear. The EEDNFs were obtained similarly for the other four targets. The results are listed in Table II. Finally, the direction to each target is determined by using these EEDNFs, as shown in Fig. 10(g). These directions (θ, ϕ) are (6.0°, 0.0°), (0.0°, 0.0°), (0.0°, 20.0°), (−8.0°, 15.0°), and (2.0°, 0.0°). These results are in good agreement with the actual target directions.

B. Two targets with the almost same distance but different direction

We now consider the case in which two targets are located at the almost same distance but different directions. The two targets are located at $(r,\theta,\phi)=(0.734\text{ m},6°,0°)$ and (0.768 m, −4°, 20°). The distances from the mouth to the ear via the first target are 1.470 m (=4325 μs) and 1.466 m (=4312 μs) for the left and right ears, respectively, and the

TABLE II. Outputs of the temporal and spectral blocks for five targets.

	Calculated results from the temporal block					
	Onset Distance[b] (Delay time[c])		Offset Distance[b] (Delay time[c])		Calculated EEDNF[d] of the spectral block	
No.[a]	Left ear	Right ear	Left ear	Right ear	Left ear	Right ear
1	1.461 (4298)	1.461 (4298)	1.461 (4298)	1.461 (4298)	43	37
2	1.706 (5018)	1.706 (5018)	1.706 (5018)	1.706 (5018)	40	40
3	2.135 (6278)	2.135 (6278)	2.135 (6278)	2.135 (6278)	36	36
4	2.400 (7058)	2.400 (7058)	2.400 (7058)	2.400 (7058)	33	41
5	2.706 (7958)	2.706 (7958)	2.706 (7958)	2.706 (7958)	41	39

[a]The numbers are in order of the onset distance.
[b]The distance unit is meters.
[c]The unit is microseconds.
[d]The unit is kHz.

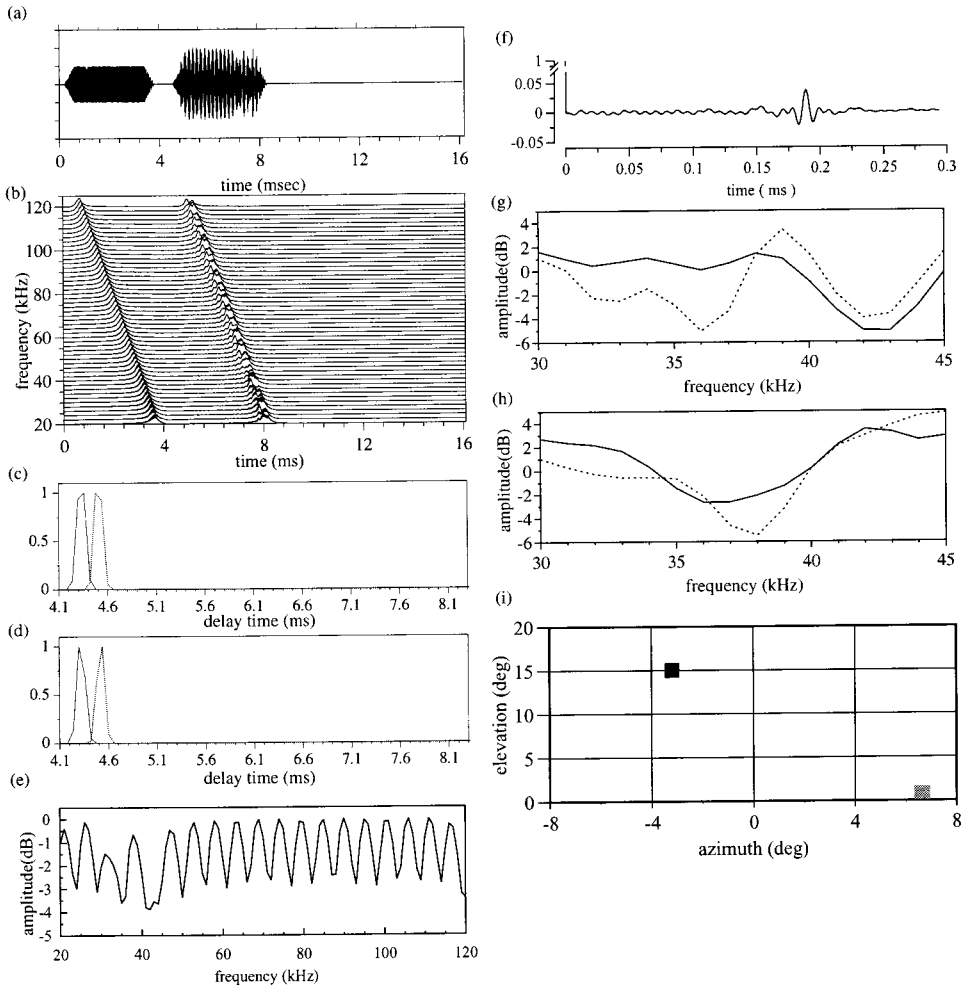


FIG. 11. Outputs in the case of two targets at almost equal distances but different directions. The two targets are located at $(r, \theta, \phi) = (0.734\ \text{m}, 6°, 0°)$ and (0.768 m, −4°, 20°). The delay times associated with the first target are 4325 μs (=1.470 m) for the left ear and 4312 μs (=1.466 m) for the right ear. The delay time associated with the second target are 4514 μs (≔1.535 m) for the left ear and 4521 μs (≔1.537 m) for the right ear. (a) The waveforms of the emission and the returning echoes. (b) Spectrogram of the emission and the echoes. (c) and (d) Outputs of the temporal block for the left ear and the right ear, respectively. The solid curve represents the output calculated from the onset pattern and the dotted curve represents the output calculated from the offset pattern. (c) The delay times corresponding to the peaks of the onset and offset patterns are 4358 μs (=1.482 m) and 4478 μs (=1.522 m). (d) The delay times corresponding to the peaks of the onset and offset patterns are 4298 μs (=1.461 m) and 4538 μs (=1.543 m). (e) The S_{echo} associated with the point half-way between the onset and offset delay times for the left ear. The point half-way between the onset and offset delay times corresponds to 1.502 m along the range axis. (f) I_{sel_pass}. The difference between delay times of echoes from nearly equidistant target is 189 μs. (g) and (h) Determination of the EEDNF for the left and right ears, respectively. The solid and dotted lines display the S_{ext_ear} associated with the nearest and farthest targets. (g) The EEDNFs are 43 kHz for the nearest target and 36 kHz for the farthest target. (h) The EEDNFs are 37 kHz for the nearest target and 38 kHz for the farthest target. (i) Calculated directions of the two targets. The direction of the nearest target (solid square) is (6.8°, 1.0°), and that of the farthest target (gray square) is (−3.2°, 15.0°).

distances from the mouth to the ear via the second target are 1.535 m (=4514 μs) and 1.537 m (=4521 μs) for the left and right ears, respectively. The differences between delay times of the two targets are 189 μs (corresponding to 64.3 mm) for the left ear and 209 μs (corresponding to 71.1 mm) for the right ear. The EEDNFs corresponding to the nearest target are 43 kHz for the left ear and 37 kHz for the right ear. The EEDNFs corresponding to the farthest target are 34 kHz for the left ear and 38 kHz for the right ear.

Figures 11(a) and (b) display the waveforms of the emission and the echoes returning from the two targets and the output of the cochlear block, respectively. Figure 11(c) displays the output of the temporal block for the left ear. The delay times at the peaks for the onset and the offset are 4358 μs (=1.482 m) and 4478 μs (=1.522 m), respectively. As shown in Fig. 11(d), for the right ear, the delay times at the onset and offset are 4298 μs (=1.461 m) and 4538 μs (=1.543 m), respectively. Figure 11(f) displays the I_{sel_pass}. This figure indicates that the average of the distances from the two targets to the left ear is 1.502 m and the difference between delay times for the two echoes returning from the targets is 189 μs (corresponding to 64.2 mm) for the left ear. In the same way, the average distance and the difference in delay times for the right ear are obtained as 1.502 m and 209 μs (corresponding to 71.0 mm). The solid and dotted curves in Fig. 11(g) for the left ear and Fig. 11(h) for the right ear represent the S_{ext_ear} for the nearest and farthest target, respectively. From the left S_{ext_ear}, the EEDNFs are obtained as 43 kHz for the nearest target and 36 kHz for the farthest one, the latter of which differs slightly from the actual value of 34 kHz. Similarly, at the right ear, the EEDNFs are found

to be 37 kHz for the nearest target and 38 kHz for the farthest target, which are consistent with their actual values. Integrating these results, we obtain the direction to the nearest target as (6.8°, 1.0°) and to the farthest target as (−3.2°, 15.0°). There are slight discrepancies between these values and the actual values. However, this discrepancies are acceptable, because the bats appear to determine the direction of an echo's source to within $\pm 2° - \pm 5°$ (Grinnell, 1995; Masters *et al.*, 1985; Simmons *et al.*, 1995a).

C. Resolution of two very closely located targets

It has been reported that the FM bat *Eptesicus fuscus* can accurately distinguish between two targets with a difference between delay times as small as 2.6 μs (Saillant *et al.*, 1993; Simmons *et al.*, 1995a). However, they also found that if the cross-correlation function method is used, two glints whose difference between delay times is 10 μs or smaller cannot be discriminated (Saillant *et al.*, 1993).

To study this situation, we consider the case in which two targets are located at $(r, \theta, \phi) = (0.7300\,\text{m}, 0°, 0°)$ and (0.7308 m, 0°, 0°). The distances from the mouth to the ear via the nearest and the farthest target are 1.460 m (=4294.1 μs) and 1.4616 m (=4298.8 μs), Thus, there is the difference between delay times of 4.7 μs (=1.6 mm). Each of the four EEDNFs corresponding to these targets is 40 kHz. Figures 12(a) and (b) display the waveforms and the spectrograms of the emission and the echoes from the two targets. Figure 12(c) exhibits the cross-correlation function for the emission and the echoes from the two targets. As seen, it contains only one peak, at the point half-way between the delay times of the echoes from the two targets. This suggests that it is difficult to determine the difference in delays between delay times for echoes from nearly equidistant targets using only the cross-correlation function. The outputs of the temporal block are shown in Fig. 12(d). This indicates that the two targets are located in the vicinity of 4298 μs (=1.461 m) along the range axis.

Figure 12(e) displays the I_{sel_pass}. The calculated difference between delay times for the two targets is 4.1 μs (=1.4 mm). Although this value differs slightly from the true value, our model has successfully determined that there are two nearly equidistant targets. Each of the four EEDNFs is found to be 40 kHz, as shown in Fig. 12(f). This implies that the directions of both targets are (0.0°, 0.0°) in agreement with the actual values.

D. Robustness in a noisy environment

Echolocating bats can reliably discriminate multiple targets in a noisy environment (Griffin, 1958). The FM bat *Eptesicus fuscus* can determine the arrival time of an echo with a precision of within 40 ns at an echo energy signal-to-noise ratio of 36 dB (Simmons *et al.*, 1990a). To investigate the robustness of our model, we added a broadband white random noise to the echoes, following Saillant *et al.* (1993). The noise is characterized by the ratio of the root mean square (rms) amplitude of the echo of a single point and the rms amplitude of the white noise during a control interval preceding the emission.

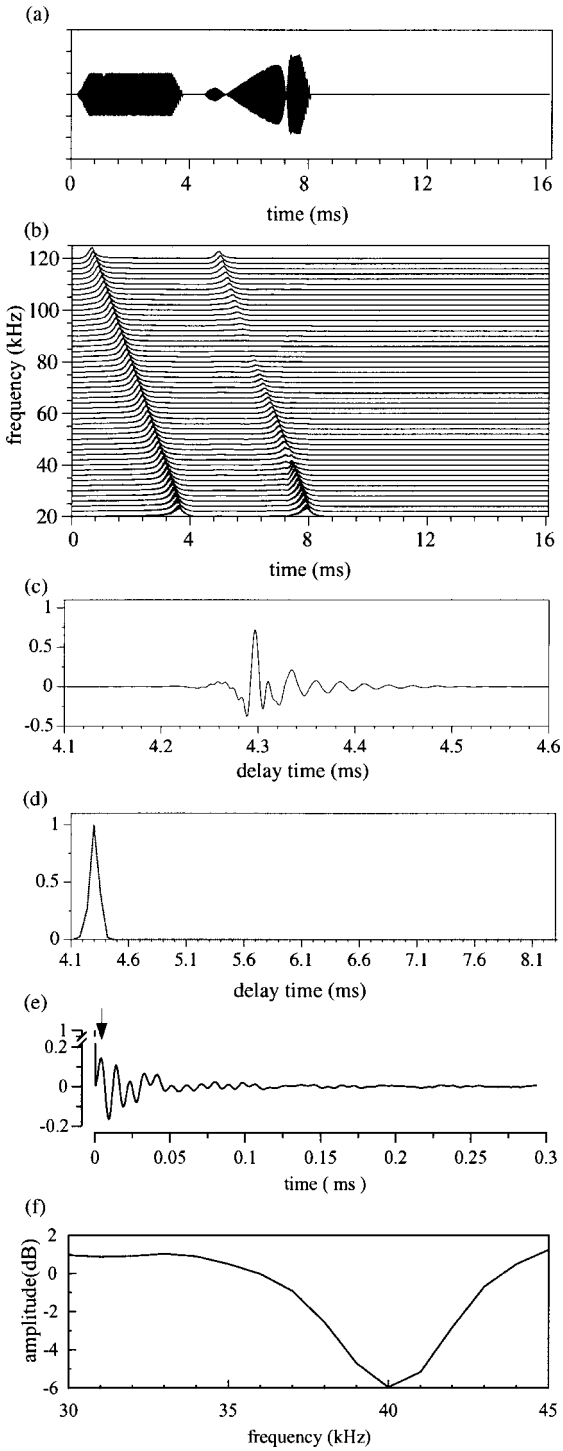


FIG. 12. Resolution of two closely located targets. Two targets are located at $(r, \theta, \phi) = (0.7300\,\text{m}, 0°, 0°)$ and (0.7308 m, 0°, 0°). The delay times are 4294.1 μs (=1.4600 m) for the first target and 4298.8 μs (=1.4616 m) for the second target. The difference between delay times of echoes from the two targets is 4.7 μs (=1.6 mm). (a) The waveforms of the emission and the echoes. (b) Spectrogram of the emission and the echoes. (c) The cross-correlation function between the emission and the echoes. This function has a single peak located at the point half-way between the onset and offset delay times of echoes from the two targets. In this case, it is impossible to determine the difference between delay times of the echoes from very closely positioned targets using only the cross-correlation function. (d) Output of the temporal block. The delay times corresponding to the peaks of the onset and offset patterns are both 4298 μs (=1.461 m). (e) I_{sel_pass}. The difference between delay times is 4.1 μs. The arrow shows the 4.1-μs location. (f) S_{ext_ear} associated with the nearest target for the left ear. The EEDNF is 40 kHz.

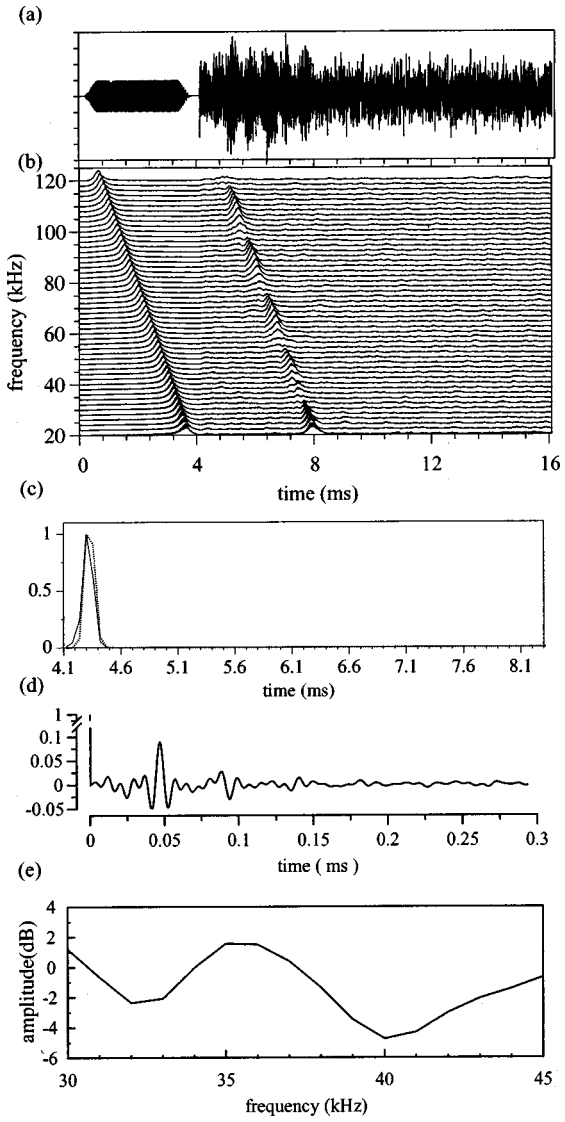


FIG. 13. Robustness in a noisy environment. The signal-to-noise ratio is 5 dB. Two targets are located at $(r,\theta,\phi)=(0.730\,\text{m},0°,0°)$ and (0.738 m, 0°, 0°). The delay times are 4295 μs (=1.460 m) for the first target and 4342 μs (=1.476 m) for the second target. The difference between delay times of echoes from the targets is 47 μs. The EEDNF is 40 kHz for both ears and both targets. (a) The waveforms of the emission and the echoes. (b) Spectrogram of the emission and the echoes. (c) Outputs of the temporal block. The solid curve represents the output calculated from the onset pattern and the dotted curve represents the output calculated from the offset pattern. The delay times corresponding to the peaks of the onset and offset patterns are both 4298 μs (=1.461 m). (d) I_{sel_pass}. The difference between delay times of echoes from the two targets is 46.8 μs (=15.9 mm). (e) S_{ext_ear} associated with the nearest target. The EEDNF for the nearest target is 40 kHz.

In this study, two targets are located at $(r,\theta,\phi)=(0.730\,\text{m},0°,0°)$ and (0.738 m, 0°, 0°). The rms amplitude signal-to-noise ratio for a single target is 5.0 dB. The distances from the mouth to the ear via the nearest and farthest targets are 1.460 m (=4295 μs) and 1.476 m (=4342 μs). The difference between delay times for the two targets is 47 μs (=16 mm). All EEDNFs for the two targets and two ears are 40 kHz.

Figures 13(a) and (b) display the waveforms and spectrograms of the emission and the returning echoes. From the outputs of the temporal block, the delay time for the onset and the offset of the group of echoes are found to be 4298 μs (=1.461 m) and 4298 μs (=1.461 m) for both ears, as shown in Fig. 13(c). Figure 13(d) displays the I_{sel_pass}. The calculated difference between delay times is 46.8 μs (=15.9 mm), which is in good agreement with the actual difference between delay times. Figure 13(e) displays the S_{ext_ear} associated with the nearest target for the left ear. The EEDNF calculated is 40 kHz. The EEDNF associated with each target for each ear was similarly calculated and found to be 40 kHz, in good agreement with the actual values. These results indicate that the direction to each target is (0.0°, 0.0°). We thus find that in spite of the existence of noise, the locations of the targets can be accurately determined.

IV. DISCUSSION AND CONCLUSION

Echolocating bats emit trains of high-frequency sounds and recognize individual targets in the real world. Since bats are able to capture a moving target in the presence of multiple objects, it appears that they may be able to determine the locations of multiple targets using the echoes produced by a single emission. In order to locate objects in 3D space, bats need to process the returning echoes to determine the distance and direction to each object.

The model to determine the distance and direction of moving, multiple targets by using both the Doppler shift and amplitude change of CF echoes was proposed (Müller and Schnitzler, 2000). However, it is difficult to discriminate echoes from closely spaced targets and to localize these targets because of the interference between echoes. In the field of the radar and sonar, for example, Liu and Xiang (1999) proposed the computational algorithm to determine the distances of unknown, multiple targets. However, this cannot determine the location of targets in 3D space because of using the information on only the distance to targets. Also, Swindlehurst and Gunther (1999) proposed the computational algorithm to localize multiple targets in 3D space by using the receiving array. However, in the case of restricting two receivers, it is hard to localize multiple targets in 3D space.

In this model, we have introduced four blocks applying different processes to analyze the information obtained from echoes to locate multiple targets in 3D space. The waveforms of the emission and the returning echoes are processed in the cochlear block, where they are transformed into a spectrogram. In the temporal block, the distance to the nearest and farthest targets of a set of targets is determined from the delay times from the emission to the onset and offset of the group of returning echoes. In the spectral block, the range-frequency pattern, which is obtained from the spectrogram using the cross-correlation function, is divided into the S_{target} and the S_{ext_ear}. The former contains information regarding the fine structure within a group of several closely positioned targets and the latter contains information regarding their directions. Finally, the outputs of the temporal block and the spectral block are integrated in the integration block, and then the location of each target in 3D space is thus determined.

When multiple targets are located at distances which differ enough that the differences between the delay times of their returning echoes are longer than the duration of the emission, 3.6 ms, the correspondence between echoes and targets can be easily discriminated and these echoes can be processed independently to locate each target. The difficult problem is to locate multiple targets whose distances are close enough that differences between the delay times of their echoes are shorter than the duration of the emission. In this case, the echoes from the multiple targets overlap. This situation can be classified into two cases, that in which the difference between delay times is longer than the integration time of the bandpass filter, and that in which it is not. In the cochlear block, the waveforms of the returning echoes are transformed into the spectrogram by the bandpass filter, the half-rectification, and the low-pass filter with an integration time of about 350 μs. When the difference between delay times of echoes is longer than the integration time of the filter, each envelope pattern in the bandpass filter channel can be separated along the time axis. In this case, the distance and the EEDNF for each target can be calculated from each separated S_{echo}. As an example of such a situation, five targets were successfully located, as shown in Fig. 10.

By contrast, when the targets are so close that the difference between their corresponding delay times is shorter than the integration time, within each filter channel the envelope patterns produced by these nearly equidistant targets cannot be separated along the time axis. The difference between delay times of echoes from nearly equidistant targets is determined from the S_{echo} associated with the average of the delay time corresponding to the onset and the offset of the group of interfering echoes. The S_{echo} associated with the average delay time is calculated from the cross-correlation function $(C(f_n, \tau_i))$. We divide the S_{echo} into two regions, the EEDNF and the non-EEDNF regions, along the frequency axis. The S_{echo} in the EEDNF region consists of both the S_{target} and the S_{ext_ear} while the S_{echo} in the non-EEDNF region contains only S_{target}. As shown in Fig. 6(a), the difference between delay times of the echoes returning from nearly equidistant targets can be accurately determined from the I_{sel_pass}. Since the I_{sel_pass} is not influenced by the interference caused by the physical process of reception at the external ear, the S_{target} in the entire frequency region can be constructed by using the I_{sel_pass} of the S_{echo} associated with the average delay time in the non-EEDNF region as shown in Figs. 7(a) and 8(a).

Determination of the target direction requires accurate determination of the EEDNF corresponding to each target. The delay times corresponding to targets, indicating the target distances, are determined from the delay time for the onset of the interfering echoes and the difference between delay times of the echoes. The S_{echo} associated with each target is calculated from the cross-correlation function $(C(f_n, \tau_i))$. The S_{ext_ear} corresponding to each target is calculated by eliminating the contribution of the S_{target} from the S_{echo} associated with that target. The contribution of the interference between the echoes from nearly equidistant targets to the S_{echo} associated with one of these targets decrease as the distance between targets increases. Because the cross-correlation function $(C(f_n, \tau_i))$ is discretized in time in steps of 60 μs in this model, we classify arrangements of targets into two cases, one in which the difference between delay times of echoes from two targets is longer than 120 μs and one in which it is shorter. When the difference between delay times is shorter than 120 μs, the peak and notch frequencies in the S_{target} coincide with the peak and notch frequencies in the S_{echo} associated with each target in the non-EEDNF region. Therefore, the EEDNF can be determined by subtracting the I_{cor_tar} from the I_{echo} along the time axis. In contrast, when the difference between delay times is longer than 120 μs, the S_{echo} associated with each target is weakly influenced by the interference between echoes. However, in this case the notch frequencies in the S_{target} are corrected from the notch frequencies in the S_{echo} associated with each target as a result of the difference between the average delay time and the delay time corresponding to each target. As shown in Fig. 7(a), the notch frequencies in the S_{echo} associated with the nearest target differ by 1 kHz from the notch frequencies in the S_{target}. Thus, when the EEDNF is determined by subtracting the S_{target} from the S_{echo} associated with the target along the frequency axis in Eq. (13), the calculated EEDNF corresponding to the nearest target is 39 kHz, which differs slightly from the actual EEDNF (=40 kHz). In our model, the EEDNF can be obtained by correcting the I_{target} in Eq. (11) and then by subtracting the I_{cor_tar} from the I_{echo} along the time axis in Eq. (12). As shown in Fig. 7(d), the calculated EEDNF (=40 kHz) is in good agreement with the actual EEDNF. We thus conclude that since in the case of nearly equidistant targets located along the same direction the EEDNF corresponding to each target can be accurately determined using subtraction along the time axis as previously described, this method is useful to determine the EEDNF of each target in this case.

In the case of nearly equidistant targets, the S_{echo} associated with the nearest target is not influenced by the S_{ext_ear} associated with the farthest target. When two targets are located in the same direction, the S_{echo} associated with the farthest target is influenced slightly by the S_{ext_ear} associated with the nearest target but more strongly by the S_{target}. When two targets are located in the different direction, the S_{echo} associated with the farthest target is influenced strongly from the S_{ext_ear} associated with the nearest target. Thus, in the case of different directions, it is difficult to accurately determine the EEDNF associated with the farthest target. As shown in Fig. 11 corresponding to this case, the calculated EEDNF of the nearest target agrees with the actual EEDNF, but the calculated EEDNF of the farthest target for the left ear is different from the actual EEDNF. It is thus found that the EEDNF of the nearest target can be accurately determined, while the error in the determination of the EEDNF of the farthest target is 2 kHz. Thus, the direction of the nearest target can be determined by using only the EEDNF. By contrast, to determine the azimuthal direction of the farthest target, both the EEDNF and the IRD must be used.

FM bats can distinguish the structure of distances of two targets which are separated by only a few microseconds

(Saillant *et al.*, 1993; Simmons, 1993; Simmons *et al.*, 1995a). When a spectrum with constant amplitude from 20 to 120 kHz is transformed through the deconvolution of the basis vector (see the Appendix), the resulting impulse response has sidelobes everywhere except at time zero. Since in the case of nearly equidistant targets the impulse response includes contributions from both the interference between echoes from nearly equidistant target and the sidelobes, the difference between delay times cannot be determined. In our model, an improved set of basis vectors is used to eliminate these sidelobes. As a result, in our model the difference between delay times can be determined from the I_{sel_pass} using this method; in the case of two targets which are separated by only 4.7 μs, the calculated difference between delay times is 4.1 μs, as shown in Fig. 12(d). We thus find that the difference between delay times can be determined within an accuracy of about 1 μs, and the resolution of the difference between delay times is a few microseconds. As shown in Fig. 12(c), none of the target delay times can be determined by using the cross-correlation function between waveforms of the emission and the returning echoes. The largest peak appears at the point half-way between the delay times of the nearest and farthest targets, and the small peaks are caused by the physical process of the reception at the external ear. However, it is impossible to discriminate between the components of these peaks that are due to the interference of echoes from nearly equidistant targets and the interference caused at the external ear, since the target direction is not known in advance. Therefore, in this case the location of the targets cannot be determined by using only the cross-correlation function.

In a noisy environment, we found that the difference between delay times can be accurately determined with a signal-to-noise ratio of 5 dB [Fig. 13(d)]. Since the difference between delay times can be accurately determined with an error of less than 1 μs through the deconvolution of the selective-pass improved basis vectors, the calculated difference between delay times can be used as crucial information for the determination of the azimuthal direction.

In the temporal block, the envelope pattern can be transformed into the onset and offset patterns using a sine-wave asymmetric filter. The onset and offset patterns both have a peak, the former at the time corresponding to the beginning of the envelope pattern and the latter at the time corresponding to the end, because the transform extracts the derivative of the envelope pattern, with ripples removed. Also, the integrated patterns have a sharp peak along the time axis [Fig. 3(b)]. While the widths of both the onset and offset patterns at 50% of peak height are approximately 150 μs along the time axis, for the integrated patterns, these widths are approximately 30 μs [Fig. 3(b)]. Thus, the integrated onset and offset patterns allow for more accurate determination of the delay times, because the width of the peaks is reduced by a factor of about $\frac{1}{5}$. This width at half peak height along the time axis equals the standard deviation of the spike latency of phasic neurons in the VNLLc (Covey and Casseday, 1991). Therefore, it seems possible that FM bats carry out something equivalent to an integration processing in band-pass filter channels. Also, we used a value of 60 μs for the time interval between coincidence-detecting cells, because this value is close to that found for the interval between BDs of delay-tuned neurons in bats, 100 μs (Suga and O'Neill, 1979). Thus, in our model, the distances to targets can be determined within a precision of 60 μs by using only the delay times corresponding to the peaks of the onset and offset patterns. To determine the distances to the targets producing interfering echoes at higher resolution requires interpolation or population coding.

In this model, the most important mechanism is the method to discriminate the S_{target} and the S_{ext_ear} from the S_{echo}. Since bats can capture the moving target in the presence of multiple objects, it is thought that this method is actually used by FM bats in the behavioral studies. However, it has been unknown that the S_{target} and the S_{ext_ear} are represented in the bat's brain in the physiological study. For this reason, it is needed to verify the existence of the neurons which can respond only the S_{target} or the S_{ext_ear}. The existence of neurons which can respond only the S_{target} may be verified by the following two experiments. The first one is to indicate that the responses of neurons depend on the difference between delay times when the difference between delay times is altered and directions of these targets are not altered. The second one is to indicate that the responses of neurons do not change when the difference between delay times is not altered and directions of these targets are altered. Contrastingly, the existence of neurons which can respond only the S_{ext_ear} may be verified by two experiments in the same way. Also, it requires the external ear's transfer function to verify the existence of these neurons.

ACKNOWLEDGMENT

This study is supported by ''Research for the Future'' Program (JSPS_RFTF97I00103) of Japan Society for the Promotion of Science.

APPENDIX: IMPROVED BASIS VECTORS

The echoes returning from two nearly equidistant targets interfere with each other. The patterns of peaks and notches in the spectra of the echoes depend on the difference between delay times of the echoes from the two targets. The peak and notch frequencies are given by

$$f_{\text{peak}}=\frac{N}{d}, \quad f_{\text{notch}}=\frac{(2N+1)}{d},$$

where d (in seconds) is the difference between delay times of the echoes from the two targets along the time axis and N $(=0,1,2,...)$ is an integer.

Cepstral analysis consists of the inverse Fourier transform of the logarithmic magnitude spectrum into the time domain representation (Oppenheim and Schafer, 1975). This time domain representation is transformed into the logarithmic magnitude spectrum by the Fourier transform. This transform is generally carried out as a fast Fourier transform (FFT) or through the convolution and deconvolution of basis vectors. The basis vectors are defined by

$$W(f_n,d)=C\cos(2\pi d f_n),$$

where f_n is the center frequency at the bandpass filter and C is a constant. When targets are not moving, the Doppler shift need not be considered, and the frequency ranges of the echoes and the emission coincide. Since the frequency range is 20–120 kHz in our model, the spectrum, which has constant amplitude along the frequency axis, is transformed into an impulse response, with sidelobes except at time zero, through the deconvolution of the basis vectors (Saillant *et al.*, 1993). The all-pass improved basis vectors are normalized as follows:

$$W_{\mathrm{all}}(f_n,d)=C\cos(2\pi df)-\left\{\sum_f \frac{C\cos(2\pi df)}{N}\right\}.$$

Here N is the number of filter channels. Thus, the transform employing the improved set of basis vectors is slightly different from the cepstral analysis.

The logarithmic amplitude spectrum in the non-EEDNF region is transformed into the impulse response through the deconvolution of the selective-pass improved basis vectors, defined by

$$W_{\mathrm{sel}}(f_n,d)=CA(f_n)\cos(2\pi df_n)-\left\{\sum_f \frac{CA(f)\cos(2\pi df)}{N}\right\},$$

where $A(f)=1$ in the non-EEDNF region and $A(f)=0$ in the EEDNF region. In our model, the EEDNF region ranges from 31 to 44 kHz.

Altes, R. A. (**1976**). ''Sonar for generalized target description and its similarity to animal echolocation systems,'' J. Acoust. Soc. Am. **59**, 97–106.

Altes, R. A. (**1978**). ''Angle estimation and binaural processing in animal echolocation,'' J. Acoust. Soc. Am. **63**, 155–173.

Brugge, J. F., Anderson, D. J., Hind, J. E., and Rose, J. E. (**1969**). ''Time structure of discharges in single auditory nerve fibers of the squirrel monkey in response to complex periodic sounds,'' J. Neurophysiol. **32**, 386–401.

Casseday, J. H., Covey, E., and Vater, M. (**1988**). ''Connections of the superior olivary complex in the rufous horseshoe bat, *Rhinolophus rouxi*,'' J. Comp. Neurol. **278**, 313–329.

Covey, E., and Casseday, J. H. (**1991**). ''The monaural nuclei of the lateral lemniscus in an echolocating bat: parallel pathways for analyzing temporal features of sound,'' J. Neurosci. **11**, 3456–3470.

Dear, S. P., Simmons, J. A., and Fritz, J. (**1993a**). ''A possible neuronal basis for representation of acoustic scenes in auditory cortex of the big brown bat,'' Nature (London) **364**, 620–623.

Dear, S. P., Fritz, J., Haresign, T., Ferragamo, M., and Simmons, J. A. (**1993b**). ''Tontopic and Functional organization in the auditory cortex of the big brown bat *Eptesicus fuscus*,'' J. Neurophysiol. **70**, 1988–2009.

Fenton, M. B., and Bell, G. P. (**1981**). ''Recognition of species of insectivorous bats by their echolocation calls,'' J. Mammal. **62**, 233–243.

Fuzessery, Z. M. (**1996**). ''Monaural and binaural spectral cues created by the external ears of the pallid bat,'' Hear. Res. **95**, 1–17.

Griffin, D. R. **(1958)**. *Listening in the Dark* (Yale U. P., New Haven, CT; reprinted by Cornell U. P., Ithaca, New York, 1986).

Grinnell, A. D. **(1995)**. ''Hearing in bats: an overview,'' in *Hearing by Bats*, edited by A. N. Popper and R. R. Fay (Springer, New York), pp. 1–36.

Grothe, B., Vater, M., Casseday, J. H., and Covey, E. (**1992**). ''Monaural interaction of excitation and inhibition in the medial superior olive of the mustached bat: an adaptation for biosonar,'' Proc. Natl. Acad. Sci. U.S.A. **89**, 5108–5112.

Hartley, D. J., and Suthers, R. A. (**1989**). ''The sound emission pattern of the echolocating bat, Eptesicus fuscus, while tracking moving prey,'' J. Acoust. Soc. Am. **91**, 1348–1351.

Jen, P. H., Sun, X., and Lin, P. J. J. (**1989**). ''Frequency and space representation in the primary auditory cortex of the frequency modulating bat *Eptesicus fuscus*,'' J. Comp. Physiol. A **165**, 1–14.

Jen, P. H., Sun, X., Kamada, T., Zhang, S., and Shimozawa, T. (**1984**). ''Auditory response properties and spatial response areas of superior collicular neurons of the FM bat *Eptesicus fuscus*,'' J. Comp. Physiol. A **154**, 407–413.

Kawasaki, M., Margoliash, D., and Suga, N. (**1988**). ''Delay-tuned combination-sensitive neurons in the auditory cortex of the vocalizing mustached bat,'' J. Neurophysiol. **59**, 623–635.

Kuc, R. (**1994**). ''Sensorimotor model of the bat echolocation and prey capture,'' J. Acoust. Soc. Am. **96**, 1965–1978.

Lawrence, B. D., and Simmons, J. A. (**1982**). ''Echolocation in bats: the external ear and perception of the vertical positions of targets,'' Science **218**, 481–483.

Liu, S., and Xiang, J. (**1999**). ''Novel method for super-resolution in radar range domain,'' IEE Proc. Radar, Sonar, Navigation **146**, 40–44.

Masters, W. M., Moffat, A. J. M., and Simmons, J. A. (**1985**). ''Sonar tracking of horizontally moving targets by the big brown bat *Eptesicus fuscus*,'' Science **228**, 1311–1333.

Miller, L. A. (**1991**). ''Arctiid moth clicks can degrade the accuracy of range difference discrimination in echolocating big brown bats, *Eptesicus fuscus*,'' J. Comp. Physiol. A **168**, 571–579.

Müller, R., and Schnitzler, H.-U. (**2000**). ''Acoustic flow perception in cf-bats: Extraction of parameters,'' J. Acoust. Soc. Am. **108**, 1298–1307.

Oppenheim, A. V., and Schafer, R. W. **(1975)**. *Digital Signal Processing* (Prentice-Hall, Englewood Cliffs).

Peremans, H. and Hallam, J. (**1998**). ''The spectrogram correlation and transformation receiver, revisited,'' J. Acoust. Soc. Am. **104**, 1101–1110.

Pickles, J. O. **(1988)**. *An Introduction to the Physiology of Hearing* (Academic, London).

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). ''Phase-locked responses to low-frequency tones in single auditory nerve fibers of the squirrel monkey,'' J. Neurophysiol. **30**, 769–793.

Russell, I., and Palmer, A. **(1986)**. ''Filtering due to the inner hair-cell membrane properties and its relation to phase-locking limit in cochlear nerve fibers,'' in *Auditory Frequency Selectivity*, edited by B. C. J. Moore and R. D. Patterson (Plenum, New York), pp. 199–206.

Saillant, P. A., Simmons, J. A., Dear, S. P., and McMullen, T. A. (**1993**). ''A computational model of echo processing and acoustic imaging in frequency-modulated echolocating bats: The spectrogram correlation and transformation receiver,'' J. Acoust. Soc. Am. **94**, 2691–2712.

Sanderson, M. I., and Simmons, J. A. (**2000**). ''Neural responses to overlapping FM sounds in the inferior colliculus of echolocating bats,'' J. Neurophysiol. **83**, 1840–1855.

Schuller, G., O'Neill, W. E., and Radtke-Schuller, S. (**1991**). ''Facilitation and delay sensitivity of auditory cortex eurons in CF-FM bats, *Rhinolophus rouxi* and *Pteronotus p. parnelli*,'' Eur. J. Neurosci. **3**, 1165–1181.

Shimozawa, T., Suga, N., Hendler, P., and Schuetze, S. (**1974**). ''Directional sensitivity of echolocation system in bats producing frequency-modulated signals,'' J. Exp. Biol. **60**, 53–69.

Simmons, J. A. (**1973**). ''The resolution of target range by echolocating bats,'' J. Acoust. Soc. Am. **54**, 157–173.

Simmons, J. A. (**1989**). ''A view of the world through the bat's ear: the formation of acoustic images in echolocation,'' Cognition **33**, 155–159.

Simmons, J. A. (**1993**). ''Evidence for perception of fine echo delay and phase by the FM bat, *Eptesicus fuscus*,'' J. Comp. Physiol. A **172**, 533–547.

Simmons, J. A., Ferragamo, M., Moss, C. F., Stevenson, S. B., and Altes, R. A. (**1990a**). ''Discrimination of jittered sonar echoes by the echolocating bat *Eptesicus fuscus*: The shape of target images in echolocation,'' J. Comp. Physiol. A **167**, 589–616.

Simmons, J. A., Ferragamo, M. J., Saillant, P. A., Haresign, T., Wotton, J. M., Dear, S. P., and Lee, D. N. (**1995a**). ''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., Kick, S. A., Lawrence, B. D., Hall, C., Bard, C., and Escudie, B. (**1983**). ''Acuity of horizontal angle discrimination by the echolocating bat, *Eptesicus fuscus*,'' J. Comp. Physiol. **153**, 321–330.

Simmons, J. A., Moss, C. F., and Ferragamo, M. (**1990b**). ''Convergence of temporal and spectral information into acoustics images of complex sonar targets perceived by the echolocating bat *Eptesicus fuscus*,'' J. Comp. Physiol. **166**, 449–470.

Simmons, J. A., Saillant, P. A., Wotton, J. M., Haresign, T., Ferragamo, M. J., and Moss, C. F. (**1995b**). ''Composition of biosonar images for target recognition by echolocating bats,'' Neural Networks **8**, 1239–1261.

Suga, N. (**1984**). ''The extent to which biosonar information is represented

in the bat auditory cortex,'' in *Dynamic Aspects of Neocortical Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 315–373.
Suga, N., and O'Neill, W. E. (**1979**). ''Neural axis representing target range in the auditory cortex of the mustached bat,'' Science **206**, 351–353.
Suga, N., O'Neill, W. E., and Manabe, T. (**1978**). ''Cortical neurons sensitive to particular combinations of information-bearing elements of biosonar signals in the mustached bat,'' Science **200**, 778–781.
Sullivan, W. E. (**1982**). ''Neural representation of target distance in auditory cortex of the echolocating bat, *Myotis lucigugus*,'' J. Neurophysiol. **48**, 1011–1032.
Surlykke, A. (**1992**). ''Target ranging and the role of time-frequency structure of synthetic echoes in big brown bats, *Eptesicus fuscus*,'' J. Comp. Physiol. A **170**, 83–92.
Swindlehurst, A. L., and Gunther, J. H. (**1999**). ''Methods for blind equalization and resolution of overlapping echoes of unknown shape,'' IEEE Trans. Signal Process. **47**, 1245–1254.
Valentine, D. E., and Moss, C. F. (**1997**). ''Spatially selective auditory responses in the superior colliculus of the echolocating bat,'' J. Neurosci. **17**, 1720–1733.
Wotton, J. M., and Simmons, J. A. (**2000**). ''Spectral cues and perception of the vertical position of targets by the big brown bat, *Eptesicus fuscus*,'' J. Acoust. Soc. Am. **107**, 1034–1041.
Wotton, J. M., Haresign, T., and Simmons, J. A. (**1995**). ''Spatially dependent acoustic cues generated by the external ear of the big brown bat, *Eptesicus fuscus*,'' J. Acoust. Soc. Am. **98**, 1423–1445.
Wotton, J. M., Haresign, T., Ferragamo, M. J., and Simmons, J. A. (**1996**). ''Sound source elevation and external ear cues influence the discrimination of spectral notches by the big brown bat, Eptesicus fuscus,'' J. Acoust. Soc. Am. **100**, 1764–1776.

On the feasibility of remote palpation using acoustic radiation force

Kathryn R. Nightingale,[a] Mark L. Palmeri, Roger W. Nightingale, and Gregg E. Trahey
Department of Biomedical Engineering, Duke University, Box 90281, Durham, North Carolina 27708

(Received 13 October 2000; revised 13 October 2000; accepted 11 April 2001)

A method of acoustic remote palpation, capable of imaging local variations in the mechanical properties of tissue, is under investigation. In this method, focused ultrasound is used to apply localized (on the order of 2 mm^3) radiation force within tissue, and the resulting tissue displacements are mapped using ultrasonic correlation based methods. The tissue displacements are inversely proportional to the stiffness of the tissue, and thus a stiffer region of tissue exhibits smaller displacements than a more compliant region. In this paper, the feasibility of remote palpation is demonstrated experimentally using breast tissue phantoms with spherical lesion inclusions, and *in vitro* liver samples. A single diagnostic transducer and modified ultrasonic imaging system are used to perform remote palpation. The displacement images are directly correlated to local variations in tissue stiffness with higher contrast than the corresponding B-mode images. Relationships between acoustic beam parameters, lesion characteristics and radiation force induced tissue displacement patterns are investigated and discussed. The results show promise for the clinical implementation of remote palpation. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1378344]

PACS numbers: 43.80.Qf, 43.80.Sh, 43.80.Vj [FD]

I. INTRODUCTION

A. The remote palpation method

It is hypothesized that an acoustic radiation force can be used to generate localized tissue displacements, and that these displacements will be directly correlated with localized variations in tissue stiffness. It is further hypothesized that this can be accomplished using a single transducer on a diagnostic ultrasound scanner to both generate the radiation force and detect the resulting displacements. These hypotheses form the basis for a new imaging method called remote palpation. In this method, acoustic radiation force is used to generate localized displacements in tissue. These displacements are measured using ultrasonic correlation based methods. The magnitude of the tissue displacement is inversely proportional to the local stiffness of the tissue. Radiation force induced displacement maps are generated at multiple locations, and combined to form a single image of variations in tissue stiffness throughout an extended field of view (FOV). A single transducer on a diagnostic scanner is used to both generate the high intensity ''pushing'' beams and track the resulting tissue displacements.

B. Purpose

Two potential clinical applications for remote palpation are lesion detection and characterization, and the identification and characterization of atherosclerosis. In this paper we focus on the former, as it pertains to the early detection of breast cancer, which has been shown to significantly improve patient survival. Existing methods of breast cancer detection include screening mammography and palpation, either by patient self-examination or clinical breast exam. Palpation relies on the manual detection of differences in tissue stiffness between breast lesions and normal breast tissue. The success of palpation is due to the fact that the elastic modulus (or Young's modulus) of breast lesions is often an order of magnitude greater than that of normal breast tissue,[1,2] i.e., breast lesions feel ''hard'' or ''lumpy'' as compared to normal breast tissue.

Differences in Young's moduli are the basis for the investigation of imaging modalities that provide information about the stiffness of tissue. Traditionally, these have fallen into two categories: (1) Sonoelasticity, in which low frequency shear wave propagation is imaged using Doppler or magnetic resonance methods. Estimates of the elastic modulus of the tissue are based upon this information.[3–7] (2) Elastography, in which local variations in tissue strain are determined by measuring local displacements that occur during global tissue compression. Reconstruction methods are used to determine the elastic modulii associated with the measured strain fields.[1,8–11] Remote palpation is similar to elastography, which has demonstrated some success in the detection of malignant breast lesions,[12] however it has several potential advantages. These include the very localized application of radiation force (as opposed to global external compression), the decrease in maximum tissue strain required for lesion visualization, and the potential for real-time implementation without the need for external compression fixtures.

II. BACKGROUND

A. Acoustic radiation force

Acoustic radiation force is a unidirectional force that is applied to absorbing or reflecting targets in the propagation path of an acoustic wave. This phenomenon is caused by a transfer of momentum from the acoustic wave to the propa-

[a]Electronic mail: kathy.nightingale@duke.edu

gation medium. The contribution of absorption is in the direction of wave propagation, whereas the contribution of scattering is dependent upon the angular scattering properties of the target. In the event that the target is both absorbing and has a scatterer that acts as a reflector with an axis of symmetry perpendicular to the direction of wave propagation (e.g., the reflecting target is spherical), the radiation force is entirely in the direction of wave propagation. In this situation, the radiation force is given by[13]

$$F=\left(\Pi_a+\Pi_s-\int \gamma\cos\theta r\,dr\,d\theta\right)\langle E\rangle, \tag{1}$$

where Π_a is the total power absorbed by the target, Π_s is the total power scattered by the target, γ is the magnitude of the scattered intensity, θ is the scattering angle, $r\,dr\,d\theta$ is an area element of the projection of the target onto the axial/lateral plane, and $\langle E\rangle$ is the temporal average energy density of the propagating acoustic wave.

If there is no absorption, and the target is perfectly reflecting (i.e., $\theta=180°$), the integral term becomes $-\Pi_s$ and the total radiation force is $\{2\Pi_s\langle E\rangle\}$. The factor of 2 can be intuitively explained by the fact that the target stops the wave from propagating, and reflects it in the opposite direction.[14] If the target is a Rayleigh scatterer (i.e., its diameter is much smaller than a wavelength), scattering occurs in all directions equally, and the integral term is zero. Thus the total radiation force on the scatterer is $\{(\Pi_a+\Pi_s)\langle E\rangle\}$.

For a tissue model comprised of a collection of Rayleigh scatterers, one can sum the contribution of the radiation force from each scatterer to determine the total force due to scattering. However, in tissue, the majority of the attenuation of an acoustic wave is due to absorption,[15] thus the contribution to the radiation force by scattering in soft tissue can, in general, be neglected. Equation (1) then becomes

$$F=\Pi_a\langle E\rangle. \tag{2}$$

Assuming plane wave propagation, the radiation force applied to tissue is thus[14,16–18]

$$F=\frac{2\alpha I}{c}, \tag{3}$$

where F (which is in the form of a body force, or force per unit volume) is the acoustic radiation force [kg/(s^2 cm^2)], or [dynes/(1000 cm^3)], c (m/s) is the speed of sound in the medium, α (m^{-1}) is the absorption coefficient of the tissue, and I (W/cm^2) is the temporal average intensity of the acoustic beam at a given point in the tissue. For a focused acoustic beam propagating through soft tissue, the ''target'' is the tissue, and the force is applied throughout the focal region of the acoustic beam.

A phenomenon that is not modeled by Eq. (3) is the enhancement of radiation force generated by nonlinear propagation of an acoustic wave.[18–21] For the same temporal average intensity, a wave with higher pressure amplitude and shorter pulse duration generates a larger radiation force than does a lower amplitude, longer duration wave. This is due to the higher order harmonics generated by nonlinear propagation, which result in an increase in absorption.[18–21] Nonlinear propagation also narrows the transmit beam and shifts the intensity field of a focused acoustic beam.[20] Nonlinearly enhanced increases in radiation force by a factor of 2.6 have been observed in breast applications *in vivo*.[22]

The potential applications of acoustic radiation force in diagnostic imaging have heretofore been widely ignored, primarily because the available energy from diagnostic ultrasound systems was low. However, advances in transducer and system design in recent years have resulted in considerable increases in the maximum possible acoustic energy output from these systems. This has sparked a renewed interest by several laboratories in the potential applications of acoustic radiation force. Our laboratory demonstrated the use of acoustic radiation force to induce acoustic streaming in cyst fluid, and thus differentiate fluid-filled from solid breast lesions *in vivo*.[22,23] Several authors have proposed the use of acoustic radiation force to remotely characterize the mechanical properties of tissue.[19,24–27] One application involves the use of a radiation force field oscillating at the beat frequency of two confocal transducers to vibrate an object; the vibrations are detected by a hydrophone, and are used to generate an image.[27] In another application, acoustic radiation force is used to displace tissue, and the speed of the shear waves generated immediately after force removal is monitored to characterize variations in tissue Young's modulus.[19] In yet another application, acoustic radiation force is used to manipulate the vitreous humor of the eye.[24]

B. Tissue response to radiation force

Remote palpation applies localized forces to tissue, which can be described for idealized cases by elasticity theory. For an infinitely small distributed force volume, the strain field can be derived from the analytic solution for a point load in an infinite elastic solid:[28]

$$\epsilon_z=\frac{-(P\,z\,(1+\nu)\,(4\,z^2\,(-1+\nu)+r^2\,(-1+4\,\nu)))}{8\,E\,\pi\,(r^2+z^2)^{5/2}\,(-1+\nu)},$$

$$\epsilon_r=\frac{-(P\,z\,(-2\,r^2+z^2)\,(1+\nu))}{8\,E\,\pi\,(r^2+z^2)^{5/2}\,(-1+\nu)}, \tag{4}$$

$$\epsilon_\theta=\frac{-(P\,z\,(1+\nu))}{8\,E\,\pi\,(r^2+z^2)^{3/2}\,(-1+\nu)},$$

where $\epsilon_{z,r,\theta}$ are the 3 dimensional normal strains. The variable z is the axial distance along the line-of-action of the applied force, P is the magnitude of the applied force, r is the radial distance from the point of applied force, E is the Young's modulus of the material, and ν is Poisson's ratio. The displacement of the material in the direction of the applied force is obtained by integration of ϵ_z:[28]

$$w=\frac{P\,(1+\nu)\,(4\,z^2\,(-1+\nu)+r^2\,(-3+4\,\nu))}{8\,E\,\pi\,(r^2+z^2)^{3/2}\,(-1+\nu)}. \tag{5}$$

Solving these equations for a magnitude of force similar to that used in RP [i.e., $P=1$ dyne, $\nu=0.49$, $E=30{,}000$ dynes/cm^2 (3.0 kPa)] gives an estimate of the size of the associated displacement and strain fields. The solution indicates that the spatial extent of the axial strain field is very

small, with strains decreasing to a negligible amount at a distance of 0.2 mm from the loading point (Fig. 1). Strains in the radial and circumferential directions are even smaller. This implies that lesions that are separated by more than 0.4 mm can be differentiated when excited with point loads at a 0.2 mm spacing.

C. High intensity acoustic beams

Potential risks associated with remote palpation arise from the fact that high intensity acoustic pulses are utilized. The FDA provides two indices to monitor the safety of diagnostic ultrasound. The mechanical index (MI) is indicative of the potential for cavitation, and the thermal index (TI) is indicative of the potential temperature rise. The high intensity pulses used in remote palpation do not exceed the MI limit of 1.9;[29] cavitation is not anticipated due to the use of relatively high frequencies (7 to 14 MHz). It is not anticipated that the TI will exceed 6, which is the limit above which the FDA requires special provisions for approval of use. Finite element models of the heating associated with the high intensity beams used for remote palpation indicate that the temperature rise will be less than 1 °C,[30] which would correspond to a TI of 1.

Throughout this paper, the derated (or *in situ* estimated) spatial peak temporal average intensity for the different acoustic beams is provided in order to calibrate the reader. The FDA currently limits the spatial peak temporal average intensity ($I_{spta.3}$) to 0.72 W/cm^2 *in situ*,[29] which is intended as an indirect indicator of potential tissue heating. However, this limit was determined assuming an indefinite application time. Short duration, high intensity acoustic pulses as are used in remote palpation (i.e., less than 10 milliseconds) were not foreseen in the development of this FDA limit, and thus the TI and/or temperature increase estimates provide a more appropriate indication of the potential for thermal effects.

D. Goals

The goals of the work presented herein are the following: first, to experimentally demonstrate the feasibility of remote palpation imaging; second, to determine whether the displacement images generated during remote palpation imaging are directly correlated with variations in tissue stiffness; and third, to determine whether a single transducer on a diagnostic scanner can be used to perform remote palpation imaging.

III. METHODS

Experiments were performed with a Siemens Elegra scanner (Siemens Medical Systems, Ultrasound Group, Issaquah, WA), that has been modified to allow user control of the acoustic beam sequences and intensities, as well as providing access to the raw radio-frequency (RF) data. A Siemens 75L40 linear array was used for these experiments. Figure 1 provides a schematic of a typical linear array, and defines the different dimensions discussed below. The 75L40 array consists of 194 elements, each of which are 5 mm tall and approximately 0.2 mm wide, with a center frequency of

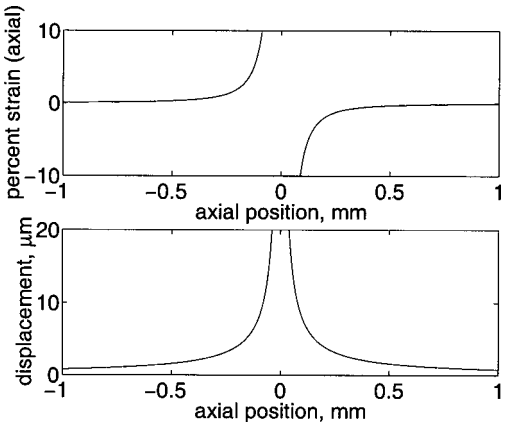


FIG. 1. Percent axial strain (top) and associated axial displacement (bottom) resulting from the application of a point load with a force magnitude similar to that used for remote palpation (i.e., $P=1$ dyne), in an elastic medium with material properties consistent with those of tissue [i.e., $\nu=0.49$, $E=30{,}000$ dynes/cm^2 (3.0 kPa)].[28] Note that the axial extent of both fields has decreased to an insignificant amount within 0.2 mm of the location of force application (0 mm).

7.2 MHz. The number of active elements can be selected electronically to adjust the lateral transmit aperture width. There is a fixed-focus acoustic lens on each element that focuses in the elevation dimension. Lateral focusing is accomplished electronically by applying the appropriate delays to each active element (Fig. 2). The interrogation of different spatial locations is accomplished by using different subapertures (or groups of elements) within the array. In all of these experiments, the lateral aperture was considerably larger than the elevation aperture. This resulted in an asymmetric focal region that was more tightly focused in the lateral dimension than in the elevation dimension (in contrast to a piston, which has an axi-symmetric focal region).

The tracking beams were standard diagnostic B-mode pulses (i.e., single cycle pulses, F/1 focal configuration, apodized, $I_{spta}<0.1$ W/cm^2, MI$\sim$0.4), and they were fired using PRFs ranging from 3 to 5 kHz.

The pulse length, transmit voltage, and transmit F-number (i.e., number of active transmit elements) of the pushing beams were varied in the experiments. The pulse lengths ranged from 0.8 to 5.8 μs, the transmit voltages ranged from 30 to 140 V (in contrast to 20 V for the tracking beams), the F-numbers ranged from F/1 to F/3, the pulse repetition frequencies (PRFs) were varied from 3 to 5 kHz,

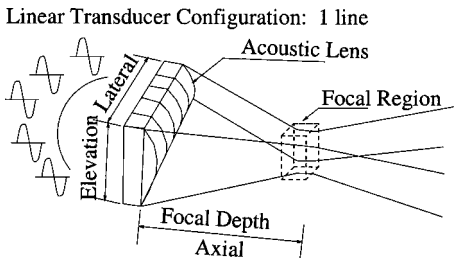


FIG. 2. Linear transducer configuration. Note that there is only 1 row of elements, and each element is much taller (elevation dimension) than it is wide (lateral dimension). There is an acoustic lens that focuses in the elevation dimension, and electronic delays are used to focus in the lateral dimension. Thus, the focal point can be varied in the lateral dimension (as was done in Fig. 5), whereas the elevation focal point is fixed.

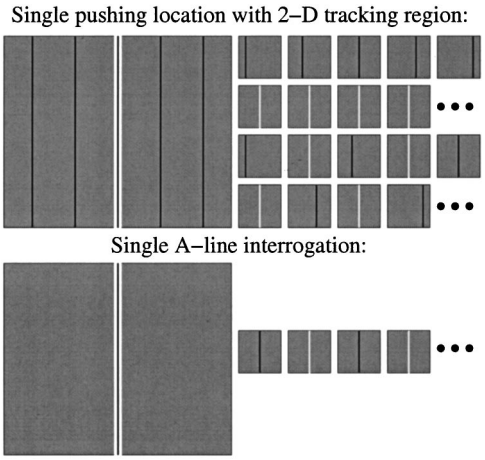


FIG. 3. Schematic of the beam sequences for the different experiments. The black lines indicate tracking lines, and the white lines indicate pushing lines. The larger image on the left presents the spatial relationship of all of the lines. The smaller images on the right indicate the line firing sequences, in order from left to right. Top: single pushing location with 2-D (B-mode) tracking region. For brevity, this schematic shows only 5 spatially separated tracking lines. The different spatial locations are interrogated by using different sub-apertures (or groups of elements) in the array. Bottom: single A-line interrogation with alternating track and push lines.

and they were not apodized. The associated *in situ* temporal average intensities ranged from 2.4 to 140 W/cm^2, which were applied for not more than 50 milliseconds. There was no interference between echoes from the pushing and tracking beams due to the low duty cycle of the pushing beams (1 to 3%).

Several beam sequences were designed to investigate the temporal and spatial response of the phantoms. The experiments are divided into three groups: single pushing location with a two-dimensional (2-D) tracking region, single A-line interrogation, and multiple pushing locations.

A. Single pushing location with 2-D tracking region

For these experiments, the transducer was held in a stationary location. The pulse sequence was designed to transmit a series of 13 spatially distributed tracking beams (i.e., B-mode interrogation using 13 laterally spaced lines, Fig. 3, top) at a PRF of 3584 Hz. This was followed by 10 milliseconds of pushing beams fired at a PRF of 3584 Hz along a single line of flight in the center of the tracking beams, followed by another series of tracking beams interspersed with pushing beams (every other beam; see Fig. 3). The spacing of the tracking beams was 0.7 mm. This sequence was repeated for 50 milliseconds. The raw RF data were stored for off-line processing.

B. Single A-line interrogation

This sequence was designed to fire all of the tracking beams in the same spatial location as the pushing beams (i.e., A-line interrogation, Fig. 3, bottom). The pulse repetition frequency (PRF) for these experiments was 10240 Hz. Every other beam was a tracking beam, thus the PRF for the tracking beams was 5120 Hz. The PRF of the pushing beams was varied by firing either every pushing pulse (PRF 5120 Hz, as shown in Fig. 3, bottom), every fourth pushing pulse (PRF 1280 Hz), every fifth pushing pulse (PRF 1024 Hz), or every sixth pushing pulse (PRF 853 Hz). The first beam fired was always a tracking beam which served as a reference for the initial position. The raw, radio-frequency (RF) data were stored and processed off-line.

C. Multiple pushing locations with 2-D tracking regions

These sequences were the same as those described in Sec. III A, however only 11 tracking lines were utilized for each pushing location (instead of 13), and these lines were spaced at 0.09 mm. This resulted in datasets from each pushing location covering ± 0.5 mm laterally. Multiple pushing locations were interrogated by connecting the transducer to an automated translation stage, and translating the transducer to interrogate different regions of tissue (using the same acoustic beam parameters and sequences at each location). These experiments were performed on two different phantoms, each with an 8 mm diameter lesion located 1.0 cm from the surface of the phantom. The phantoms were submerged in a liquid slurry phantom material, that was created by blending a homogeneous phantom. This was done in order to achieve the same absorption and attenuation of the beam for each axial transducer position, without applying compression to the phantom itself. The pushing point locations spanned 11 mm laterally, and 9 mm axially with 1 mm spacing in both dimensions (Fig. 8 top). The raw RF data from each pushing location were stored for off-line processing.

D. Data processing

The summed RF echo data is 16 bit data, acquired from the Elegra scanner at a sampling rate of 36 MHz. The data is up-sampled to 1.8 GHz, thus given an assumed acoustic velocity of 1540 m/s the minimum measured displacement is 0.4 microns. Off-line data processing was accomplished by performing 1-D cross-correlation in the axial dimension between the up-sampled sequentially acquired tracking lines.[31,9] Each tracking line was divided into a series of search regions (0.7 mm in length), and the location of the peak in the cross-correlation function between a kernel in the first tracking line (0.5 mm in length) and the corresponding search region in the next tracking line was used to estimate the axial tissue displacement in that region. No overlap of the search regions was used. One-wavelength translations were removed and all displacements associated with a correlation coefficient greater than 0.99 were considered valid.

E. Intensity measurements

Pressure and intensity calibration measurements were made in accordance with the guidelines provided by the American Institute of Ultrasound in Medicine.[32] These measurements are made complicated at higher transmit voltages due to acoustic saturation.[33] Saturation occurs at lower pressures in water than it does in phantoms and tissue, thus accurate measurement of the intensities and pressures used for the highest transmit voltages was not possible. However, the relationship between the acoustic radiation force and the re-

TABLE I. Phantom fabrication recipes for 300 milliliters of solution (Ref. 34).

	Gelatin (g)	Water (ml)	N-propanol (ml)	Graphite (g)	Glutaraldehyde (25% soln, ml)	E (kPa)
Tissue (A,B,C)	13.0	230.0	18.0	18.75	52.0	~0.1
Lesion	17.0	222.0	18.0	18.75	60.0	0.5
Tissue (D)	20.0	222.0	18.0	18.75	60.0	1.6

sulting displacement is linear.[26] Therefore, one can obtain an estimate of the magnitude of the radiation force generated by a very high intensity acoustic beam by comparing the corresponding peak displacement obtained in the same phantom using a lower intensity, quantifiable beam. Where Eq. (3) applies, the ratio of the forces is equal to the ratio of the intensities. Thus one can obtain an estimate of the intensity in the higher energy beam under the assumption of linear propagation by evaluating the ratio of the peak displacements.

The above approach does not account for the enhancement in radiation force caused by nonlinear propagation.[18–20,22] This results in an overestimate of the intensity of the beam, due to the assumption of a single frequency and absorption coefficient in Eq. (3). Even so, this method provides a reasonable approximation of the intensities, and thus is utilized to estimate the intensities used in the experiments where actual measurements were not possible (i.e., transmit voltages of 80 V and higher).

F. Phantom construction

Experiments were performed both in elastography phantoms, and in thawed, de-veined calf liver. The elastography phantoms were fabricated from gelatin, graphite, alcohol, water, and glutaraldehyde.[34] The lesions in the phantoms were all generated from a single batch, with a recipe corresponding to a higher Young's modulus than the background material. Table I provides the recipes for the phantoms. The calf liver was purchased frozen from the local grocery store in a vacuum-sealed package. It was thawed under water, and the experiments were performed on the same day.

G. Phantom mechanical property characterization

Characterization of the Young's modulus of tissue and tissue-like media is very challenging. Only a few reports exist in the literature on the subject.[2,35,36] The reported values vary considerably, possibly because Krouskop *et al.* applied pre-compression while making their measurements, whereas the other groups did not;[2,35,36] the stiffness of tissue increases with compression.[2] The numbers provided here were obtained in a method consistent with those reported. This involved modeling of the test apparatus to account for boundary condition deviations from the assumed theoretical solution, and a comparison of the experimental measurements with both the theoretical model (which had inherent assumptions) and the boundary condition correction model.[34,36] A thorough description of these methods is outside of the scope of this paper. It is possible that our measurements suffer from a constant offset, however, the relative stiffnesses of the phantoms are consistent (i.e., in order from the most compliant to the most stiff: Phantom C, Phantom A, Phantom B, Lesion Material, Phantom D).

IV. RESULTS

A. Phantom fabrication

Phantoms A, B, and C were designed using the same recipe shown in Table I. Phantom A had an 11 mm diameter lesion, Phantoms B and C each had an 8 mm diameter lesion. Although the ingredients were the same for each of these phantoms, the background material varied in its Young's modulus slightly between batches. This is likely due to limited control of temperature fluctuations during fabrication.

The measured Young's modulus of the lesion material was 0.5 kPa. The Young's modulus of the background material in Phantoms A, B, and C was 0.07 kPa, 0.13 kPa, and 0.05, respectively. These values are approximately an order of magnitude lower than those reported for soft tissue,[35,36] however the lesion to tissue stiffness ratios (LTSR) are between 3.8 and 10, which are consistent with LTSRs reported for breast tissue.[35]

The Young's modulus of Phantom D was 1.6 kPa. Phantom D was designed without a lesion. The Young's modulus of the liver sample used for the experiments was not measured, however reported values in the literature range from 0.4 to 1.7 kPa.[36] A comparison of the steady state displacements generated with the same remote palpation imaging sequence in both Phantom D and the liver sample indicate that Phantom D was approximately 4 times more stiff than the liver sample.

B. Single pushing location with 2-D tracking region

Figure 4(a) portrays the two-dimensional displacement profile generated in a homogeneous portion of Phantom A for an F/1 focal configuration with a lateral focal point of 20 mm, and a relatively low transmit voltage (29 V). This image was generated after 10 milliseconds of force application, and represents the tissue displacement profile (in two dimensions) resulting from a single force application. If the acoustic beam had generated a point load, one would expect a small circular displacement profile, consistent with Fig. 1. However, the focused acoustic beam used for these experiments is associated with an intensity field that is approximately the size of its focal region. The extent of the inner contour (-2 dB) shown in Fig. 4(a) is approximately the size of the intensity field of the acoustic beam (including intensity values down to -20 dB).

The derated spatial peak temporal average intensity of the pushing beams used to generate Fig. 4(a) was 2.4 W/cm^2,

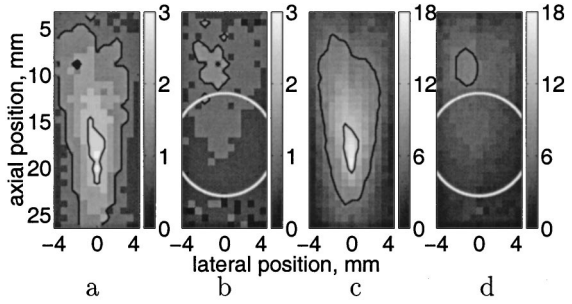


FIG. 4. Displacement maps after 10 milliseconds of force application in phantom A. For each of the images, the black contours represent levels of −2 dB and −10 dB of the peak displacement, the lateral focal point was 20 mm, the transducer was located at the top of the images, and the colorbar scale is in microns. (a) Displacement profile generated in a homogeneous region of the phantom for an F/1 focal configuration using a single pushing location and a relatively low transmit voltage (29 V), and a pulse length of 2.2 microseconds. (b) Displacement profile generated using the same beam sequence as in (a), with the pushing location centered in an 11 mm diameter lesion. The lesion boundaries are highlighted in white. Both (a) and (b) have the same colorbar scales, ranging from 0 to 3 microns. (c) Displacement profile generated in the same location as (a), also using an F/1 focal configuration but with a higher transmit voltage (80 V), and a pulse length of 0.8 microseconds. Note that the displacement profile is similar to (a), however it is shorter axially, with the peak shifted slightly closer to the transducer. This is to be expected in the presence of nonlinear propagation. (d) Displacement profile generated in the same location as (b), using the same beam sequence as in (c). Both (c) and (d) have the same colorbar scales, with displacements ranging from 0 to 18 microns.

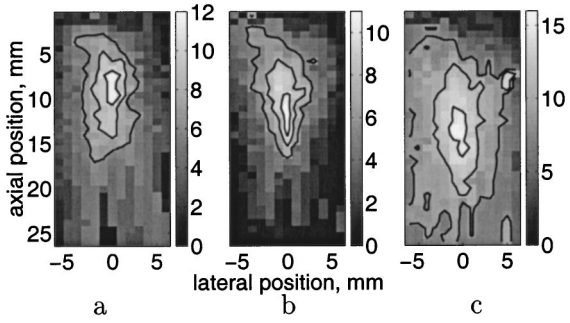


FIG. 5. Displacement maps generated in calf liver after 10 milliseconds of force application. For each of the images, the black contours represent levels of −1 dB, −3 dB, and −6 dB of the peak displacement. The units of displacement on the colorbars are microns. The transducer was located at the top of the images. A constant focal configuration of F/1 was used to generate these images, with a transmit voltage of 140 V, and a pulse length of 3.4 microseconds. (a) lateral focal point at 10 mm. (b) Lateral focal point at 15 mm. (c) Lateral focal point at 20 mm. Note that the peak in the displacement map moves with the lateral focal point, and is slightly closer to the transducer than the focal point.

and the associated radiation force, as computed using Eq. (3), was 100 dynes/cm^3. The maximum displacement is 2.9 microns. Figure 4(b) shows the two-dimensional displacement profile generated when the same beam sequence is focused in the center of an 11 mm diameter lesion in the same phantom. The maximum displacement is smaller than in the homogeneous case (1.4 microns), and the shape of the intensity field is not apparent in the contours of displacement.

Figure 4(c) portrays the displacement map with the pushing beams focused in the same phantom location as in Fig. 4(a), however with a much higher transmit voltage (80 V). The shape of the contours of displacement are similar to Fig. 4(a), however they are shorter axially, and the peak has shifted slightly closer to the transducer. The estimated *in situ* peak intensity ($I_{\text{spta.3}}$) of this beam is 15 W/cm^2 (Sec. III E). The maximum displacement is 18 microns.

Figure 4(d) was obtained using the same beam sequence as in Fig. 4(c), but focused within the lesion [in the same location as in Fig. 4(b)]. The maximum displacement is 7 microns. A comparison of image pairs a–b and c–d indicates that their differences are similar. In both image pairs, the −10 dB contours begin axially in similar locations. As predicted, the presence of the lesion decreases the peak displacement, and alters the displacement profile from the homogeneous case. In both Fig. 4(b) and Fig. 4(d) the maximum displacement is approximately half of that in the corresponding homogeneous cases [Figs. 4(a) and (c)].

Figure 5 portrays the displacement profiles generated in the calf liver sample, using different focal positions (10, 15, and 20 mm). In each case, an F/1 focal configuration was maintained (i.e., the number of active transmit elements was increased as the focal position depth increased). The estimated derated peak intensity of the beam used for the 20 mm focus is 80 W/cm^2 (Sec. III E). In each plot, the peak in the displacement profile occurs slightly in front of the focal position. Although the focal configuration was constant (F/1), the maximum displacement differs in each plot: 12, 11, and 16 microns, for plots a, b, and c respectively.

C. Single A-line interrogation

Figure 6 portrays the maximum displacement through time in Phantom A. In this sequence, all of the tracking lines were fired along the same line (spatially) as the pushing lines, and the tracking and pushing lines were interspersed

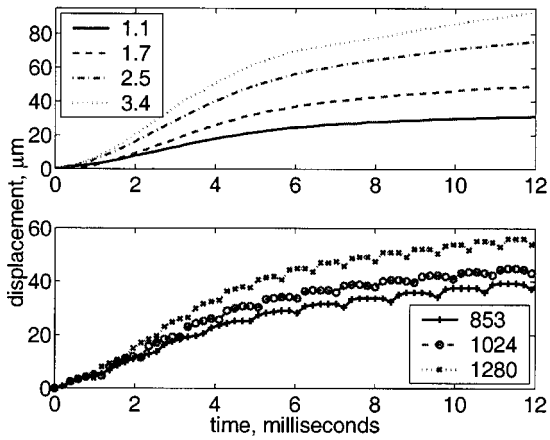


FIG. 6. Displacement through time near the position of peak displacement (17.5 mm from the transducer) in phantom A. Top: The pulse length of the pushing pulses was varied for each curve as indicated in the legend in units of microseconds, but the PRF was held constant (3584 Hz), as was the focal configuration (F/1) and transmit voltage (140 V). Note that, as expected, the steady state displacement for the 3.4 microsecond pulse is double that of the 1.7 microsecond pulse. Bottom: For each of these curves, the same pushing pulses were used (F/1 focal configuration, a pulse length of 5.8 microseconds, and a transmit voltage of 140 V) but with different pushing pulse PRFs. The frequency of the 'rectified sine wave' apparent in each of the three plots corresponds to the PRF of the high intensity beams for each experiment (as indicated in the legend, in Hz). This is due to the recovery of the tissue between pushing beams. Note that the magnitude of the recovery increases over time, with increasing mean tissue displacement.

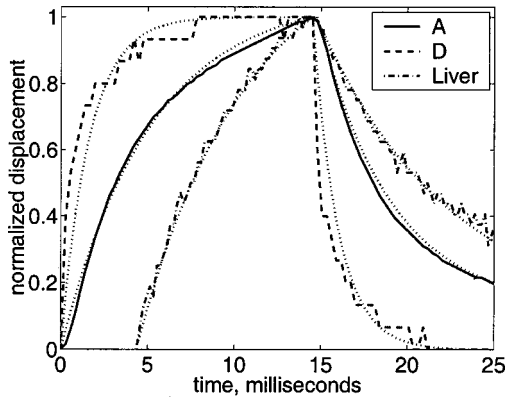


FIG. 7. Displacement over time in the focal region of the transducer for Phantoms A and D in a homogeneous portion of each phantom, as well as in the liver sample. In both Phantom A and the liver sample, the pushing pulses were 3.4 microseconds long, with an F/1 focal configuration and a transmit voltage of 140 V. In Phantom D, the pushing pulses were 5.8 microseconds long, with an F/1 focal configuration and a transmit voltage of 140 V. In all experiments the pushing pulses and tracking pulses were interspersed, and the pushing pulses were fired for a total of 52 lines, and then turned off. The PRF in the liver sample was 5120 Hz, whereas in phantoms A and D it was 3584 Hz. The time at which the pushing pulses were turned off has been aligned in these plots for ease of analysis. The dashed lines represent exponential fits to the data. The time constants for the ramp up time are 5.1, 1.6, and 6.2 milliseconds for phantoms A, D, and the liver sample, respectively. The time constants for the recovery are 4.2, 1.5, and 7.9 milliseconds for phantoms A, D, and the liver sample, respectively.

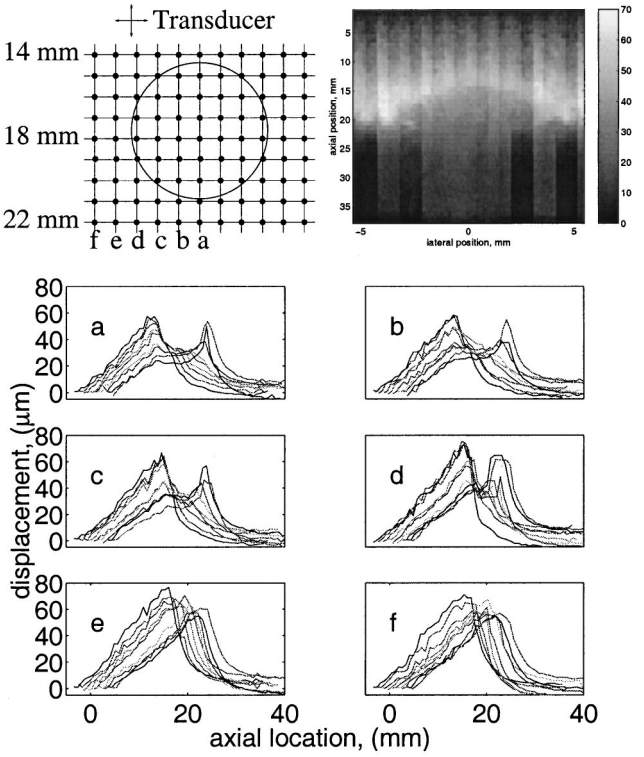


FIG. 8. Top left: Grid of pushing locations (9×11 mm, 1 mm spacing in both the axial and lateral directions) and their relationship to the lesion. Each dot at the grid intersections represents an independent pushing location focal point (and thus a different position of the translation stage and transducer). At each position, the beam sequence described in Sec. III A was fired (with line spacings as described in Sec. III C). The 2-D tracking region surrounding each pushing focal point extends across the entire ROI axially, and laterally it extends ±0.5 mm, halfway between each of the neighboring focal point positions (i.e., dots) to the left and right. Top right: Raw data from the fourth row of pushing locations (including all of the lateral tracking lines) used to generate the image shown in Fig. 9. The colorbar scale is in microns, and the transducer is located at the top of the image. Note that although the focal position of the pushing beams is at a fixed range (19 mm), the lesion is well defined from 12 to 21 mm axially. Bottom: Each subplot represents the displacement along the axis of the high intensity pushing beam for the 9 axial pushing locations at each lateral position shown on the grid above [(a)–(f)]. The transducer is at the left side of these plots. The lesion is clearly apparent as a ''dip'' in the displacement profiles in lines a-d, and the width of the ''dip'' is consistent with the cross-sectional dimension of the lesion at each lateral location.

(i.e., track, push, track, push). The system was in an F/1 focal configuration, and the transmit voltage was 140 Volts.

The curves in the top plot were generated using the same PRF for both the pushing and tracking pulses (3584 Hz), but different pulse lengths. As expected, the increase in steady state displacement is linearly related to the increase in pulse length (i.e., the increase in intensity). For example, the steady state displacement for the 3.4 microsecond pulse is double that of the 1.7 microsecond pulse. Also as expected, the time constants for displacement were similar for the different pulse lengths ($\sim$6 milliseconds, determined by fitting the curves to an exponential).

The bottom plot was generated using pushing beams with the same pulse length and transmit voltage, but with different PRFs (the PRF of the tracking beams was held constant at 3584 Hz). The frequency of the ''rectified sine wave'' apparent in each of the three curves corresponds to the PRF of the high intensity beams for each experiment. This temporal response is due to the recovery of the tissue between pushing beams. Note that the magnitude of the recovery increases with increasing mean tissue displacement. This recovery is not apparent in the top plot because the PRF of the tracking beams was the same as that for the pushing beams, thus the recovery response was undersampled. Again, as expected, the time constant for each of these curves was similar to those in the top plot ($\sim$6 milliseconds).

Figure 7 portrays the temporal response of Phantoms A, D, and the liver sample during force application and after removal of the force. The ramp-down time constants are 4.2, 1.5, and 7.9 milliseconds, respectively. This parameter is indicative of the damping present in the tissue. The ramp-up time for Phantom D is about the same (1.6 milliseconds), however it is slightly slower for Phantom A (5.1 milliseconds), and slightly faster for the liver sample (6.2 milliseconds).

D. Multiple pushing locations with 2-D tracking regions

Figure 8 (top) provides a schematic showing the position of the different pushing locations (i.e. focal points) used in the multiple pushing location experiments, and their relationship to the lesion in Phantom C. The sub-plots show the displacement in the same spatial location as the pushing beam after 10 milliseconds of force application in each grid location. Each subplot represents the displacements obtained at all of the axial positions in one lateral location in the grid (as indicated by the letters a–f). The presence of the lesion is clear in these plots, where a ''dip'' in the axial profile occurs

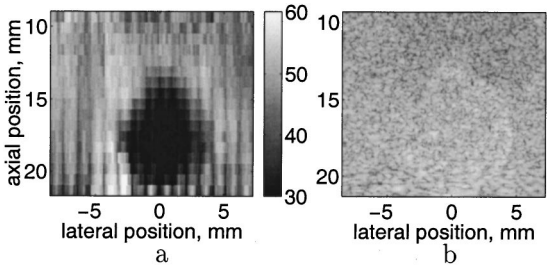


FIG. 9. (a) Displacement image created from multiple pushing locations in phantom C, corresponding to the raw data shown in Fig. 8. The pushing pulses were 1.7 microseconds long, with an F/1 focal configuration and a transmit voltage of 140 V. The PRF was 3584 Hz. The image was created by combining the displacements in each spatial location after 10 milliseconds of force application. The units of displacement on the colorbar are microns, and the transducer is at the top of the images. (b) Corresponding B-mode image.

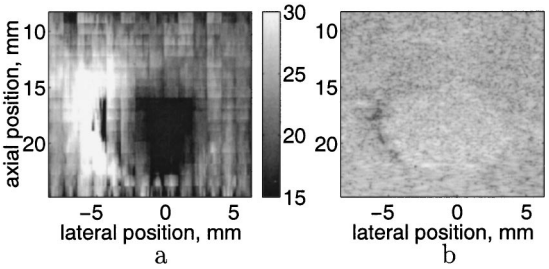


FIG. 10. (a) Displacement image created from multiple pushing locations in phantom B. This phantom had a tear between the lesion and the background material on the left side of the lesion. This resulted in some of the liquid phantom leaking into the tear, and hence a much more compliant region in the vicinity of the tear. The pushing pulses were 0.83 microseconds long, with an F/1 focal configuration and a transmit voltage of 140 V, and a PRF of 3584 Hz. The force applied in this experiment was half of that used to generate the image in Fig. 9. The image was created by combining the displacements in each spatial location after 10 milliseconds of force application. The units of displacement on the colorbar are microns, and the transducer is at the top of the images. (b) Corresponding B-mode image.

in each lateral location that intersected the lesion. The extent of the ''dip'' corresponds to the length of the interrogated cross-section of the lesion.

Figure 9(a) portrays the combination of all of the displacement data from the different pushing locations (from Fig. 8) into a single displacement image. In this image, the displacement at each pixel is determined by computing the average displacement from all of the pushing locations whose focal regions overlap with that pixel. The lesion is clearly apparent in the displacement image, while it is not as evident in the conventional B-mode image [Fig. 9(b)].

Figure 10(a) represents a combined displacement image of Phantom B that was generated in the same way as Fig. 9(a). Phantom B had a slight tear along the interface between the lesion and the background on the left side of the image. This region filled with the liquid slurry phantom material that surrounded the phantom. Therefore, it was much more compliant than the phantom itself. This is apparent in the displacement map shown in Fig. 10(a), as well as in the corresponding B-mode image [Fig. 10(b)].

V. DISCUSSION

Several insights can be gained from the simplified point load analysis of radiation force (Fig. 1). This model is applicable to the lateral and elevation dimensions, where the size of the acoustic focal region is a fraction of a millimeter. It implies that RP can locally sample tissue stiffness with minimal confounding artifacts from neighboring tissue stiffness variations in these dimensions. In addition, because the strain and displacement fields are highly localized, displacements measured at the point of force application are highly correlated with tissue stiffness at that location.

The point load model is not directly applicable in the axial dimension, where the radiation force is distributed along a line (i.e., the focal region is approximately 1 cm axially). This does not, however, result in poor axial resolution. As shown in Fig. 8, the lesion boundary is preserved in all of the data where the axial extent of the focal region (i.e., the line of distributed force) crosses the lesion boundary.

Figures 9 and 10 clearly demonstrate the feasibility of acoustic remote palpation imaging. The homogeneous background medium is a fairly uniform gray color, and the stiffer lesion is clearly apparent as a darker region. The spatial extent of the lesions shown in these images is comparable to that in the corresponding B-mode images. The presence of the lesion is not apparent in the displacement maps beyond the lesion boundaries. This is clearly evident in Fig. 8, where the displacements on either side of the ''dip'' are very similar to those obtained in the same axial location in the absence of a lesion [e.g., subplots 8(d) and (f)]. This supports the hypotheses that remote palpation imaging can be performed in complex media without significant artifacts induced by neighboring structures, and that the displacements achieved in remote palpation imaging are directly correlated to tissue stiffness.

As predicted by finite element simulations,[26] the displacement profiles remain approximately the same shape for increases in intensity; and the displacement magnitude scales with intensity [Figs. 4(a) and (c)]. The slight decrease in the axial extent of the displacement profile, and the slight shift of the displacement peak towards the transducer in Fig. 4(c) is likely a result of nonlinear propagation of the higher pressure amplitude acoustic beam.[20,37] In both cases, the presence of the lesion results in a considerably different displacement profile.

In comparing Figs. 4(a) and (c), and Figs. 4(b) and (d), it is clear that the images generated at lower intensities exhibit more noise. This is due to the lower displacements in these images. The minimum detectable displacement for a correlation based tracking algorithm is dependent upon several factors: system signal-to-noise ratio (SNR), pulse-to-pulse correlation, signal bandwidth, as well as the center frequency and pulse length of the tracking beams. Utilizing values for the above parameters that are applicable to the experiments performed herein (i.e., 40 dB, 0.99, 70%, 7.2 MHz, 0.3 μs, respectively), the lower limit of displacement that can be tracked is computed to be 0.5μm.[38] The majority of the displacements in the low intensity cases [Figs. 4(a) and (b)] are close to this minimum detectable value, thus these displacement images have a lower signal-to-noise ratio (SNR) than those corresponding to the higher intensity beams [Figs. 4(c)

and (d)]. Clearly, larger displacements are preferable to achieve better SNRs.

Figure 5 represents the first *in vitro* remote palpation displacement maps generated in a tissue sample. Displacements of up to 15 microns were achieved in this tissue sample. The varying lateral focal point experiments were performed in order to investigate the limitations presented by a linear array with a fixed elevation focus. Each of the three maps was generated using an F/1 focal configuration and the same transmit voltage and pulse length. If the arrays were axi-symmetric (i.e., pistons), one would expect identically shaped displacement profiles, with decreases in peak displacement at deeper focal positions due to the attenuation of the overlaying tissue. However, because a linear array (asymmetric, $5\,\text{mm}\times 20\,\text{mm}$) was used, the peak displacement actually occurs for the deepest focal position [Fig. 5(c)]. This focal position is closer to the elevational focal point than the other two focal positions, which results in increased intensities in this location. Remote palpation will ideally be implemented by holding the transducer in a stationary position, and using electronic focusing to apply radiation force to multiple pushing locations within an extended field of view (FOV). The implications of Fig. 5 are that some type of depth/focal position dependent scaling may be necessary to achieve a uniform displacement map in a homogeneous medium. This would be similar to time gain control (TGC), which is available on diagnostic scanners to account for tissue attenuation and focal gain. These results also suggest that a two-dimensional array capable of electronically focusing in both the lateral and elevation dimensions is preferable to a linear array.

Evaluation of the steady state displacements shown in the top of Fig. 6 indicates a linear relationship between steady state displacement and acoustic pulse length. Because increases in radiation force are proportional to increases in temporal average intensity (which are proportional to increases in pulse length), the steady state displacements shown in this plot portray a linear relationship between radiation force and temporal average intensity, which was predicted by finite element simulations.[26]

The bottom plot in Fig. 6 was generated using the same pushing pulses, with different PRFs. Therefore, the intensity was lower for the lower PRFs, which resulted in smaller steady state displacements. Because the PRF of the tracking pulses was fixed at 3584 Hz, which is above Nyquist for each of the pushing pulse PRFs, the bottom plot in Fig. 6 portrays the recovery of the phantom after each individual pushing pulse. This is apparent in the ''rectified sine wave'' that appears on each curve. The PRF of the pushing pulses is identical to the frequency of the ''rectified sine wave'' for each curve. It is interesting to note that for each curve, the magnitude of the recovery increases with increasing mean displacement (i.e., the amplitude of the rectified sine wave increases with time). In addition, the magnitude of the ''rectified sine wave'' is slightly larger for the lower PRFs. This is due to the recovery of the tissue between pushing pulses. The longer the time between pushing pulses, the more the tissue recovers resulting in a net ''loss'' in maximum displacement. This suggests that a more efficient method to implement remote palpation will be to use a single, very long pushing pulse, instead of the series of shorter duty cycle pushing pulses (1% to 3% duty cycles) used in these experiments.

This point is further demonstrated by evaluating the displacement after the same number of pulses have been fired at two different PRFs (Fig. 6, bottom). For example, at a PRF of 1280 Hz, after 6 milliseconds, 8 pushing pulses have been fired, and the displacement is 45 microns. Correspondingly, 8 pushing pulses have been fired at a PRF of 853 Hz after 8.5 milliseconds, and the displacement is 36 microns. If one computes the temporal average intensity used in both scenarios over 8.5 milliseconds, it is the same. However, by applying the energy initially, instead of evenly spacing it over the entire time, an increase of 20% is achieved in the maximum displacement.

Figure 7 allows the evaluation of the recovery time of two different phantoms and the liver sample. This quantity is indicative of the damping present in the tissue. The liver sample is clearly more damped than either of the phantoms, as it has a longer recovery time. Phantom D exhibits considerably less damping than the other two, in that its recovery curve appears almost linear.

As stated in the background section, remote palpation is similar to elastography, however it has many potential advantages. The fact that the radiation force field is localized and can be applied in selected remote locations is significant. Because the force is applied directly to the position of interest, considerably smaller forces are required than in the case of global compression. Thus one can achieve lesion detectability without exposing the tissue to potentially damaging larger strain fields. In addition, because displacement is directly correlated with local variations in stiffness, the method does not require complex reconstruction algorithms. Also, in contrast to elastography, the displacements generated during remote palpation are relatively small (microns). Thus the correlation based tracking routines are not subject to problems associated with relative scatterer motion in the presence of large strain.[39] Finally, remote palpation provides the ability to evaluate and image localized variations in the temporal response, or damping, of tissue.

Clinical implementation of remote palpation will be similar to conventional mixed modes, such as B-mode/color-mode. The physician will identify a smaller FOV within a B-mode image where remote palpation will be implemented. Then, upon entering the remote palpation mode, the physician will hold the transducer in a stationary position while the FOV is interrogated via electronic focusing and steering. Depending upon the size of the FOV, the spacing of the pushing locations, and the time spent at each location, frame rates as high as 2 to 5 frames per second may be achieved while interrogating a 1.5 cm square FOV. This frame rate will allow the superposition of remote palpation displacement images with the conventional B-mode image in semi-real-time, which will aid clinical evaluation.

The resolution achievable by remote palpation imaging systems will depend upon a variety of factors, including the size and spacing of the pushing beams, the resolution and spacing of the tracking beams, the target tissue mechanical

properties [Eqs. (4) and (5)], and the methods of image formation utilized. The impact of these factors on resolution is under investigation, however it is outside the scope of this paper.

Given the comparably high intensities and unique pulsing sequences of remote palpation imaging, the safety of the method is of concern. The phantoms used to create Figs. 9 and 10 had Young's moduli that were an order of magnitude lower than those of tissue, whereas phantom D was approximately four times more stiff than the liver sample. Detectable displacements were achieved in all of these media. The derated intensity used in the liver sample was approximately $80\,W/cm^2$. Accounting for the variability of tissue stiffnesses, this suggests that a reasonable upper limit on the *in situ* intensity required for *in vivo* remote palpation will be around $300\,W/cm^2$ at each pushing location (assuming an application time of several milliseconds). This is well in excess of the FDA limit ($0.72\,W/cm^2$).[29] However, these high intensity beams need only be generated for time scales on the order of milliseconds. In this short time period, the heating that would be generated is less than 1 °C,[30] and is limited spatially to a region that is slightly smaller than the focal region of the acoustic beam. Therefore, assuming that adjacent pushing beams do not significantly overlap, the maximum expected temperature rise will be less than 1 °C. It is widely accepted that temperature increases of less than 1 °C do not pose a danger to the patient.[40] Thus, remote palpation imaging as discussed herein should not pose a danger to the patient.

VI. CONCLUSION

The work presented herein demonstrates the feasibility of acoustic remote palpation. Displacement maps that are directly correlated with variations in tissue stiffness were generated and presented. It is possible to perform remote palpation using a single transducer on a diagnostic scanner to both generate the required high intensity beams, and track the resulting displacements. This will facilitate ease of implementation on transition to the clinical setting. The intensities that are required to induce detectable displacements *in vivo* (i.e., tens of microns) will be on the order of $200-300\,W/cm^2$, and the required dwell time less than 10 milliseconds. Due to the short duration of exposure, these beams should not cause appreciable tissue heating. Given these findings, acoustic Remote Palpation clearly exhibits clinical potential.

ACKNOWLEDGMENTS

This work was supported by DOD BCRP Grant No. BC972755. We thank Siemens Medical Systems, Ultrasound Group for their system support. We thank Intel Corporation for their technical and in-kind support.

[1] A. Sarvazyan *et al.*, Acoust. Imaging **21**, 223 (1995).
[2] T. Krouskop *et al.*, Ultrason. Imaging **20**, 260 (1998).
[3] Y. Yamakoshi, J. Sato, and T. Sato, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **17**, 45 (1990).
[4] R. Lerner, S. Huang, and K. Parker, Ultrasound Med. Biol. **16**, 231 (1990).
[5] K. Parker, S. Huang, R. Musulin, and R. Lerner, Ultrasound Med. Biol. **16**, 241 (1990).
[6] T. Krouskop, D. Dougherty, and S. Levinson, J. Rehabil. Res. Dev. **24**, 1 (1987).
[7] S. Levinson, J. Biomech. **20**, 251 (1987).
[8] J. Ophir *et al.*, Ultrason. Imaging **13**, 111 (1991).
[9] M. O'Donnell, A. Skovoroda, B. Shapo, and S. Emelianov, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **41**, 314 (1994).
[10] J. Ophir *et al.*, Proc. Inst. Mech. Eng. **213**, 203 (1999).
[11] L. Gao, K. Parker, R. Lerner, and S. Levinson, Ultrasound Med. Biol. **22**, 959 (1996).
[12] B. Garra *et al.*, Radiology **202**, 79 (1997).
[13] P. J. Westervelt, J. Acoust. Soc. Am. **23**, 312 (1951).
[14] G. Torr, Am. J. Phys. **52**, 402 (1984).
[15] D. Christensen, *Ultrasonic Bioinstrumentation* (Wiley, New York, 1988).
[16] W. Nyborg, in *Physical Acoustics*, edited by W. Mason (Academic, New York, 1965), Vol. IIB, Chap. 11, pp. 265–331.
[17] D. Dalecki, Ph.D. thesis, University of Rochester, 1993.
[18] H. Starritt, F. Duck, and V. Humphrey, Phys. Med. Biol. **36**, 1465 (1991).
[19] A. Sarvazyan *et al.*, Ultrasound Med. Biol. **24**, 1419 (1998).
[20] O. Rudenko, A. Sarvazyan, and S. Emelianov, J. Acoust. Soc. Am. **99**, 2791 (1996).
[21] D. Dalecki, E. Carstensen, K. Parker, and D. Bacon, J. Acoust. Soc. Am. **89**, 2435 (1991).
[22] K. Nightingale, P. Kornguth, and G. Trahey, Ultrasound Med. Biol. **25**, 75 (1999).
[23] K. Nightingale and G. Trahey, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **47**, 201 (2000).
[24] W. Walker, J. Acoust. Soc. Am. **105**, 2508 (1999).
[25] T. Sugimoto, S. Ueha, and K. Itoh, in *Proceedings of the 1990 Ultrasonics Symposium* (Publisher, City, 1990), pp. 1377–1380.
[26] K. Nightingale, R. Nightingale, M. Palmeri, and G. Trahey, Ultrason. Imaging **22**, 35 (2000).
[27] M. Fatemi and J. Greenleaf, Proc. Natl. Acad. Sci. U.S.A. **96**, 6603 (1999).
[28] S. Timoshenko and J. Goodier, *Theory of Elasticity* (McGraw-Hill, New York, 1987).
[29] Center for Devices and Radiological Health (CDRH), U.S. Dept of Health and Human Services, 1985, Rev. 1993 (1994).
[30] M. Palmeri, G. Trahey, R. Nightingale, and K. Nightingale, IEEE UFFC (in preparation).
[31] G. Trahey, J. Allison, and O. VonRamm, IEEE Trans. Biomed. Eng. **BME-34**, 965 (1987).
[32] NEMA Standards Publication/No. UD 2, *Acoustic Output Measurement Standard for Diagnostic Ultrasound Equipment* (National Electrical Manufacturers Association, Washington, DC, 1992).
[33] F. Duck, Ultrasound Med. Biol. **25**, 1009 (1999).
[34] T. J. Hall, M. Bilgen, M. F. Insana, and T. A. Krouskop, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **44**, 1355 (1997).
[35] A. Skovoroda *et al.*, Biophysics (Engl. Transl.) **40**, 1359 (1995).
[36] E. J. Chen, J. Novakofski, W. K. Jenkins, and W. D. O'Brien, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 191 (1996).
[37] A. Baker, Ultrasound Med. Biol. **23**, 1083 (1997).
[38] W. Walker and G. Trahey, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42**, 301 (1995).
[39] M. Bilgen and M. F. Insana, Phys. Med. Biol. **43**, 1 (1998).
[40] NCRP, *Report No. 113: Exposure Criteria for Medical Diagnostic Ultrasound: I. Criteria Based on Thermal Mechanisms* (National Council on Radiation Protection and Measurements, NCRP Publications, Bethesda, MD 20814, 1992).

LETTERS TO THE EDITOR

This Letters section is for publishing (a) brief acoustical research or applied acoustical reports, (b) comments on articles or letters previously published in this Journal, and (c) a reply by the article author to criticism by the Letter author in (b). Extensive reports should be submitted as articles, not in a letter series. Letters are peer-reviewed on the same basis as articles, but usually require less review time before acceptance. Letters cannot exceed four printed pages (approximately 3000–4000 words) including figures, tables, references, and a required abstract of about 100 words.

Linear elastic properties of anisotropic open-cell foams

S. Sahraoui,[a] E. Mariez, and M. Etchessahar
Laboratoire d'Acoustique de l'Université du Maine, UMR–CNRS 6613, 72085 Le Mans Cedex 9, France

(Received 4 April 2000; accepted for publication 30 March 2001)

Plastic foams are commonly used as sound-absorbing materials. These foams can exhibit a low-frequency resonant vibration of a skeleton where fluid damping is important. The two-phase material dynamic behavior is then described by the Biot equations. This paper examines the elasticity of these kinds of porous materials through their microstructure. The foam is modeled by a periodic network of elastic bars. It is shown that the orientation and distribution in space of the bars and their geometry control the macroscopic properties and the anisotropy. This study is an extension of a model developed for isotropic foams by Warren and Kraynick [J. Appl. Mech. **55**, 341–346 (1988)]. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1378351]

PACS numbers: 43.40.Qi, 43.35.Mr, 43.20.Jr, 43.40.Tm [PJR]

I. INTRODUCTION

The Biot theory of fluid-saturated porous media provides a description of the waves propagating in soils (water-saturated rocks).[1] Several authors have extended this theory to sound-absorbing materials, such as glass wool and plastic foams, for noise control applications in engineering activities such as aeronautics and the automotive industries.

In the Biot–Johnson–Allard model, acoustic materials are characterized by many physical parameters which depend on the geometry of the porous material, the properties of the air, and the coupling effects.[2] When a skeleton is motionless, the pressure field in the saturated media is described by the so-called equivalent fluid theory. At low frequencies or when the skeleton is fixed to a vibrating structure, the air pressure field and the solid phase displacement are coupled and, hence, the elastic properties of the porous material become significant. In order to interpret the modal behavior of finite-size structures, three-dimensional finite-element formulations based on extended Biot theory have been developed.[3]

The theories of foam mechanics have variously involved mechanically isolated struts,[4] isolated joints,[5] or isolated cells.[6] Less attention has been given to understanding anisotropic cellular materials.[7] Low-density foams have irregular cell structures with cells that vary in size. The percentage of void content is so high that, effectively, there are only polyhedral cells with well-defined edges, and small islands of material at the vertices. Polymer foams made by pouring the polymer mixed with hardener and a foaming agent into a mould (so that it rises rather like a loaf of bread) usually have cells which are elongated in the rise direction and equiaxed in the plane normal to it, giving them an axisymmetric structure. The purpose of this paper is to propose an extension of the Warren and Kraynick model which will be applicable to isotropic transverse foams.

II. WARREN AND KRAYNIK MODEL[5]

Four half-strut GM_i $(i=1,4)$ of length L which meet at the tetrahedral angle constitute the so-called unit joint illustrated by Fig. 1. A regular tetrahedron containing the unit cell (Fig. 2) represents the investigated equivalent continuous medium. Cutting this volume element (Fig. 2) on the plane $x=0$, Warren and Kraynick introduce the effective stress that must act upon the exposed face to maintain equi-

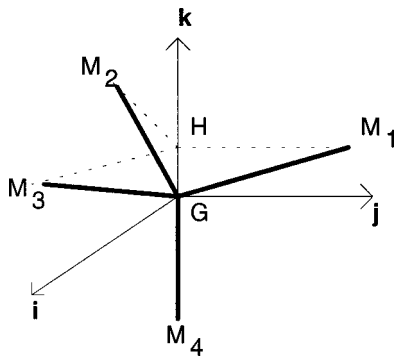


FIG. 1. Unit joint.

[a]Electronic mail: sohbi.sahraoui@univ-lemans.fr

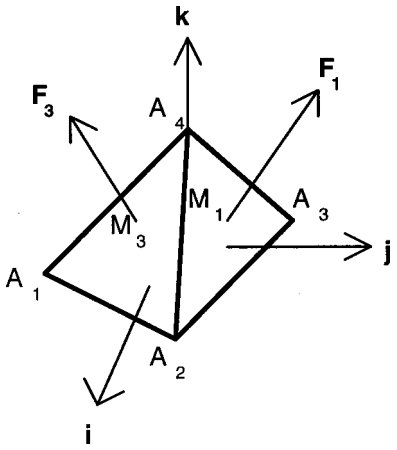


FIG. 2. Tetrahedral element.

librium with the forces on the remaining volume. Similarly, cutting the volume element on the plane $y=0$ then on the bottom triangular face in the plane $z=L$, we can obtain six independent components of the stress tensor. This tensor is averaged for all possible orientations defined by the three Euler angles $\psi\in[0,2\pi]$, $\theta\in[0,\pi]$, $\varphi\in[0,2\pi]$ of the tetrahedral element.

The elastic behavior of the unit cell depends on the axial elastic compliance M and the bending compliance N of the half-struts which are assumed to be the same, respectively, in all directions.

After some calculations, the elastic constants (the Young modulus and Poisson ratio) of the equivalent medium are shown to be related to M and N by

$$E=\frac{11N+4M}{2\sqrt{3}L(10N^2+31MN+4M^2)},$$
$$\nu=\frac{(N-M)(5N+4M)}{(10N^2+31MN+4M^2)}, \tag{1}$$

with

$$M=L/E_0A, \quad N=L^3/3E_0I, \tag{2}$$

where E_0, A, and I represent, respectively, the Young modulus matrix, the area of the strut cross section, and its quadratic moment.

III. MODEL EXTENSION FOR ANISOTROPIC FOAMS

We now consider the elastic deformation of foams which exhibit a transverse isotropy in a plane normal to the ''rise'' direction (corresponding to the **k** direction in Fig. 1).

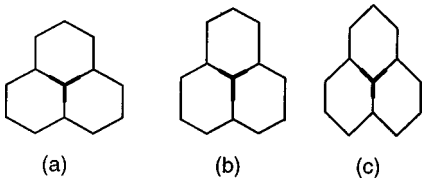


FIG. 3. Geometric parameters of the joint (in bold). (a) $\alpha=\beta=1$, same length, same angle between the half-struts. (b) $\alpha>1$; $\beta=1$. (c) $\alpha=1$; $\beta<1$.

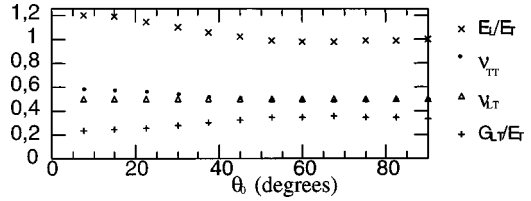


FIG. 4. Elastic constants vs θ_0 (in degrees) for $\alpha=\beta=1$.

Following Warren and Kraynik's analysis summarized above, we consider a new unit joint (four half-strut GM_i', $i=1-4$) with three additional assumptions:

(a) Their exists a preferential orientation for the unit joints. The possible orientations are defined by the three Euler angles:
$\psi\in[0,2\pi]$, $\theta\in[0,\theta_0]$, $\phi\in[0,2\pi]$,
where the angle θ_0 can be considered as an orientation parameter.

(b) The half-strut GM_4' is not equal to the others:
$GM_4'=\alpha GM_4$, $\alpha>1$.

(c) The angles between two half-struts (except the longitudinal one) are equal and smaller than the tetrahedral angle (where β is the reduction factor).

In order to illustrate the α and β parameters in a two-dimensional configuration, we consider a joint (with three half-struts), which is the repeating element of a hexagonal honeycomb (Fig. 3).
The geometric factors α and β, as illustrated, control the shape of the cells constituting the elongated polyhedron and which is oriented within a specific solid angle [the parameter θ_0 in assumption (a)].

Obviously, the original Warren and Kraynik model is obtained when $\alpha=\beta=1$ and $\theta_0=\pi$. In this paper we focus our attention on foams which can be described by simple combinations of these parameters.

For $\alpha=\beta=1$ and $\theta_0=0$, we obtain the five elastic constants of which their definition is given in Ref. 8:

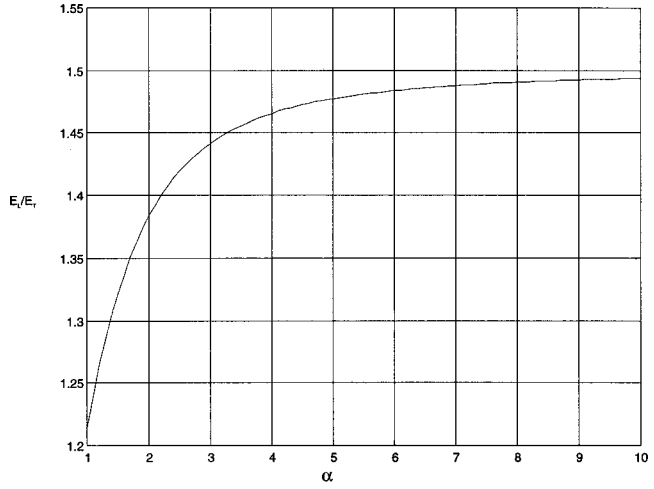


FIG. 5. E_L/E_T vs α for $\beta=1$ and $\theta_o=0$.

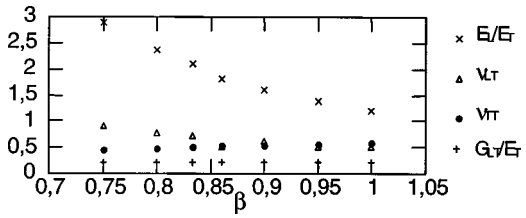


FIG. 6. Elastic constants vs β for $\alpha=1$ and $\theta_o=0$.

$$
\begin{aligned}
E_T &= \frac{1}{2}\,\frac{\sqrt{3}(7N+2M)}{(17N^2+56MN+8M^2)L},\\
E_L &= \frac{1}{2}\,\frac{\sqrt{3}}{M(2N+7M)L},\\
\nu_{TT} &= 2\,\frac{(N-M)(5N+4M)}{17N^2+56MN+8M^2},\\
\nu_{LT} &= \frac{N-M}{2N+7M},\\
G_{LT} &= \frac{1}{108}\,\frac{\sqrt{3}(5N+4M)}{(2M+N)NL},
\end{aligned}
\tag{3}
$$

where the subscripts L and T denote the longitudinal and the transverse directions. For high-porosity foams (more than 0.96), these expressions may be simplified to

$$
\begin{aligned}
&E_T = \frac{7\sqrt{3}}{34}\,\frac{1}{NL}, \quad E_L = \frac{\sqrt{3}}{4}\,\frac{1}{NL},\\
&G_{LT} = \frac{5\sqrt{3}}{108}\,\frac{1}{NL}, \quad \nu_{TT} = \frac{10}{17}, \quad \nu_{LT} = 0{,}5,
\end{aligned}
\tag{4}
$$

where $N(\gg M)$ denotes the bending compliance of the half-strut [Eq. (2)]. The influence of the parameter θ_0 on the elastic constants is shown in Fig. 4.

If we consider $\beta=1$ and $\theta_0=0$, we find

$$
\frac{E_L}{E_T} = \frac{55+81\alpha^2}{2(29+27\alpha^2)}. \tag{5}
$$

The ratio E_L/E_T increases with α (Fig. 5) and can reach a maximum value of 1.5 [Eq. (5)]. The parameters α and θ_0 (see Fig. 3) cannot generate higher E_L/E_T ratios.

Finally, for $\alpha=1$ and $\theta_0=0$, we observe a wide range variation of elastic constants versus β (Fig. 6).

IV. CONCLUSION

The Warren and Kraynick model extension can simulate the anisotropic elastic behavior of high-porosity open-cell foams. The elastic constants depend upon the Young's modulus matrix and the unit-cell shape and orientation. The most marked implications of this model are summarized as follows:

(i) E_L, E_T and G_{LT} are proportional to the Young's modulus matrix E_o.

(ii) ν_{TT} and ν_{LT} are not dependent on the matrix mechanical properties.

For polymeric foams, the viscoelastic behavior can be predicted through the above results: the moduli E_L, E_T, and G_{LT} will have the same frequency dependence as the matrix Young's modulus; the Poisson ratios ν_{TT} and ν_{LT} are real and independent of the frequency. Published experimental results confirm these predictions at low frequencies (1–100 Hz).[9]

[1] M. A. Biot, ''The theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low frequency range. II. Higher frequency range,'' J. Acoust. Soc. Am. **28**, 168–191 (1956).

[2] J. F. Allard, *Propagation of Sound in Porous Media* (Elsevier, New York, 1993).

[3] N. Dauchez, S. Sahraoui, and N. Atalla, ''Convergence of poroelastic finite elements based on Biot displacement formulation,'' J. Acoust. Soc. Am. **109**, 33–40 (2001).

[4] S. V. Kanakkanatt, ''Mechanical anisotropy of open-cell foams,'' J. Cell. Plast. **9**, 50–53 (1973).

[5] W. E. Warren and A. M. Kraynick, ''The linear elastic properties of open cells foams,'' J. Appl. Mech. **55**, 341–346 (1988).

[6] W. E. Warren and A. M. Kraynick, ''Linear elastic behavior of low-density Kelvin foam with open-cells,'' J. Appl. Mech. **64**, 787–794 (1997).

[7] A. T. Huber and L. J. Gibson, ''Anisotropy of foams,'' J. Mater. Sci. **23**, 3031–3040 (1988).

[8] R. M. Christensen, *Mechanics of Composite Materials* (Wiley, New York, 1979), p. 74.

[9] M. Melon, E. Mariez, C. Ayrault, and S. Sahraoui, ''Acoustical and mechanical characterization of anisotropic open-cell foams,'' J. Acoust. Soc. Am. **104**, 2622–2627 (1998).

Estimation of an open-loop compact adaptive passive noise control system with microstructures

Mitsuhiro Yoda[a] and Satoshi Konishi
Department of Mechanical Engineering, Ritsumeikan University, 1-1-1 Nojihigashi, Kusatsu-City, Shiga 525-8577, Japan

(Received 5 May 2000; accepted for publication 18 April 2001)

An adaptive passive noise control system with high-aspect-ratio apertures in a thin perforated plate is presented. An implemented absorber forms an array of Helmholtz resonators which consists of micro-order diameter apertures in a perforated plate and a tunable cavity gap. The simple open-loop control algorithm was employed as a control strategy in order to reduce extra hardware and realize the compact system. The developed adaptive passive noise control system employed high-aspect-ratio apertures in an absorber and adapted for a lower-frequency range without increasing the cavity gap. In the experiment, the compact system could absorb less than 1000 Hz frequency noise with a thin cavity gap, 23 mm. This paper reports the possibilities of a compact adaptive passive noise control system with taking into account of advantages of microstructures based on microelectromechanical systems technology. © *2001 Acoustical Society of America.*
[DOI: 10.1121/1.1379077]

PACS numbers: 43.50.Gf, 43.50.Ki [MRS]

I. INTRODUCTION

Active noise control methods with loud speakers as actuators for wave cancellation have been investigated in the field of acoustic noise control. The advantage of active noise control is attenuate complex signal noise in a broadband frequency. However, those methods have drawbacks including instability, a need of additional energy, and a complicated control algorithm. On the other hand, Helmholtz resonators have been utilized to attenuate an acoustic noise since ancient times. The advantages of passive devices such as Helmholtz resonators are simplicity. Passive devices are effective in the narrow frequency range. Passive devices (Helmholtz resonators) can be improved to adaptive passive devices by tuning the structural parameters in order to achieve effective noise reduction under changing environment conditions. The structural parameters of passive devices should be tuned precisely to achieve effective acoustic noise attenuation. The advantages of adaptive passive noise control are that additional energy is not required and that the control strategy is simple. We are interested in a compact adaptive passive noise control system with a simple control algorithm. Over the past 15 years, many researchers have exploited adaptive passive devices such as mufflers in automobiles, the side branch of ducts in many applications. The following shows representative research on adaptive passive noise control. Koopman and Neise investigated the method by which an adjustable acoustic resonator can be used to attenuate the aerodynamic noise generated by centrifugal fans.[1,2] Tuning of the resonators was achieved by changing the cavity length via a movable Teflon piston. Walker *et al.* also developed active control Helmholtz resonators to attenuate excessive turbofan engine noise.[3] The active Helmholtz resonator provided increased acoustic bandwidth and efficient tonal noise suppression in a multimode, multifrequency environment in an axial fan inlet duct.

Matsuhisa and co-workers developed the variable volume resonator which was used as a side branch in a duct.[4,5] The volume was tuned by displacing a piston in a cavity. In the system control, the resonator cavity was adjusted such that the phase lag between the pressures in the duct and resonator is 90°.

Little *et al.* developed the fluidic intelligent Helmholtz resonator which was applied to the hydraulic engine mount.[6] Tuning of the resonator was achieved by controlling the neck cross-sectional area.

Lamancusa proposed the use of tunable Helmholtz resonators as an alternative to expansion chamber mufflers in automobiles.[7] Two types of tuning methods were proposed in order to vary the resonator volume. The volume variable resonator was controlled according to the engine rpm signal. The developed muffler could provide more than 30 dB of transmission loss of a single pure tone at the resonant frequency.

De Bedout *et al.* proposed the adaptive Helmholtz resonator under changing environmental conditions.[8] The proposed control strategy is comprised of the combination of an open-loop coarse tuning scheme and a closed-loop precise tuning scheme.

Most of the previous studies on adaptive passive acoustic systems employ thick structures with such a macroscale adaptive mechanism as a piston. It becomes important to achieve compact system structures when several attractive results on adaptive passive acoustic systems are applied for practical uses. Especially, the problem of size and weight reduction cannot be bypassed in aeroacoustic applications. In order to solve this problem, we have studied the application of microelectromechanical system (MEMS) technology for adaptive passive acoustic systems.[9] Furthermore, the recent

[a]Electronic mail: sme30297@se.ritsumei.ac.jp

 0001-4966/2001/110(1)/638/4/$18.00

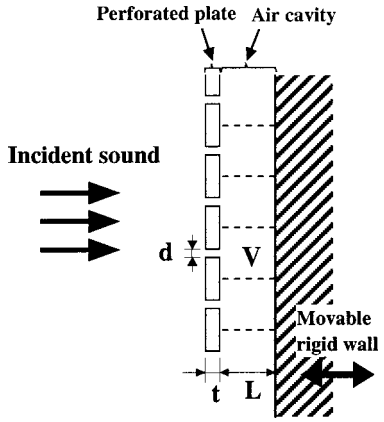


FIG. 1. Tunable acoustic absorber.

improvement in MEMS technology makes it possible to supply a large number of device structures in a batch process with reasonable cost. MEMS technology can contribute the following benefits:

(1) size and weight reduction and
(2) implementation of an intelligent system integrated with passive and active devices.

For instance, an increase of the aspect ratio of apertures in perforated plates of the Helmholtz resonators decreases the resonant frequency when cavity volumes are fixed. This means that it is possible to decrease the resonant frequency without increasing the cavity gap. We are interested in compact structures for adaptive passive acoustic systems working in a low-frequency range. In this paper, we estimate the possibilities of adaptive passive acoustic systems with microperforated devices as one of the steps towards the application of MEMS technology using a simple control method.

II. TUNABLE ACOUSTIC ABSORBER

A. Device design

The tunable acoustic absorber illustrated in Fig. 1 consists of a perforated plate with microapertures, an air cavity, and a rigid wall which performs as an array of the Helmholtz resonator. The resonator-type acoustic absorber transforms incident acoustic energy into thermal energy as a result of the resonance.[10,11] A single Helmholtz resonator consists of a cavity with a neck entrance. The neck part at the entrance of the structure works as acoustic resistance and inertance. An air cavity works as acoustic capacitance. The equivalent

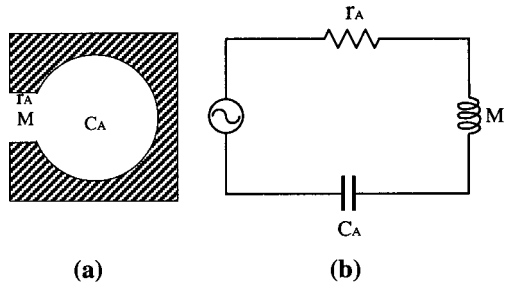


FIG. 2. Equivalent electrical circuit of a single Helmholtz resonator. (a) Helmholtz resonator. (b) Equivalent electrical circuit.

electrical circuit of the Helmholtz resonator is expressed as shown in Fig. 2. The total acoustic impedance Z_A is defined by the following equation:

$$Z_A = r_A + j\left(\omega M - \frac{1}{\omega C_A}\right), \tag{1}$$

where r_A is acoustic resistance, M is inertance, and C_A is acoustic capacitance.

A screw structure with a stepping motor is equipped behind the rigid wall in order to move the rigid wall. The screw structure transforms the rotation of the stepping motor to linear actuation of the rigid wall. The resonant frequency of the absorber is expressed below:[10]

$$f_0 = \frac{c}{2\pi}\sqrt{\frac{S}{V(t+0.85d)}} = \frac{c}{4}\sqrt{\frac{d^2}{\pi V(t+0.85d)}}, \tag{2}$$

where f_0 is resonant frequency, c is velocity of sound in air, S and d is area and diameter of an aperture, t is thickness of the perforated plate, and V is volume of the air cavity divided into each aperture. In this paper, we focus on volume V of the cavity, though other parameters can also be utilized. The relation between f_0 and V is simplified as

$$f_0 \propto \frac{1}{\sqrt{V}} \propto \frac{1}{\sqrt{L}}, \tag{3}$$

where L is the gap of the air cavity.

We designed narrow apertures with micrometer-order diameter in order to increase the aspect ratio of the apertures in a thin perforated plate. The absorber with narrower apertures can adapt lower-frequency noise at a fixed cavity gap. MEMS technology can achieve more than 100 in the aspect ratio in the microdomain. We succeeded in fabricating an aperture array with 0.1 mm in diameter, 1 mm in depth. Fundamental results using these microstructures with 10 in the aspect ratio were reported previously.[9] In this experiment, we used an aperture array with a 0.5 mm diam, 1 mm depth, in order to estimate the possibilities of adaptive passive acoustic systems with microstructures.

Based on Eq. (3), the resonant frequency of the absorber could be tuned by controlling the volume that is the gap of the air cavity. The stepping motor and the screw structure were used to control the gap of the air cavity. A single pulse step of the stepping motor can move the gap with 0.006 mm accuracy. The initial gap was set up at 5 mm when the number of pulses of the stepping motor is zero.

B. Control strategy

The conversion method $L(f)$ is employed in our developed system as mentioned at first. The optimal gap L of the absorber is determined when the desired frequency f of the sound is given. For instance, it is possible to relate L and the frequency of the detected incident sound wave through the microphone sensor. In order to construct the conversion method, it is required to clear the relation between the resonant frequency and the gap of the absorber through a data-

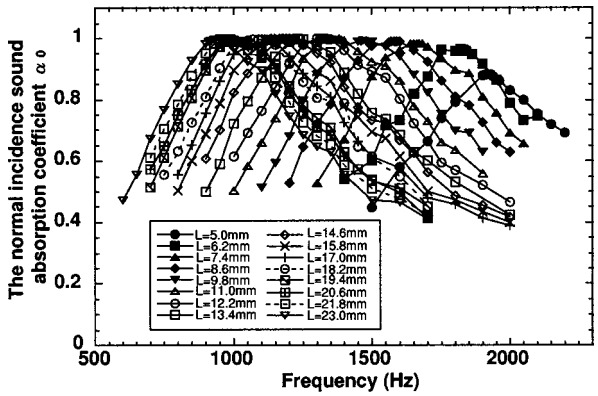


FIG. 3. Frequency response of α_0 at several gaps of the air cavity. L is the cavity gap.

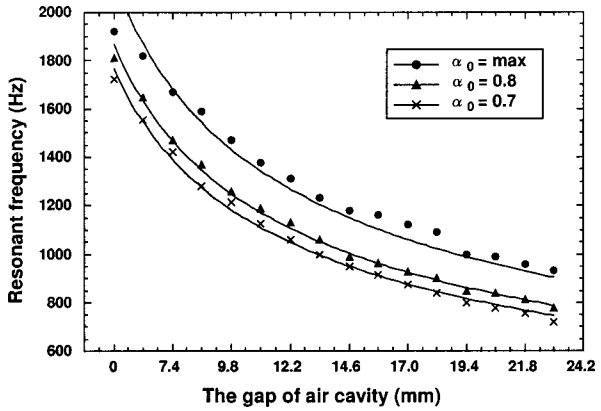


FIG. 4. Relation between the resonant frequency and the gap of the air cavity.

base construction. In addition, the function between the resonant frequency and the gap of the absorber can be obtained by applying the curve fit method to the database. The obtained function is to be putted into the computer in the system as described in the following section. It is effective for the system to calibrate the function $L(f)$ in order to keep constant characteristics periodically.

C. Database construction

As a first step, we tried fundamental experiments in order to construct the database for the control system. The normal-incidence sound absorption coefficient was used to estimate our developed absorber. The conventional standing-wave method with a traveling microphone was employed to measure the normal-incidence sound absorption coefficient. The impedance tube is a circular pipe with a 40 mm diam. The diameter of aperture d is 0.5 mm, and depth t is 1 mm. The test absorber is set at one end of the impedance tube where the movable rigid wall is equipped. At the other end, a loud speaker is set as a sound source. When a sine acoustic wave is transmitted from the sound source, a standing wave generates as a result of the resonance of the incident and reflected waves. Standing-wave ratio n is obtained by measuring the maximum and minimum of the sound pressure in the impedance tube through a microphone sensor. The normal-incidence sound absorption coefficient α_0 is obtained by using the following equations:[10]

$$\alpha_0 = 1 - |R|^2 = 1 - \left(\frac{n-1}{n+1}\right)^2 = \frac{4}{n + (1/n) + 2}, \tag{4}$$

$$n = P_{\max}/P_{\min}, \tag{5}$$

where R is the sound reflection coefficient, n is the standing-wave ratio, and P is the sound pressure.

In the experiment, we measured each normal-incidence sound absorption coefficient α_0 at 16 points. The points were set at every 1.2 mm gap of the absorber by controlling the stepping motor. Figure 3 shows the experimental results. The resonant frequency of the absorber corresponding to each gap of the air cavity was estimated as a peak of α_0. The resonant frequency decreased according to the gap of the air cavity. For example, the resonant frequency was shifted about 1000 Hz by tuning the gap of the air cavity from 5 to 23 mm. In addition, most of the normal-incidence sound absorption coefficients of the estimated absorber were more than 0.9. Thus, we see that it is possible to tune the resonant frequency of the absorber by this system.

III. FITTING

In the developed system, we employed the conversion method from frequency to gap $(f-L)$ in order to determine the desired structural parameters of the absorber. This method means that the optimal gap of the absorber (that is, the number of pulses of the stepping motor) can be calculated by function $L(f)$ when the desired frequency of the sound wave is given. Therefore, it was necessary to define the function between the resonant frequency of the absorber and the number of pulses of the stepping motor from the zero point. Figure 4 shows the relation between the resonant frequency and the gap of the air cavity. The plots indicate the data from the database experiments in Fig. 3. The solid curves indicate fitted functions based on Eq. (2). In this experiment, three kinds of functions (α_0 = maximum, 0.8, 0.7) of the sound absorption coefficient were calculated, because it is also significant for a system to adapt α_0 to various values except the maximum α_0. There exist two candidates of $L_i(f)$ for a certain (f_i, α_i) except the case of maximum α_0. The objective sound absorption coefficient α_i at frequency f_i can be realized by two conversion functions $L_1(f)$ and $L_2(f)$. The difference between these two candidates is that the cavity gap L_1 is thicker than L_2. For example, $(f_i = 1650\,\text{Hz}, \alpha_i = 0.8)$ is provided by $L_1 = 11.0\,\text{mm}$ or $L_2 = 6.2\,\text{mm}$ in our system, as shown in Fig. 3. We chose a lower frequency at cavity gap L_2 from the two candidates because the system can be adapted to lower frequencies with thinner cavity gaps.

IV. SYSTEM CONTROL

A. Adaptive passive noise control system

Figure 5 shows a schematic drawing of our developed adaptive passive noise control system using a compact tunable acoustic absorber. The measurement setups are shown together in Fig. 5. The calculated functions between the resonant frequency and the gap of the air cavity by the curve fit method were put into a computer in advance, as shown in

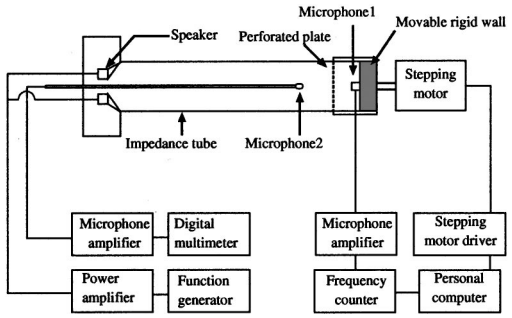


FIG. 5. Adaptive passive noise control system with a measurement setup.

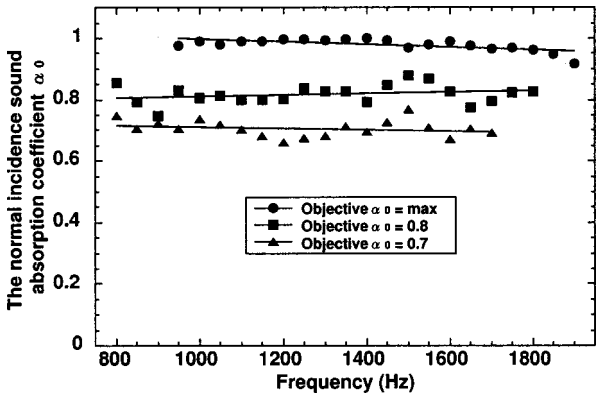


FIG. 6. Control results on α_0 by the proposed system.

Fig. 5. The frequency of the incident sound detected by microphone 1 on the rigid wall is obtained through the frequency counter. After the conversion calculation in the computer, the movable rigid wall is adjusted to the optimal position so as to reduce the noise efficiently. Furthermore, this system can adjust the parameter of the absorber not only to the maximal sound absorption coefficient but also to desired one. For example, $\alpha_0=0.8$ and 0.7 are investigated in this report. It becomes possible for the system to provide the desired sound absorption characteristics of the absorber.

The response speed of the system control is one of the important factors in estimating system performance. The response time to move the rigid wall to the next position depends on the speed of the stepping motor. In the experiment, it took around 1 s to move the rigid wall from 5 to 23 mm. The sampling time to detect the frequency of the incident sound wave was, of course, more than 1 s. Improvement of the response time of an actuator makes it possible to increase the response speed of the adaptive passive noise control system.

B. Control results

We measured the normal-incidence sound absorption coefficient α_0 under changing frequency conditions in order to estimate our system performance. The developed system absorbed over a fixed range of frequencies with varying cavity gap L. We can see the resonance of the Helmholtz resonator through measuring the peak of the sound absorption coefficient. Generally, it is regarded that a higher α_0 is better for noise reduction. It, however, is also significant to consider a system that can adapt α_0 of our system to various values: maximum, 0.8, and 0.7. Figure 6 shows the experimental control results when objective α_0 was maximum, 0.8, and 0.7. Both the resonant frequency and the sound absorption coefficient of the system can be controlled by changing cavity gap L. The developed system tuned absorption frequencies from 800 to 1800 Hz by controlling the cavity gap from 5 to 23 mm. Data were measured every 50 Hz. The measured sound absorption coefficients were in error by less than 10% in each experiment. It can be seen from these results that the developed system could be controlled effectively.

V. CONCLUSIONS

We investigated the possibilities of a compact adaptive passive noise control system taking into account of advantages of microstructures based on MEMS technology. A simple open-loop method for system control was employed for our estimation. The developed system employed high-aspect-ratio microapertures in a perforated plate of a resonator and adapted for the lower-frequency range without increasing the cavity gap. The system could absorb less than 1000 Hz frequency noises with a thin cavity gap L, 23 mm. We employed the simple conversion method to decide the structural parameters for the required system characteristics. Our employed control algorithm was so simple that the system required few extra hardware. The system control could be accomplished in error by less than 10% using the simple open-loop control algorithm in the control experiments. We can say that our developed compact adaptive passive noise control system is effective as a side branch of a duct and an aeroacoustic application. More compact adaptive passive noise control systems attenuating lower frequency can be obtained with improvement of the aspect-ratio apertures by MEMS technologies.

[1] W. Neise and G. Koopman, ''Reduction of centrifugal fan noise by use of resonators,'' J. Sound Vib. **73**, 297–308 (1980).
[2] G. Koopman and W. Neise, ''The use of resonators to silence centrifugal blowers,'' J. Sound Vib. **82**, 17–27 (1982).
[3] B. E. Walker, A. S. Hersh, L. J. Heidelberg, D. L. Sutliff, and M. E. Spencer, ''Active resonators for control of multiple spinning modes in an axial flow fan inlet,'' AIAA Paper 99-1853, Proceedings of the Fifth AIAA/CEAS Aeroacoustic Conference, 339–348 (1999).
[4] S. Sato and H. Matsuhisa, ''Semiactive noise control by a resonator with variable parameters,'' Proc. Internoise **90**, 1305–1308 (1990).
[5] W. H. Matsuhisa, B. Ren, and S. Sato, ''Semiactive control of duct noise by a volume-variable Resonator,'' Jpn. Soc. Mech. Eng. Int. J. **35**, 223–228 (1992).
[6] E. Little, A. Kashani, J. Kohler, and F. Morrison, ''Tuning of an electrorheological fluid-based intelligent Helmholtz resonator as applied to hydraulic engine mounts,'' ASME DSC Transp. Syst. **54**, 43–51 (1994).
[7] J. S. Lamancusa, ''An actively tuned, passive muffler system for engine silencing,'' Proc. Noise Control **87**, 313–318 (1987).
[8] J. 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).
[9] S. Konishi, M. Yoda, M. Hosaka, S. Sugiyama, and S. Akishita, ''Tunable acoustic absorbing system using a deep hole array,'' Proceedings of the 11th IEEE International Conference on Microelectromechanical Systems (1998), pp. 655–660 (unpublished).
[10] M. Nagata *et al.*, *Architectural Acoustics* (Corona, 1995).
[11] M. Ohta *et al.*, *Physics of Acoustic Engineering* (Asakura, 1990), p. 208.