CODEN: JASMAN

THE JOURNAL of the Acoustical Society of America

ISSN: 0001-4966

Vol. 101, No. 1 January 1997

(Continued)

CONTENTS—*Continued from preceding page*

AEROACOUSTICS, ATMOSPHERIC SOUND [28]

UNDERWATER SOUND [30]

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

TRANSDUCTION [38]

STRUCTURAL ACOUSTICS AND VIBRATION [40]

(Continued)

CONTENTS—*Continued from preceding page*

(Continued)

CONTENTS—*Continued from preceding page*

NOTES CONCERNING ARTICLE ABSTRACTS

1. The number following the abstract copyright notice is a Publisher Item Identifier (PII) code that provides a unique and concise identification of each individual published document. This PII number should be included in all document delivery requests for copies of the article.
2. PACS numbers are for subject classification and indexing. See June and December issues for detailed listing of acoustical classes and subclasses.
3. The initials in brackets following the PACS numbers are the initials of the JASA Associate Editor who accepted the paper for publication.

ACOUSTICAL NEWS—USA

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

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

Announcement of the 1997 Election

In accordance with the provisions of the bylaws, the following Nominating Committee was appointed to prepare a slate for the election to take place on 6 June 1997:

Jiri Tichy, *Chair* — John Erdreich
Ira Dyer — Christy K. Holland
David M. Green — Alexandra I. Tolstoy

The bylaws of the Society require that the Executive Director publish in the *Journal* at least 90 days prior to the election date an announcement of the election and the Nominating Committee's nominations for the offices to be filled. Additional candidates for these offices may be provided by any Member or Fellow in good standing by letter received by the Executive Director not less than 60 days prior to the election date, and the name of any eligible candidate so proposed by 20 Members or Fellows shall be entered on the ballot.

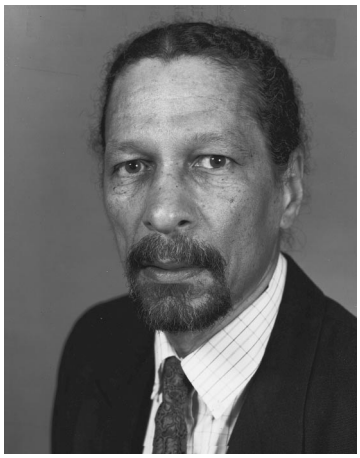
Biographical information on the candidates and statements of objectives of the candidates for President-Elect and Vice President-Elect will be mailed with the ballots.

CHARLES E. SCHMID
Executive Director

The Nominating Committee has submitted the following slate:

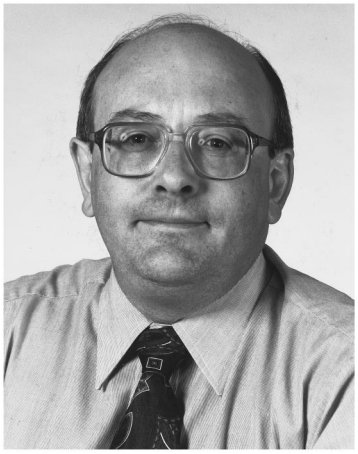
FOR PRESIDENT-ELECT

William J. Cavanaugh

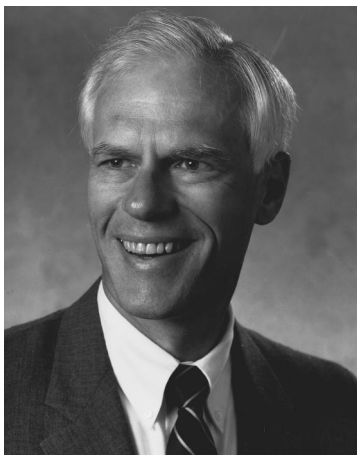


James E. West

FOR VICE PRESIDENT-ELECT

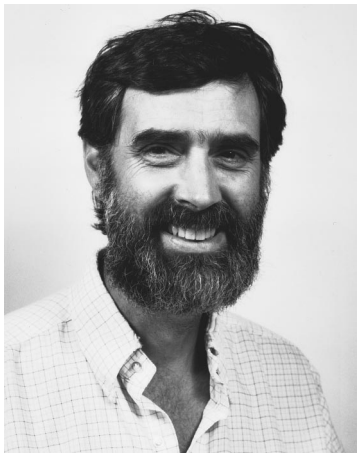


Gilles A. Daigle

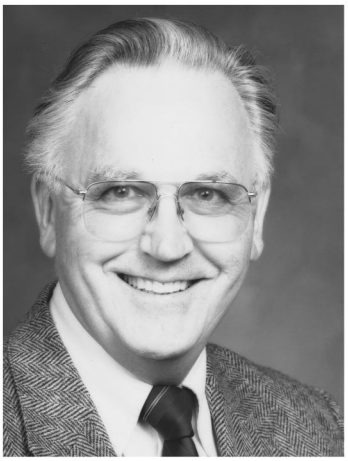


William M. Hartmann

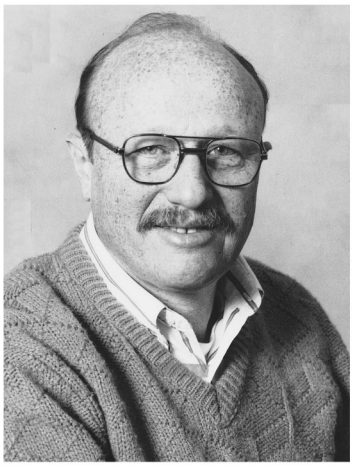
 0001-4966/97/101(1)/1/11/$10.00



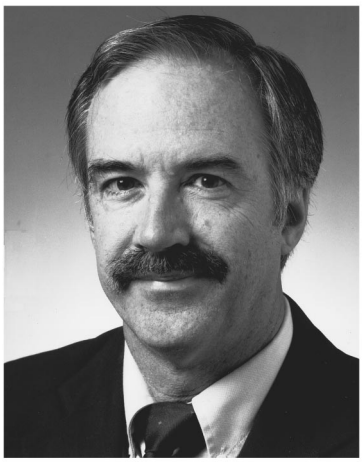
Joseph W. Dickey



Uwe J. Hansen



William A. Kuperman



Julian D. Maynard

Sigfrid D. Soli

Janet M. Weisenberger

Frederick H. Fisher receives IEEE award

The IEEE Oceanic Engineering Society Distinguished Technical Achievement Award was awarded to Frederick H. Fisher in October 1996 for his outstanding contributions to the field of Ocean Acoustics.

Fred Fisher is a Fellow of the Acoustical Society of America (ASA) and served as its President from 1983–84. He has also held the offices of President-Elect (1982–83), Vice President (1980–81), and Vice President-Elect (1979–80) and served as a Member of the ASA Executive Council (1976–79). Dr. Fisher has served on several ASA Committees and was Chair of the Fall 1990 meeting held in San Diego.

Fred Fisher received B.S. and Ph.D. degrees in 1949 and 1957 from the University of Washington. He became head of a research group at the Marine Physical Laboratory (MPL) of the Scripps Institution of Oceanography, University of California, in 1958, became Associate Director of MPL in 1974, and was Deputy Director of MPL from 1989 to 1994. He was Scientific Officer and codesigner of the manned ocean buoy, FLIP, and was responsible for working out the ''flipping'' operation with 350-ft-long 1/10 scale models.

Since 1965 Dr. Fisher has been a Principal Investigator for National Science Foundation grants devoted to high-pressure measurements related to the physical chemistry of sound absorption in seawater due to magnesium sulfate and other salts. He has also been a Principal Investigator on various Office of Naval Research contracts related to measurements of sound propagation in the ocean.

James Flanagan receives honorary doctorate in computer science

Dr. James Flanagan, former President of the Society and ASA Gold Medalist, was awarded Doctor Honoris Causa from the University of Paris-Sud in ceremonies at the Orsay campus on 4 October 1996. President of the University, Dr. Alain Gaudemer, noted Flanagan's research and academic contributions in bestowing the degree. Flanagan currently serves as Vice President for Research at Rutgers University, New Brunswick, NJ. Earlier this year he was presented the National Medal of Science by President Clinton in ceremonies at the White House. He was cited ''for his contributions to research in speech communication and for leadership in applications to telecommunications.''

Regional Chapters news

Delaware Valley: The Delaware Valley Chapter held the first dinner meeting of the season on Wednesday, 16 October 1996 at Williamson's in Bala Cynwyd, PA, with 12 members and guests in attendance. Dr. John Barry of the Philadelphia Office of OSHA discussed the new NIOSH recommendation for changes in the OSHA noise standards. This involves lowering the threshold to 85 dB and changing from a 5- to 3-dB exchange rate for computing exposure. He showed that the noise dose grows very rapidly

with level increase with the 3-dB rule. The difficulties in obtaining noise dosimeter data for emergency response workers under field conditions were discussed; a major problem was with radio transmitters which caused large errors in some dosimeters. Some interesting slides of fire stations and equipment were shown.

JON SANK

Greater Boston: ''The 25th Anniversary Celebration of the Greater Boston Chapter of the Acoustical Society of America (GBC-ASA) was held at Northeastern University in Boston, Massachusetts. About 70 members attended a dinner and lecture, including pioneers of the Chapter such as Dick Bolt and Leo Beranek. Professor Kenneth Stevens presented a fascinating lecture entitled, ''Human and Machine Generation of Speech.'' Following a lively discussion, he was awarded a life membership to the Greater Boston Chapter of ASA (see Fig. 1). William Cavanaugh was also honored with life membership for his untiring dedication to the Chapter. The celebration was sponsored by GBC-ASA, the Department of Speech–Language Pathology and Audiology at Northeastern University, and the Communication Digital Signal Processing center at Northeastern University.

FIG. 1. Professor Kenneth Stevens (center) is awarded life membership in the Greater Boston Chapter of the Acoustical Society of America by Mary Florentine and William Cavanaugh.

MARY FLORENTINE

Madras: The birth of the Madras Regional Chapter of the Acoustical Society of America took place at I.I.T., Madras, India in April 1995. Hari Paul, Department of Mathematics, I.I.T. (M), is the Chapter's Chair and serves also as Chapter Secretary (see Fig. 1). Patrons are Prof. N. V. C. Swamy and Prof. R. Natarajan.

Chapter supporters are D. P. Raju, K. Natarajan, S. N. Majhi, C. P. Vendhan, S. Narayanan, N. Muthiyalu, A. Namasivayam, K. S. Subramanian, Baldev Raj, and Gautan Suri. Thirty other Indian scientists joined in the formation of the chapter. Twenty-seven papers were presented in the first technical sessions on 19 August 1995, which included invited papers, a distinguished lecture on Vedhic Acoustics, and a music concert.

FIG. 1. Robert Apfel (l, ASA President 1995–96) congratulates Hari S. Paul (r) on organization of Madras Regional Chapter of the Acoustical Society of America.

HARI S. PAUL

Position Open

Director of Research: Central Institute for the Deaf (CID) seeks a Director of Research to head the department of research. This position offers exciting opportunities to lead this growing department through new research frontiers. CID is a major research and education center founded in 1914 to address the needs of the hearing impaired and is an affiliate of Washington University Medical Center. Current research programs involve applied studies in childhood deafness, sensory neuroscience, speech and hearing, noise, and sensory electrophysiology. The successful candidate will be expected to organize the department into centers of excellence. The center for childhood deafness and oral rehabilitation is already well organized and is ongoing. Three new faculty members are currently being recruited for the second center emphasizing molecular and cellular approaches to degeneration, rescue, and repair. A third center focused on complex auditory cognition is planned. Additionally, the Director will be responsible for overseeing research programs and obtaining external funding. A full-time research administrator will assist the Director in administering departmental policies. Teaching in the professional education program is an option. Maintenance of an independent research program is expected. Candidates must have a doctoral degree in hearing sciences or related field and have an outstanding record of scientific accomplishments and other achievements, a strong record of research funding, excellent communication skills, knowledge of research techniques, methodology and equipment, and the capacity to provide leadership essential to the continued development of research and creative activity at the Institute.

We offer a competitive salary commensurate with experience and qualifications along with an excellent benefit package. The position is available immediately. Please send a letter of interest, curriculum vitae, and names and address of three references to:

Search Committee
Research Director Position
Donald W. Nielsen, Ph.D., Chair
Central Institute for the Deaf
818 S. Euclid
St. Louis, MO 63110

EOE/M/F/H/V

USA Meetings Calendar

Listed below is a summary of meetings related to acoustics to be held in the U.S. in the near future. The month/year notation refers to the issue in which a complete meeting announcement appeared.

1997

27–28 Feb.	Penn State Ultrasonic Transducer Engineering Workshop, Newport Beach, CA [Donna Rode, SPIE, P.O. Box 10, Bellingham, WA 98227-0010, Tel.: 360-676-3290; E-mail: donnar@mom.spie.org or K. Kirk Shung, 231 Hallowell Bldg., Penn State Univ., University Park, PA 16802, Tel.: 814-865-1407; E-mail: kksbio@engr.psu.edu].
13–16 April	23rd International Symposium on Acoustical Imaging, Boston, MA [Sidney Lees, Bioengineering Dept., Forsyth Dental Ctr., 140 Fenway, Boston, MA 02115; FAX: 617-262-4021; E-mail: slees@forsyth.org]. *7/96*
12–14 May	Third AIAA/CEAS Aeroacoustic Conference, Atlanta, GA [Dr. Stephen Engelstad, Lockheed Martin Aeronautical Systems, D/73-47, Z/O-685, Marietta, GA 30063, Tel.: 770-494-9178; FAX: 770-494-3055; E-mail: sengelstad@fs2.mar.1mco.com].
15–20 June	Eighth International Symposium on Nondestructive Characterization of Materials, Boulder, CO [Debbie Harris, The Johns Hopkins University, Ctr. for Nondestructive Evaluation, 102 Maryland Hall, 3400 N. Charles St., Baltimore, MD 21218, Tel.: 410-516-5397; FAX: 410-516-7249, E-mail: cnde@jhuvms.hcf.jhu.edu].

15–17 June	NOISE-CON 97, State College, PA [Institute of Noise Control Engineering, P.O. Box 320, Arlington Branch, Poughkeepsie, NY 12603, Tel.: 914-891-1407; FAX: 914-463-0201].
16–20 June	133rd meeting of the Acoustical Society of America, State College, PA [ASA, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; FAX: 516-576-2377; E-mail: asa@aip.org, WWW: http://asa.aip.org].
9–13 July	International Clarinet Association, Texas Tech Univ., Lubbock, TX [Keith Koons, Music Department, Univ. of Central Florida, P.O. Box 161354, Orlando, FL 23816-1354, Tel.: 407-823-5116; E-mail: kkoons@pegasus.cc.ucf.edu].
7–11 Sept.	American Academy of Otolaryngology—Head and Neck Surgery, San Francisco, CA [American Academy of Otolaryngology—Head and Neck Surgery, One Prince St., Alexandria, VA 22314, Tel.: 703-836-4444; FAX: 703-683-5100].
22–24 Sept.	Second Biennial Hearing Aid Research and Development Conference, Bethesda, MD [National Institute of Deafness and Other Communication Disorders, 301-970-3844; FAX: 301-907-9666; E-mail: heaingaid@tascon.com]. Deadline for abstracts is 15 March.
1–5 Dec.	134th meeting of the Acoustical Society of America, San Diego, CA [ASA, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; FAX: 516-576-2377; E-mail: asa@aip.org, WWW: http://asa.aip.org].

1998

22–26 June	135th meeting of the Acoustical Society of America/ 16th International Congress on Acoustics, Seattle, WA [ASA, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; FAX: 516-576-2377; E-mail: asa@aip.org, WWW: http://asa.aip.org].
13–17 Sept.	American Academy of Otolaryngology—Head and Neck Surgery, San Francisco, CA [American Academy of Otolaryngology—Head and Neck Surgery, One Prince St., Alexandria, VA 22314, Tel.: 703-836-4444; FAX: 703-683-5100].
12–16 Oct.	136th meeting of the Acoustical Society of America, Norfolk, VA [ASA, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; FAX: 516-576-2377; E-mail: asa@aip.org, WWW: http://asa.aip.org].

Revision to Membership List

New Associates

Absher, Richard G., 2 Crescent Road, Burlington, VT 05405

Acosta, Luis M., 11510 Victory Boulevard, North Hollywood, CA 91606

Akahane-Yamada, Reiko, Dept. 1, ATR Human Information Processing, Research Labs., 2-2 Hikaridai, Seika-cho, Soraku, Kyoto, 619-02 Japan

Alach, D. Robert, Alactronics, Inc., 192 Worcester Street, Wellesley Hills, MA 02181

Alpizar, Javier, N y E Omicron S.A. de C.V., Irapuato #14, Col. Condesa, Mexico DF C.P., 06170 Mexico

Apelian, Chahe V., 6660 Zelzah Avenue, Reseda, CA 91335

Appelbaum, Irene, Philosophy Dept., Washington University, Campus Box 1073, One Brookings Drive, St. Louis, MO 63130

Arrieta, Lisa L., Code 130B, Coastal Systems Station, 6703 West Highway 98, Panama City, FL 32407-7001

Ball, Martin J., School of Behavioural & Communication Sciences, Univ. of Ulster, Jordanstown, Newtownabbey BT37 0QB, N. Ireland

Bautista, Edgar O., 35-50 169th Street, Flushing, NY 11358

Betancourt, Octavio L., Computer Science, CUNY, 138th Street & Convent Avenue, New York, NY 10025

Bi, Ning, Systems Engineering, Qualcomm, Inc., 6455 Lusk Boulevard, San Diego, CA 92121

Bishop, Joseph R., Ocean Acoustics Division, NOAA/AOML, 4301 Rickenbacker Causeway, Miami, FL 33149

Blihi, Said, Expandable Polystyrene, Shell Research, Avenue Jean Monnet 1, Ottignies, Lovain 1a Neuve, B-1348 Germany

Bobrovnitskii, Yuri I., Russian Academy of Sciences, Inst. of Engineering Research, Lab. of Structural Acoustics, Griboedov Street 4, Moscow 101830, Russia

Breed, David S., Automotive Technologies, Inc., P.O. Box 8, Denville, NJ 07834

Burgemeister, Kym A., Uve Arup & Partners, Arup Acoustics, P.O. Box Q116, Q. V. B. Post Office, Sydney, NSW 1230, Australia

Burlage, Kirk G., Digisonix, Inc., 8401 Murphy Drive, Middleton, WI 53562

Bushek, Don, St. Croix Medical, 5155 East River Road, Minneapolis, MN 55421

Carlson, Thomas J., Land & Water Resources, Pacific Northwest National Lab., MSIN:K6-85, Battelle Boulevard, Richland, WA 99352

Carroll, Linda M., 133 West 78th Street, Suite 2, New York, NY 10024-6748

Cazals, Yves A. G., Lab. d'Audiologie Experimentale, Univ. de Bordeaux 2, Hopital Pellegrin, Bordeaux 33076, France

Chan, Daisy W. Y., B/2, 12/F 905 King's Road, Quarry Bay, Hong Kong

Chapman, Robert C., 11721 West Brandt Avenue, Littleton, CO 80127

Choi, Myoung S., Hanbit Apt. 103-1002, Eun-Dong, Yusong-gu, Taejon 305-333, Korea

Choi, Min Joo, College of Medicine, Dept. of Biomedical Engineering, Seoul National Univ., 28 Yongon-Dong, Chongno-gu, Seoul 110-744, Korea

Connolly, Sean M., David L. Adams Associates, 1701 Boulder Street, Denver, CO 80211

Creamer, Dennis B., Naval Research Lab., Code 5583, 4555 Overlook Avenue, South, Washington, DC 20375

Crocker, Steven E., 7695 White Oak Lane, Delavan, IL 61734

D'Angelo, William R., U. S. Airforce Armstrong Lab., AL/CFBA Bldg, 441, 2610 7th Street, Wright-Patterson AFB, OH 45433

Davies, Peter O. A. L., Inst. of Sound and Vibration Research, Southampton Univesrity, Highfield, Southampton S017 1BJ, U.K.

de Cheveigne, Alain, CNRS/Univ. Paris 7, Case 7003, 2 place Jussieu, Paris 75251, France

De Leon, Phillip L., Klipsch School of Electrical and Computer Engineering, New Mexico State Univ., Box 30001/Dept. 3-0, Las Cruces, NM 88003-8001

Delannoy-Arriagada, Jaime, Acoustics, Univ. Tecnologica Vicente, Perez Rosales, Brown Norte 290, Santiago, Nunoa, Chile

Drzewiecki, Tadeusc M., Defense Research Technologies, Inc., 354 Hungerford Drive, Rockville, MD 20850

Dunn, Margaret H., 25 Clark Street, New Haven, CT 06511

Dymkin, Gregory Ya., Ultrasonic Nondestructive Testing, Bridge Research Inst., Fontanka 113, St. Petersburg 190031, Russia

Ebbini, Emad S., Electrical Engineering & Computer Science, Univ. of Michigan, 1301 Beal Avenue, Ann Arbor, MI 48109-2122

Eldridge, Mark C., 1656 Birchmont Lane, Keller, TX 76248

Elias, Bartholomew, AL/OEBN, 2610 Seventh Street, Wright-Patterson AFB, OH 45433-7901

Elias-Juarez, Alfredo A., Acoustics & Vibration Div., Centro Nacional de Metrologia, Rm. 4.5 carr. a los cues, Queretaro 76000, Mexico

Engdahl, Bo L., Velandsgt. 54-4, Oslo 0457, Norway

Fetridge, Guild A., 507 Old Country Road, Elmsford, NY 10523

Fink, Rainer J., Engineering Technology, Texas A & M University, Fermier Hall 111B, College Station, TX 77843

Flemming, Edward S., Dept. of Linguistics, Stanford University, Building 460, Stanford, CA 94305-2150

Flynn, Patrick J., Electrical Engineering & Computer Science, Washington State University, Pullman, WA 99164-2752

Fung, Wai Shun, 15 Watson Road, Victoria Centre, 35th Floor, Flat B, Block 2, Hong Kong

Gallotta, Richard A., Marine Acoustics, Inc., 2345 Crystal Drive, Suite 901, Arlington, VA 22202

Garrison, Laurie F., AT&T, Voice Quality Assessment Lab., 101 Crawfords Corner Road, Room 3D501A, Holmdel, NJ 07733

Gatehouse, Stuart, 6 Bedale Road, Garrowhill, Glasgow G69 7LU, U.K.

Gavin, Joseph R., 166 Litton Avenue, Groton, CT 06340

Giangrande, Janice, 2301 South Congress Avenue, #1613, Boynton Beach, FL 33426

Goacher, Geoffrey K., Sound Advice Acoustical Consultants, 15625 Alton Parkway, Suite B, Irvine, CA 92718

Godfrey, Richard D., Science & Technology, Owens-Corning, 2790 Columbus Road, Route 16, Granville, OH 43023-1200

Gonzalez Suarez, Mercedes, Marques De La Ensenada No, 3-3-1a, Castellon de la Plana 12003, Spain
Goold, John C., 24 Gwel Eryri, Llandegfan, Anglesey LL59 5PY, U.K.
Gow, David W., 24 Dearborn Road, Medford, MA 02155
Grieser, George A. M., Andre Grieser Arquitetura & Consultoria, Acustica Ltda., Av. Senador Virgilio Tavora 50, Sobreloja 3, Fortaleza, Ceara 60170-250, Brazil
Han, Ruijing, MEMC Electronic Materials, Inc., 501 Pearl Drive, St. Peters, MO 63376
Handel, Stephen J., Psychology Dept., Univ. of Tennessee, Austing-Peay Bldg., Knoxville, TN 37996-0900
Hauenstein, John R., JRH Acoustical Consulting, Inc., 312 West 51st Street, #1R, New York, NY 10019
Hayward, Gordon, Electronics & Electrical Engineering, The Univ. of Strathclyde, 204 George Street, Glasgow G11XW, Scotland, U.K.
Hill, Ryan B., Espey, Huston & Associates, Inc., 206 Wild Basin Road, Suite 300, Austin, TX 78746
Hirobayashi, Shigeki, Dept. of Electrical & Computer Eng., Kanazawa University, 2-40-20 Kodatsuno, Kanazawa-shi, Ishikawa-ken, 920 Japan
Hoefler, Jeffrey J., EPD, Bose Corporation, The Mountain, Mail Stop 2, Framingham, MA 01701
Homola, Christopher A., Mechanical Engineering, Byers Engineering Company, 1621 Euclid Avenue, Suite 300, Cleveland, OH 44115
Hu, Qiang, Dept. of Mechanical Engineering, Wayne State Univ., Detroit, MI 48202
Hudig, J. Mauritz, 44 Edgemore Road, Montclaire, NJ 07042
Ike, Frank W., 156 Cardinal Road, Chalfont, PA 18914
Imhof, Matthias G., MIT Ear Resources Lab., E34-370, 42 Carleton Street, Cambridge, MA 02142-1324
Johnson, Ralph R., NRaD, Code 44213, 53245 Patterson Road, San Diego, CA 92152
Kalb, Joel T., 735 Marian Drive, Aberdeen, MD 21001
Kallistratova, Margarita A., Udaltsova 10, ap. 131, Moscow 117415, Russia
Kato, Hiroaki, ATR Human Information Processing, Research Labs., Dept. 1, 2-2 Hikaridai, Seikacho, Sorakugun, Kyoto, 619-02 Japan
Kennedy, Rodney A., Telecommunications Engineering, Australian National University, Canberra, ACT 0200, Australia
Kennedy, Joel A., St. Croix Medical, 5155 East River Road, Minneapolis, MN 55421
Kessel, Ronald T., 95 Caton Place, Victoria, BC V9B 1L1, Canada
Kidd, Dennis E., 6141 Archer, NE, Rockford, MI 49341
Kim, Bong-Chae, Physical Oceanography Div., Korea Ocean Research & Development Inst., 1270, Sa-Dong, Ansan-Shi, Kyounggi-Do 425-170, Korea
Kim, Tae-Gun, International Processional Associates, International Business Dept., c/o Smart Cargo Service, 1302 Mahalo Place, Campton, CA 90220
Kitagawa, Hiroo, Research and Development Center, Ricoh Co. Ltd., 415 Research Group, 3-2-3 Shin-yokohama, Kohoku-ku, Yokohama, 222 Japan
Kompis, Martin, Goldenthalweg 16, Oberwill BL 4104, Switzerland
Korber, Dennis J., Glass Engineering, PPG Industries, P.O. Box 11210, Guys Run Road, Pittsburgh, PA 15238-0472
Kroll, Kai, St. Croix Medical, 5155 East River Road, Suite 418, Minneapolis, MN 55421
Kwok, Chung Wai, Mechanical Engineering, Hong Kong Univ. of Science and Technology, Clear Water Bay, Kowloon, Hong Kong
Kyriakakis, Chris, 1250 Elizabeth Street, Pasadena, CA 91104
Labrozzi, Danielle M., 118 Bell Vista Drive, Cary, NC 27513
Lage, Maria Oti'Lia P., Inst. Politecnico do Porto, Ser. de Documentacao e Publicaocoes, Rue dr. Roberto Frias, Porto 4200, Portugal
Lai, Peter C.-C., 2335-F Collins Drive, Sidney, OH 45365
Lally, Michael J., The Modal Shop Inc., 1776 Mentor Avenue, Suite 170, Cincinnati, OH 45212
Landis, Donald H., Epic Metals Corp., 11 Talbot Avenue, Rankin, PA 15104
Langley, Robin S, Dept. of Aeronautics and Astronautics, Univ. of Southampton, Southampton, Hamsphire SO17 1BJ, England
Large, Edward W., Inst. for Research in Cognitive Science, Univ. of Pennsylvania, 3401 Walnut Street, Suite 301C, Philadelphia, PA 19104
Lee, James J., 44785 Glengarry Road, Canton, MI 48188
Lee, Robert A., AL/OEBN, Noise Effects Branch, Armstrong Laboratory, 2610 Seventh Street, Wright-Patterson AFB, OH 45433-7901
Lei, Ming, 9308 Macallan Road, NE, Albuquerque, NM 87109
Lenzi, Arcanjo, Rue Dos Jasmins, NO. 539, Corrego Grande, 88037-140 Florianopolis SC, Brazil
Leotta, Antonio, Via Bainsizza No. 30/4, Genova 16147, Italy
Lilley, Geoffrey M., Dept. of Aeronautics & Astronautics, Univ. of Southampton, Highfield, Southampton SO17 1BJ, U.K.
Lima, Scott K., Geco Defence A.S., Kvassenseveien 27, Isdalsto 5100, Norway
Mackersie, Carol L., Center for Research in Speech and Hearing Sciences, Graduate Center, CUNY, 33 West 42nd Street, New York, NY 10036
Marcus, Larry A., Lucent Bell Laboratories, P.O. Box 50427, 6602 East 75th Street, Indianapolis, IN 46250-0427
Martin, Steven B., 3384 Lawrencetown Road, RR #2, Porters Lake, NS B0J 2S0, Canada
Mauney, Daniel W., 618 Bonair Place, La Jolla, CA 92037
Maze, Gerard G., Rue Lacepede 6, Le Havre 76600, France
McDermott, Hugh J., Dept. of Otolaryngology, Univ. of Melbourne, 384-388 Albert Street, East Melbourne, VIC 3002, Australia
McKinley, Bruce L., 12003 Johns Place, Fairfax, VA 22033
Mees, Paul, Albrecht Rodenbachstraat 71, Leuven 3010, Belgium
Messner, Douglas R., Corporate Quality Engineering, Shure Brothers, Inc., 222 Hartrey Avenue, Evanston, IL 60202-3696
Meyerson, Scott C., St. Croix Medical, 5155 East River Road, Minneapolis, MN 55421
Miles, Ronald N., Mechanical Engineering, SUNY, P.O. Box 6000, Binghampton, NY 13902-6000
Miller, Joel D., Sterling Software, NASA Ames Division, MS 262-6, Moffett Field, CA 94035-1000
Moon, Won-kyu, Precision Mechanics Lab., Samsung Advanced Inst. of Tech., P.O. Box 111, Suwon 440-600, Korea
Moquin, Philippe, 59 Acklam Terrace, Kanata, ON K2K 2H7, Canada
Moulin, Annie M., 13 Chemin du Barray, Brignais 69530, France
Munin, Anatoli G., Myasnitskay str, 40-a, ap. 48, Moscow 101000, Russia
Munoz-Madrigal, Norma C., Centro Nacional de Metrologia, Metrologia Fisica, Div. de Vibraciones y Acustics, Rm. 4.5, Carretera a los cues, El Marques, Queretaro 76900, Mexico
Muzzey, Gary C., Research & Development, Tibbetts Industries, Inc., P.O. Box 1096, Colcord Avenue, Camden, ME 04843
Nakayama, Kiyoshi, Dept. Electrical & Electronic Eng., Sophia University, 7-1 Kioicho, Chiyodaku, Tokyo, 102 Japan
Nelson, Bradley A., Sound Solutions Northwest, Inc., 3180 West Clearwater Avenue, Suite J, Kennewick, WA 99336
Nguyen, Phuong-Tao N., 91-1033 Kuea Street, Kapolei, HI 96707
Ni, Jun, Electrical and Computer Engineering, Univ. of Wollongong, Northfields Avenue, Wollongong, NSW 2522, Australia
Obeng, Samuel G., Linguistics Dept., Indiana University, Memorial Hall 326, Bloomington, IN 47405
Odom, Robert I., Applied Physics Lab., Univ. of Washington, 1013 NE 40th Street, Seattle, WA 98105
Ogawa, Hitoshi, Electronic Control Engineering, Hiroshima National College of Maritime, Technology, Higashino-cho 4272-1, Toyota-gun, Hiroshima Pref., 725-02 Japan
Olson, Bruce C., Olson Sound Design, 8717 Humboldt Avenue, North, Brooklyn Park, MN 55444-1320
Orth, Scott D., Audio and Acoustics, Delco Electronics, 1800 East Lincoln, M/S R104, Kokomo, IN 46904
Ostashev, Vladimir E., Physics Dept., New Mexico State Univ., Box 30001, Dept. 3D, Las Cruces, NM 88003-8001
Palakal, Mathew J., Computer Science Dept., Purdue University, 723 West Michigan Street, SL280, Indianapolis, IN 46202
Parzych, David J., 756 Riverboat Circle, Orlando, FL 32828
Pasterkamp, Hans, Pediatrics Dept., Univ. of Manitoba, Room CN 503 820 Sherbrook Street, Winnipeg, MB R3A 1R9, Canada
Patat, Frederic J., Biphysique Medicale, Faculte de Medecine, 2 Bis Bvd. Tonnelle, BP 3223, Tours 37032, France
Patel, Samir B., Communications Business Unit, Hughes Space and Communications, 2000 East Imperial Highway, SC/S24/D520, El Segundo, CA 90019
Penney, David P., Pathology & Lab. Medicine, Univ. Rochester Medical Center, Box 626, 601 Elmwood Avenue, Rochester, NY 14642-0001
Peretz, Isabelle S. L., Psychology Dept., Univ. de Montreal, C.P. 6128, Succ. Centreville, Montreal, PQ H3C 3J7, Canada
Peterman, Karl L., Special Technologies Group, Newcomb & Boyd, One Northside 75, Atlanta, GA 30318

Petrykowski, John C., Mechanical & Aerospace Engineering, Univ. of Dayton, 300 College Park, Dayton, OH 45469-0210
Phelps, Andy D., Inst. of Sound and Vibration, Research, Fluids and Acoustics Group, Univ. of Southampton, Southampton, Hants. SO17 1BJ, U.K.
Pouliquen, Eric, FSG/ERD, SACLANTCEN, CMR 426, APO, AE 09613
Prausnitz, Mark R., School of Chemical Engineering, Georgia Inst. of Technology, Atlanta, GA 30332-0100
Raphael, David T., Texas Tech University Health, Science Center, Anesthesiology Dept., 4800 Alberta Avenue, El Paso, TX 79905
Recanzone, Gregg H., Center for Neuroscience, Univ. of California at Davis, 1544 Newton Court, Davis, CA 95616
Reller, Austin F., 90 Morningside Drive, Apt. 6J, New York, NY 10027
Rickert, Martin E., Dept. of Psychology, Indiana University, Bloomington, IN 47405
Rodgers, Robert A., 5949 Rose Avenue, Long Beach, CA 90805
Romero, Damion, 6132 Romaine Street, Los Angeles, CA 90038
Russell, William A., 703 Beards Hill, Aberdeen, MD 21001-1776
Salvi, Richard J., Hearing Research Lab., SUNY Univ. at Buffalo, 215 Parker Hall, Buffalo, NY 14214
Sander, Andres, Audiology Research, Starkey Laboratories GmbH, Rugenbarg 69, Norderstedt 22848, Germany
Sarrao, John L., Florida State Univ., NHMFL, 1800 East Paul Dirac Drive, Tallahassee, FL 32306
Sautermeister, Per, Spoken Language Research, Telia Research, 136-80 Haninge, Sweden
Sayeed, Mohammed, Bently Nevada Corporation, 7651 Airport Boulevard, Houston, TX 77061
Schonthal, Ernst, 6656 West Ridge Drive, Brighton, MI 48116
Schwartz, Mark, RR #1, Box 400, Effort, PA 18330
Sen, Dipanjan, Speech Processing Software and Technology Research, AT&T Bell Labs., 600 Mountain Avenue, Murray Hill, NJ 07974
Serpa, Jose A., N y E Omicron S.A. de C.V., Irapuato #14, Col. Condesa, Mexico DF C.P. 06170, Mexico
Sevener, Nathan B., Arup Acoustics, 2440 South Sepulveda Boulevard, Suite 180, Los Angeles, CA 90064
Shearer, Kenneth W., 9107 Lantern Lane, Indianapolis, IN 46256-2249
Sheffert, Sonya M., Psychology Dept., Indiana University, Bloomington, IN 47403
Shigeo, Hase, Theatre Design Co., Ltd., 1-16-1 Shibata, Kita-ku, Osaka-shi, Osaka, 530 Japan
Shigong, Ye, Inst. of Acoustics, Nanjing University, 22 Han Kou Road, Nanjing, Jiangsu 210093, P.R. China
Sjursen, Walter P., 6 Bankers Drive, Washington Crossing, PA 18977
Slaney, Malcolm, Internal Research Corp., 1801 Page Mill Road, Bldg. C, Palo Alto, CA 94304
Sola, Tony, Acoustical Consulting Services, P.O. Box 41182, Mesa, AZ 85274
Stepinski, Tadeusz, Uppsala University, P.O. Box 534, Uppsala 751 21, Sweden
Steurer, Martin, AKH-Wien, ENT (HNO), Waehringer Guertel 18-20, Vienna A-1030, Austria
Stratman, James A., 14706 Village Park West Drive, Carmel, IN 46033
Swarnamani, Seetharaman, Indian Inst. of Technology, Applied Mechanics, Madras, Tamlindadu 600-036, India
Szeri, Andrew J., Dept. of Mechanical & Aerospace Eng., Univ. of California, Irvine, Irvine, CA 92697-3975
Tavares, Rick, 10438 Rancho Carmel Drive, San Diego, CA 92128
Tavossi, Hassan M., Engineering Science & Mechanics, Pennsylvania State Univ., 227 Hammond Building, University Park, PA 16802-1401
Ter Haar, Gail R., 24 Spencer Hill, Wimbledon SW19 4NY, U.K.
Tidball, Glynnis A., School of Audiology & Speech Sciences, Univ. of British Columbia, 5804 Fairview Crescent, Vancouver, BC V6T 1Z3, Canada
Trout, J. D., Parmly Hearing Inst., Loyola Univ. of Chicago, 6525 North Sheridan Road, Chicago, IL 60626
Tsuru, Hideo, Nittobo Acoustic Engineering, 1-13-12 Midori, Sumidaku, Tokyo, 130 Japan
Ugolotti, Emanuele, R & D Dept., ASK S.p.A., Via 7 F.lli Cervi 79, Reggio Emilia, 42100 Italy
Uhlendorf, Volkmar, Ultrasound Contrast Media Research, Schering AG, Muellrstr. 178, Berlin 13342, Germany
Vakakis, Alexander F., Mechanical & Industrial Eng., Univ. of Illinois, 1206 West Green Street, Urbana, IL 61801
Van Lieshout, Pascal H. H. M., ENT, Voice & Speech Pathology, Academic Hospital Nijmegen St. Radboud, P.O. Box 9101, Nijmegen 6500HB, The Netherlands
Veen, Jerry R., Prince Corporation, Advanced Engineering, One Prince Center, Holland, MI 49423
Vega-Campos, Hernan, Sound and Acoustics School, Univ. Tecnologica Vicente, Perez Rosales, Brown Norte 290, Santiago, Nunoa, Chile
Verona, Enrico, CNR, Istituto Di Acustica, Via Cassia 1216, Roma I-00189, Italy
Vivyurka, Terry W., J. L. Richards & Associates Ltd., 864 Lady Ellen Place, Ottawa, ON K1Z 5M2, Canada
Wampler, Scott D., Technology Assessment & Development, Ethicon Edno-Surgery Inc., 4545 Creek Road, Cincinnati, OH 45242-2839
Wang, Shuozhong, Electronics & Information Eng., Shanghai University, 149 Yangchang Road, Shanghai 20072, P.R. China
Wang, Chao-Nan, Dept. of Naval Architecture and Ocean Engineering, National Taiwan Univ., College of Engineering, 73 Chow-Shan Road, Taipei, Taiwan, R.O.C.
Weber, Peter K., Neugrabenweg 18, Saarbrucken D-66123, Germany
Whistler, Peter, 12 Miller Street, Needham, MA 02192
Wilen, Larry A., Physics & Astronomy Dept., Ohio University, Athens, OH 45701
Winker, Douglas F., Research and Development, Dynamic Control of North America, Inc., 3042 Symmes Road, Hamilton, OH 45015
Woo, Hoon Young, Sanggye Paik Hospital, INJE Univ., Otolaryngology, Nowon ku Sanggye 7 Dong, Seoul 139-707, Korea
Writer, Eric R., Colin Gordon & Associates, 411 Borel Avenue, Suite 425, San Mateo, CA 94402
Yamashita, Mitsuyasu, Kobayasi Inst. of Physical Research, 3-20-41 Higashi-Motomachi, Kokubunji, Tokyo 185, Japan
Yan, Albert K. T., Room 819, Wah Shun House, Wah Fu Estate, Hong Kong
Yao, Weiping, Otolaryngology, Head and Neck Surgery, Medical College of Ohio, 3000 Arlington Avenue, Toledo, OH 43699-0008
Yasui, Kyuichi, Dept. of Physics, Waseda University, 3-4-1 Ohkubo, Shinjuku, Tokyo, 169 Japan
Ye, Shigong, Inst. of Acoustics, Nanjing University, 22 Han Kou Road, Nanjing, Jiangsu 210093, P.R. China
Yeh, Jyi-Tyan, Acoustics & Vibration Project, Material Research Labs., Industrial Technology Research Inst., Bldg. 77, 195 Chung Hsing Rd., Sec. 4, Chutug, Hsinchi 310, Taiwan
Zhang, Xiaorong, Dept. of Electronics Science & Eng., Inst. of Acoustics, Nanjing University, 22 Han Kou Road, Nanjing, Jiangsu 210093, P.R. China
Zhang, Chao Ying, House Ear Institute, 2100 West Third Street, Los Angeles, CA 90057
Zhdanov, Vladimir, 1581 East 24th Street, Apt. 5B, Brooklyn, NY 11229

New Students

Andalaft, Elias A., Casilla 214 Correo 30 Vitacura, Santiago, Chile
Andrews, Erik W., Engineering Dept., Brown University, Box D, Providence, RI 02912
Balkany, Jourdan T., 13061 Mar Street, Coral Gables, FL 33156-6427
Banerjea, Raja, 35 Beaumont Court, Tinton Falls, NJ 07724
Bangayan, Philbert T., 17417 Tuscan Drive, Granada Hills, CA 91344
Barney-Tomalin Joy L., 3725 Federal Avenue, Everett, WA 98201
Berndt, Tobias P., Materials Science & Engineering, The Johns Hopkins University, 3400 North Charles Street, Maryland Hall 102, Baltimore, MD 21218
Bezemek, Jacklyn D., 132C Palmer, #111, West Lafayette, IN 47906
Bland, Angela R., P.O. Box 5156, Elko, NV 89802
Bunton, Kate E., Dept. of Communicative Disorders, Univ. of Wisconsin Madison, 491 Waisman Center, 1500 Highland Avenue, Madison, WI 53705
Burnett, Teri A., 632 Sheridan Square, #2, Evanston, IL 60202
Chan, Tin Man, ISVR, Univ. of Southampton, Highfield, Southampton SO17 1BJ, U.K.
Cherukuri, Aravind, 25 Aberdeen Street, #2F, Boston, MA 02215
Dainora, Audra E., 5532 South Kenwood Avenue, #411, Chicago, IL 60637
Dankovicova, Jana, Phonetics Lab., Univ. of Oxford, 41 Wellington Square, Oxford OX1 2JF, U.K.
Deffenbaugh, Max, 127 Palmer Avenue, Apt. 1, Falmouth, MA 02540
Degertekin, Fahrettin L., Electrical Engineering Dept., Stanford University, E. L. Ginzton Lab., Stanford, CA 94305-4085
Dodson, John M., 40646 Rock Hill, Novi, MI 48375

Dolan, Daniel H., Physics Dept., Univ. of Utah, 201 JFB, Salt Lake City, UT 84112
Estaphan, Michelle K., 13 Brinley Street, #2, Newport, RI 02840
Evers, Penny J., 801 Brookmere Drive, Edmonds, WA 98020
Field, Christopher D., Architectural & Design Science, Univ. of Sydney, Wilkinson Building 904, Sydney, NSW 2006, Australia
Fortin, Martin, 6225 Place Northcrest, #L-19, Montreal PQ H3S 2T5 Canada
Fozo, Michael S., 20 North Broadway, #N165, White Plains, NY 10601
France, Daniel J., Biomedical Engineering, Vanderbilt University, Box 1671, Station B, Nashville, TN 37235
Garbe, Colleen M., Kresge Hearing Research Inst., Otolaryngology Dept., 1301 East Ann Street, Ann Arbor, MI 48109-0506
Gaygen, Daniel E., 2571 Youngstown-Lockport Road, #146, Ransomville, NY 14131
Groenenboom, Jeroen, Section of Applied Geophysics, Delft Univ. of Technology, Faculty of Applied Earth Sciences, Mijnbouwstraat 120, Delft 2628RX, The Netherlands
Gupta, Asit K., Applied Geophysics, Indian School of Mines Dhanbad, Dhanbad, Bihar 826004, India
Halberstam, Benjamin, 1577 East 17th Street, Apt. 5J, Brooklyn, NY 11230-6747
Hampson, Michelle, 115 Nottinghill Road, Brighton, MA 02135
Hassan, Waled T., Aerospace Engineering and Engineering Mechanics, Univ. of Cincinnati, ML #0070, Cincinnati, OH 45221
Hemphill, Rachel M., Linguistics Dept., Univ. of Chicago, 1010 East 59th Street, Chicago, IL 60637
Henderson, Paul D., 15817 Thompson Road, Charlotte, NC 28227
Hirata, Yukari, 1605 East 50th Street, #12D, Chicago, IL 60615
Howard, Danny D., Mail Code 205-45 CALTECH, Pasadena, CA 91125
Huff, Jr., John E., 3218 New London Court, #105, Lafayette, IN 47905
Hunter, Eric J., 695 North 800 East, Provo, UT 84606
Husain, Fatima T., Dept. of Cognitive & Neural Systems, Boston University, 677 Beacon Street, Boston, MA 02215
Isaac, Michael C., Turk Hill Road, Brewster, NY 10509
Jones, Christopher D., 1422 33rd Avenue, Seattle, WA 98122
Jung, Yi Bong, 546 Hyo Jung 1, Hwa San Myun, Young chun, Kyung pook 771-850, Korea
Kannan, Srilatha K., 1111 East University Drive, #224, Tempe, AZ 85281
Khosla, Sunny R., 1736 McIntyre Drive, Ann Arbor, MI 48105
Kibre, Nicholas, 401 West Pine, #31, Lompos, CA 93436
Kim, Hyun Chul, Sensor Technology Research Center, Kyungpook National Univ., 1370 San Kyuk dong, Taegu 702-701, Korea
Kim, Young-Key, Mechanical Engineering, Korea Advanced Inst. of Science and Technology, Yusong-gu guson-dong 373-1, Taejon 305-701, Korea
Kleinknecht, Sabina A., 22 Quarry Path, Stonington, CT 06378
Krause, Jean C., Massachusetts Inst. of Technology, Research Lab. of Electronics, Rom 36-747, Cambridge, MA 02139
Kuo, Hong-Kwang J., Speech Communication Lab., MIT Research Lab. of Electronics, 36-511, 50 Vassar Street, Cambridge, MA 02139
Kwon, Bomjun, Speech Pathology & Audiology, Univ. of Iowa, Wendell Johnson Speech & Hearing Center, Iowa City, IA 52242-1012
Lane, John E., 400 Wilson Avenue, Satellite Beach, FL 32937
Lee, Young Jin, Sensor Engineering, Kyungpook National Univ., 1370 Sankyuk Dong Pukgu, Taegu 702-701, Korea
Lee, June-Yule, Inst. of Sound and Vibration Research, Southampton University, Signal Processing & Control Group, Southampton SO17 1BJ, England, U.K.
Lentz, Jennifer, Bioengineering Dept., Univ. of Pennsylvania, 120 Hayden Hall, Philadelphia, PA 19104-6392
Logan, Thomas M., 2829 Weston Brook Lane, Duluth, GA 30136
McCullough, Earl T., 3699 Lockport-Olcott Road, Lockport, NY 14094
Mercado, Eduardo, III, Psychology Dept., Kewalo Basin Marine Mammal Lab., 1129 Ala Moana Blvd., Honolulu, HI 96814
Meyouhas, Sela M., 2456 Massachusetts Avenue, #201, Cambridge, MA 02140
Moidu, Abdulaleel K., Mechanical & Industrial Eng., Univ. of Toronto, 5 King's College Road, Toronto, ON M5S 1A4, Canada
Monahan, Peter, 21 Spireview Lane, Rathgar Road, Dublin 6, Ireland
Montgomery, Joshua M., 25 Aberdeen Street, #2F, Boston, MA 02215
Murray, Scott O., 1519 One'ele Place, Honolulu, HI 96822
Nadeau, Richard C., 117 Wildflower Drive, Cranston, RI 02921
Nadel, Rebecca E., Box 99250, Durham, NC 27403
Nathwani, Kishore M., Applied Physics Lab., Ocean Physics Dept., 1013 NE 40th Street, Seattle, WA 98105
O'Neill, James S., 731 South Plymouth Court, Apt. 611, Chicago, IL 60605
Olivieri, Marc P., 301 SW 1st Street, #D112, Boca Raton, FL 33432
Olson, Kenneth S., Linguistics Dept., Univ. of Chicago, Classics Bldg., 1010 East 59th Street, Chicago, IL 60637
Ostroff, Jodi M., 120 West 44th Street, New York, NY 10036
Patka, Robert R., Parmly Hearing Inst., Loyola Univ. of Chicago, 6525 North Sheridan Road, Chicago, IL 60626
Pedersen, Michael, Beatrixstraat 150, Enschede NL-7511 KV, The Netherlands
Pedersen, Torstein K., 1530 Cambridge Street, Apt. 3, Cambridge, MA 02139
Poplawski, James E., 3624 Partridge Path, #2, Ann Arbor, MI 48108
Rabinkin, Daniel V., Electrical Engineering, CAIP Center, Core Bldg., Rutgers University, Piscataway, NJ 08855
Redford, Melissa A., Psychology Dept., Univ. of Texas at Austin, 330 Mezes, Austin, TX 78712
Rieche, Bernd, Rudolf-Haym Str. 5, Halle D-06110, Germany
Roberts, Richard A., 154-C Norton Place, Mobile, AL 36607
Roup, Christina M., Communicative Disorders, Audiology, Univ. of Wisconsin—Madison, 1975 Willow Drive, Madison, WI 53706
Ryerson, Erik J., 17 Sears Place, Wayne, NJ 07470
Safavi-Ardebili, Vahid, 35 Charles Street, West, #2012, Toronto, ON M4Y 1R6, Canada
Salindong, Arthur R., 251 Ardennes Circle, Seaside, CA 93955
Sampath, Arunkumar M., Mechanical Engineering Dept., Univ. of Maryland, College Park, MD 20742
Samuels, Timothy O., Physics and Astronomy, Brigham Young University, 263 Fletcher Building, P.O. Box 24360, Provo, UT 84602
Schrade, II, William A., 173 Warren Avenue, Plymouth, MA 02360
Scott, Mark I., P.O. Box 7685, Stanford, CA 94309-7685
Shaffer, Thomas R., Psychology Dept. 296, Univ. of Nevada, Reno, NV 89557
Shockley, Wesley J., Dept. of Biology, Univ. of South Florida, 422 East Fowler Avenue, Tampa, FL 33620
Singh, Shajila A., 800 Hinman Avenue, #413, Evanston, IL 60202
Smith, Sandra A., 101 Davis Street, Apt. D-8, Athens, GA 30605
Smith, Gordon P., 1802 West Jackson Avenue, #96, Oxford, MS 38655
Stenfelt, Stefan P. Y., Dept. of Applied Electronics, Chalmers Univ. of Technology, Horsalsvagen 11, Goteborg 412 96, Sweden
Swiergosz, Matthew J., Psychology Dept., Univ. of Nevada, Reno, NV 89557
Tantum, Stacy L., 1319 Norton Street, Durham, NC 27701
Tasko, Stephen M., Communication Disorders, Univ. of Wisconsin—Madison, Waisman Center, 1500 Highland Avenue, Rm. 493, Madison, WI 53705-2280
Terrill, Eric J., Marine Physical Lab., Scripps Inst. of Oceanography, 8605 La Jolla Shores Drive, Mail Stop 0230, La Jolla, CA 92093-0230
Thongprasirt, Rachod, P. O. Box 522230, Univ. of Memphis, Memphis, TN 38152
Tuthill, John D., ATRI, Curtin Univ. of Technology, GPO Box U1987, Perth, WA 6001, Australia
Uppal, Samir, Aerospace Engineering & Engineering Mechanics, Iowa State University, 2019 Black Engineering Bldg., Ames, IA 50011
Vine, Daniel S. G., 254 Malmesbury Park Road, Bournemouth, Dorset BH8 8PR, England
von Berger, Stephanie M., Speech and Hearing Science, Arizona State Univ., Box 871908, Tempe, Az 85287
Walker, V. Ashley, Dept. of Artificial Intelligence, Univ. of Edinburgh, 5 Forrest Hill, Edinburgh EH1 2QL, Scotland, U.K.
Warner, Natasha L., 3331 Belmont Avenue, El Cerrito, CA 94530
Whitman, Teresa A., Agricultural & Biological Eng., Purdue University, West Lafayette, IN 47907
Withnell, Robert H., Physiology Dept., Univ. of Western Australia, Mounts By Road, Nedlands, Perth WA 6907, Australia
Wong, Willy, Inst. of Biomedical Engineering, Univ. of Toronto, 4 Taddle Creek Road, Toronto, ON M5S 3G9, Canada
Yasutaka, Ueda, Graduate School of Science and Engineering, Kobe University, 1-1 Rokkodai Nadaku, Kobe, Hyogo, 657 Japan

Associates to Members

J. C. Adamowski, G. Alker, M.J. Beam, J. J. L. Beard, S. S. Black, D. J. Blasinsky, S. L. Broschat, T. A. Brungart, S. A. Chapman, L. S. Deliman, E. F. Elkins, L. Ercoli, M. Fink, J. A. Fitzpatrick, S. N. Gurbatov, R. H.

Gutierrez, C. M. Hedberg, K. Heutschi, D. J. Hill, D. A. James, O. Jiricek, J. P. Kelly, H-G. Kim, D. W. Lee, P. Y. Lee, C. P. Lichtenwalner, Q-H. Liu, R. W. Lungstrum, B. Moebius, J. H. Mullins, Stephen J. O'Leary, V. G. Pimshtein, O. Soiji, J. R. Smith, Jr., C. Soize, F. E. Stanke, T. L. Szabo, O. T. O. Toru, T-C. Yang

Students to Associates

K. A. Burgemeister, V. Dutt, R. F. Fehlmann, D. H. Kruger, A. J. Oxenham, S-H. Peng, R. V. N. Rao, J. A. A. Rott, H. Yang

Reinstated

B. Cai, D. R. Jones, J. H. Huckans—*Members*

Deceased

M. Heckl, H. L. Oestreicher, P. S. Veneklasen—*Fellows*
C. L. Badley, R. Beresford, A. E. Bisson, M. L. Harbold, L. G. Ramer—*Members*

Fellows	785
Members	2779
Associates	2712
Students	881
	7157

REVIEWS OF ACOUSTICAL PATENTS

Daniel W. Martin
7349 Clough Pike, Cincinnati, Ohio 45244

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

Reviewers for this issue:

GEORGE L. AUGSPURGER, *Perception Incorporated, Box 39536, Los Angeles, California 90039*
SAMUEL F. LYBARGER, *101 Oakwood Road, McMurray, Pennsylvania 15317*
D. LLOYD RICE, *11222 Flatiron Drive, Lafayette, Colorado 80026*
CARL J. ROSENBERG, *Acentech Incorporated, 125 Cambridge Park Drive, Cambridge, Massachusetts 02140*

5,514,841

43.38.Ja REFLEX COMPRESSION VALVE-DIVIDED CHAMBER LOUDSPEAKER CABINET

Donald C. Rochon, Surry, BC, Canada
7 May 1996 (Class 181/156); filed 30 March 1994

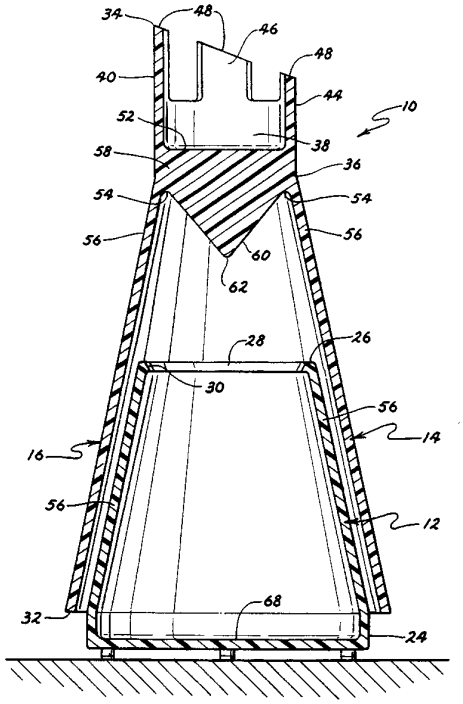
First there is the matter of the patent title. It is probably meant to scan as: reflex compression-valve, divided-chamber, loudspeaker cabinet. Second, one may well ask why the world needs yet another vented box, and the patent text modestly provides the answer: ''...the primary object of the invention is to advance the art of speaker cabinet design in a radical, pioneering way.'' A total of four vents, a compression batting valve, and a free-flow air pass slot are all that are needed to reach this happy state of affairs.—GLA

5,525,767

43.38.Ja HIGH-PERFORMANCE SOUND IMAGING SYSTEM

Walter Fields, Minneapolis MN
11 June 1996 (Class 181/155); filed 22 April 1994

Coaxial mid- and high-frequency loudspeakers **20** and **22** are mounted in their own chamber atop base unit **12**. Woofer **18** drives its upper chamber which is then coupled to the room through the vent formed by the gap between skirts **12** and **14**. If the drawing is at all to scale then the low-frequency section functions as a single-tuned bandpass system and ''acoustic lens'' **60** serves no useful purpose.—GLA



5,416,847

43.38.Lc MULTI-BAND, DIGITAL AUDIO NOISE FILTER

Steven E. Boze, assignor to The Walt Disney Company
16 May 1995 (Class 381/94); filed 12 February 1993

The patent describes an operator-controlled, multiband, digital sound track scrubber that appears to be both practical and useful. Appendix A is the complete source code listing, using a Motorola XSP56001RC33 microprocessor.—GLA

5,440,639

43.38.Lc SOUND LOCALIZATION CONTROL APPARATUS

Yasutake Suzuki and Junichi Fujimori, assignors to Yamaha Corporation
8 August 1995 (Class 381/17); filed in Japan 14 October 1992

Monophonic signals from a synthesizer or the like can be electronically panned with regard to angular location and distance. As in most Yamaha patents, the apparatus is clearly described in the text and numerous illustrations are included.—GLA

5,442,712

43.38.Lc SOUND AMPLIFYING APPARATUS WITH AUTOMATIC HOWL-SUPPRESSING FUNCTION

Akihisa Kawamura *et al.*, assignors to Matsushita Electric Industrial Company
15 August 1995 (Class 381/83); filed in Japan 25 November 1992

Digital processor chips make it possible to perform complicated audio analysis and control at relatively low cost. So, why can't there be a public address system that doesn't howl? A number of patents are devoted to this goal. A major problem is identifying the onset of howl in the first place; a sound system that repeatedly sings and then corrects itself is not acceptable to most listeners. The patent at hand clearly explains an interesting approach to howl detection. Unfortunately, the goal of howl prediction remains elusive.—GLA

5,467,393

43.38.Lc METHOD AND APPARATUS FOR VOLUME AND INTELLIGIBILITY CONTROL FOR A LOUDSPEAKER

Jim A. J. Rasmusson, assignor to Ericsson, Incorporated
14 November 1995 (Class 379/388); filed 24 November 1993

High-power sound amplification systems often include proprietary signal processing that combines compression with frequency response contouring to get the greatest possible sound output without overdriving the loudspeakers. In this patent a similar approach is taken to get maximum intelligibility from a small, low-power loudspeaker such as that used in a cellular speakerphone. The arrangement is well thought out and clearly described in the patent document.—GLA

5,487,113

43.38.Lc METHOD AND APPARATUS FOR GENERATING AUDIOSPATIAL EFFECTS

Steven D. Mark and David Doleshal, assignors to Spheric Audio Laboratories, Incorporated
23 January 1996 (Class 381/17); filed 12 November 1993

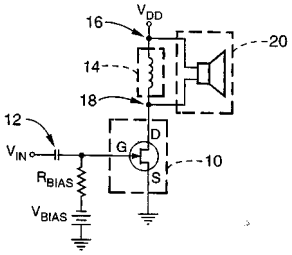
Frequency-contoured background noise (a gray noise template) is added to an audio signal for the purpose of inhibiting pinna-related spatial cues. When a spatially localized effect is desired, this disorienting template is modified or augmented by ''reorienting'' signal processing. Since interaural differences are not exploited, such effects can be perceived equally well by all members of a large audience, presumably even those who may be deaf in one ear.—GLA

5,498,997

43.38.Lc TRANSFORMERLESS AUDIO AMPLIFIER

Cristopher F. Schiebold, Palo Alto, CA
12 March 1996 (Class 330/277); filed 23 December 1994

A field effect output transistor **10** is combined with inductor **14** to match the low impedance of loudspeaker **20** yet provide distortion and



overload characteristics typical of vacuum tube circuitry.—GLA

5,515,446

43.38.Lc ELECTRONIC AUDIO ACCURATE REPRODUCTION SYSTEM AND METHOD

George Velmer, Studio City, CA
7 May 1996 (Class 381/98); filed 22 August 1994

To achieve natural sounding audio compression, the signal may be split into two or three bands which are processed independently and then recombined. Well if three is good, then ''a plurality'' (eight maybe?) must be even better. The patent document teaches that any audio signal will somehow be reproduced more accurately if its tonal balance is thus dynamically flattened.—GLA

5,530,761

43.38.Lc AUTOMATIC PROCESS OF ADJUSTMENT OF THE VOLUME OF SOUND REPRODUCTION

Stephane M. d'Alayer de Costemore d'Arc, assignor to Staar S.A.
25 June 1996 (Class 381/57); filed in Belgium 30 June 1993

By now there must be at least two dozen patents dealing with automatic adjustment of sound level in relation to background noise. What has been missing is the inclusion of a computer using fuzzy logic. By making use of the simple formula: $S = \mathrm{MICRO} - aLS - S(OM)$ the inventors have advanced the state of the art in such a way that, ''...the adjustment faithfully follows the level of the disturbing noise and the user hears no sudden variation in level.'' The patent document is clearly written and presents interesting information.—GLA

5,524,058

43.38.Si APPARATUS FOR PERFORMING NOISE CANCELLATION IN TELEPHONIC DEVICES AND HEADWEAR

William T. Moseley, assignor to MNC, Incorporated
4 June 1996 (Class 381/71); filed 12 January 1994

The patent shows a handset with noise canceling circuitry for the receiver. Oppositely phased microphones are placed in the receiver, one facing the cavity behind the diaphragm and the other facing toward the ear. Air vents are provided in the cavity behind the diaphragm. The circuitry provided adds the outputs of the microphones to give significant cancellation of ambient noise.—SFL

5,533,122

43.38.Si ARTICULATED HEADSET SUPPORT

Phillip A. Gattey *et al.*, assignors to ACS Wireless, Incorporated
2 July 1996 (Class 379/430); filed 26 April 1995

The patent shows an over-the-ear headset that includes two curved housings that rotate with respect to each other. One housing supports a microphone boom; the other a receiver element. An elastomer hinge between the housings reduces feedback.—SFL

5,545,859

43.38.Si ANTI-VIRAL ACOUSTICALLY TRANSPARENT EARPHONE COVER

Kenneth A. Ullrich, Wenatchee, WA
13 August 1996 (Class 181/129); filed 7 March 1994

A cover of thin (e.g., 1 mil) substantially antiviral material such as polyethylene is to be placed between an audiometer earphone and the ear being tested. The thin material used does not alter the acoustical performance of the earphone. A circular elastic band fastened to the cover to hold it in place may be used.—SFL

5,493,697

43.38.Tj COMMUNICATIONS SYSTEM FOR THE GAME OF FOOTBALL INCLUDING PLAYER-CARRIED TRANSMITTER AND SIDE LINES SPEAKERS FOR OVERCOMING SPECTATOR NOISE

Randall L. May, Huntington Beach, CA
20 February 1996 (Class 455/66); filed 27 August 1993

The quarterback's helmet includes a small microphone and transmitter. If the spectators make too much noise for his signals to be heard, he simply presses a talk switch to energize powerful loudspeakers along the side lines.

According to the patent document, the system has been tested successfully.—GLA

5,497,425

43.38.Vk MULTI CHANNEL SURROUND SIMULATION DEVICE

Robert J. Rapoport, St. Petersburg, FL
5 March 1996 (Class 381/18); filed 7 March 1994

A resistive matrix apportions suitable mixtures of left and right loudspeaker-level signals to a total of five loudspeakers and/or additional power amplifiers. Circuitry is inserted in plus (high) and minus (common) signal paths, which should lead to interesting results with single-ended amplifiers.—GLA

5,521,982

43.38.Vk ELECTRONIC DEVICE FOR THE GENERATION OF ACOUSTIC SPATIAL EFFECTS

Yair Schiftan, Effretikon, Switzerland
28 May 1996 (Class 381/25); filed in Switzerland 19 July 1993

The system takes left and right signals from a stereophonic source, then distributes them to a number of left and right outputs with independent processing added to each channel. One channel feeds stereo headphones and the remaining channels drive vibrators. ''The bone vibrator transducers may be placed on the head or on other parts of the body... to generate various acoustic and vibrational effects.''—GLA

5,533,129

43.38.Vk MULTI-DIMENSIONAL SOUND REPRODUCTION SYSTEM

Herbert I. Gefvert, Lake Forest, IL
2 July 1996 (Class 381/24); filed 24 August 1994

This home theatre loudspeaker system includes front-left and front-right ambience loudspeakers, which are mounted in the same enclosure as the main center loudspeaker. Signals for the two additional loudspeakers are derived by combining main left plus ambience left, and main right plus ambient right.—GLA

5,517,173

43.50.Ed APPARATUS GENERATING NOISE SOUND FOR ELECTRIC CAR

Young-Whan Cha and Hyun-Soo Ahan, assignors to Samsung Heavy Industries Company
14 May 1996 (Class 340/404.1); filed in Republic of Korea 30 June 1993

Electric automobiles and lift trucks are almost noiseless, making the world a more dangerous place for unwary pedestrians. An air scoop, a tuned pipe, and an adjustable throttle are combined to provide a possible remedy.—GLA

5,504,282

43.50.Gf SOUND TRANSMISSION AND ABSORPTION CONTROL MEDIA

Joseph F. Pizzirusso and John R. Rucher, assignors to Foamex L. P.
2 April 1996 (Class 181/290); filed 24 August 1994

Traditional sound transmission barrier systems (as used in the passenger cabin of motor vehicles and boats) use a high-mass intermediate layer. This patent describes a system that replaces the single intermediate layer with two or more layers that are separated by a strong yet light material, such as polyurethane foam or rebonded polyurethane foam. The total mass per unit area of the system consisting of the two separated layers and the foam often can be less than that of the traditional intermediate layer, without sacrificing any of the effectiveness of the system as a sound transmission barrier.—CJR

5,505,344

43.55.Ev ACOUSTIC CEILING PATCH SPRAY

John R. Woods, assignor to Spraytex, Incorporated
9 April 1996 (Class 222/394), filed 30 November 1994

This spray is intended to provide an inexpensive and practical way to match the surface texture of an acoustical ceiling. The spray patch material is stored under pressure, and is dispensed from a hand-held pressurized container for direct application as a liquid or semiliquid foam onto a drywall supporting ceiling. The textured surface of this patch will then match the surrounding acoustical tile ceiling visually and mechanically.—CJR

5,512,715

43.55.Ev SOUND ABSORBER

Hiroyuki Takewa and Yutaka Torii, assignors to Matsushita Electric Industrial Company
30 April 1996 (Class 181/295); filed in Japan 15 June 1993

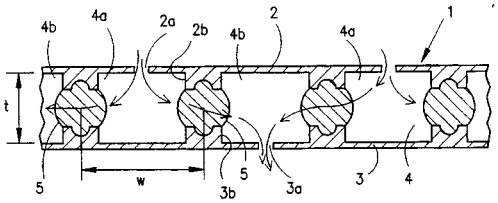
The patent states that the sound absorber has a constant coefficient of absorption over a wide frequency spectrum. To do this, the absorber has a laminated structure with layers of porous materials and high-polymer films. The laminations are in parallel to the incidence plane of sound, and are of increasing thickness away from the sound.—CJR

5,532,440

43.55.Ev LIGHT TRANSMISSIVE SOUND ABSORBING MEMBER

Kyoji Fujiwara, assignor to Nitto Boscki Company
2 July 1996 (Class 181/289); filed 7 December 1994

Light transmissive sheets are made into a sound-absorbing panel by spacing two sheets apart and then introducing holes in the outer sheets. The



holes are not connected through directly but instead introduce sound pressure into smaller interstices, which absorb sound.—CJR

5,545,861

43.55.Ev MEMBRANOUS-VIBRATION SOUND ABSORBING MATERIALS

Toro Morimoto, assignor to Toru Morimoto and Unix Corporation
13 August 1996 (Class 181/290); filed in Japan 13 March 1992

A transparent resin membrane film is sandwiched between perforated metal or expanded metal plates (or other sheets with a large number of openings of different sizes), and the resulting combination exhibits sound-absorbing properties. The sound-absorbing properties of the thin film are determined by its tensile strength, thickness, and tension.—CJR

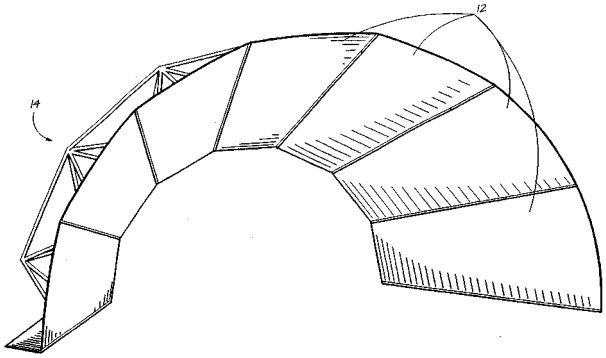
5,525,766

43.55.Fw PORTABLE ACOUSTICAL SHELL STRUCTURE

Richard R. Atcheson *et al.*, assignors to R. & A. Acoustical Structures

11 June 1996 (Class 181/30); filed 23 November 1994

This portable acoustical shell comprises sound-reflecting panels joined together in a modular frame to create an open, arch-like configuration. The shell is suitable for either indoor or outdoor use, irrespective of irregularities in the ground surface. The size or arch diameter may be increased or decreased by adding or removing panels and their respective frame modules. The shell would use readily available inexpensive materials such as standard



plastic sheet material for the panels, and light-metal conduit or tubing as the modular frame assembly. No special tooling would be required for fabrication; and the shell could be quickly assembled and disassembled.—CJR

5,530,211

43.55.Fw SOUND REFLECTING SHELL TOWER AND TRANSPORTER STRUCTURE AND METHODS OF ERECTING AND STORING THE TOWERS

Orley D. Rogers, James F. Jenne, and Phillip R. Blaisdell

25 June 1996 (Class 181/30); filed 27 December 1994

The patent describes an improved shell tower panel structure which is readily assembled in modules to form an optimal acoustical enclosure, which uniquely nests very compactly in a stored position, and which can be moved rapidly and easily between assembly and storage. The tower has a central panel, hinged wing panels, a counter-weighted base frame, and elements for connecting to a stage.—CJR

5,504,281

43.55.Rg PERFORATED ACOUSTICAL ATTENUATORS

Leland R. Whitney *et al.*, assignors to Minnesota Mining and Manufacturing Company

2 April 1996 (Class 181/286); filed 21 January 1994

The acoustical attenuator (barrier material) uses a porous material which is made up of tiny particles bonded or sintered together. The porous material has passages that are open through the material, but it still acts as a barrier because of the interstitial porosity (about 20% to 60%), average pore diameter (about 5 to about 280 μm), density (about 5 to about 60 lbs. per cubic foot), and so forth. The patent states that these properties make the material useful for appliances, sound equipment, etc. The voids between particles that form the porosity through the material are so small, they have to be measured with an electron microscope.—CJR

5,509,247

43.55.Wk VIBRATION-DAMPING INSIDE ROOF CONSTRUCTION

Maurice Fortez and Thorsten Alts, assignors to Matec Holding AG, Switzerland

23 April 1996 (Class 52/630); filed in Switzerland 23 September 1992

This patent presents a product that is designed to be used in all kinds of vehicles to provide an inside-roof construction that effectively damps the vibrations in the range below 200 Hz without much increase in weight, and at the same time retains completely its sound-absorbing efficiency. The approach incorporates an inside-roof construction comprising at least one member and a sound-absorbing moulded part. The moulded part is made of a porous material, preferably a compressed fibrous web. In order to dampen the low-frequency vibrations of the body member, the side of the moulded part that faces the body member has an anisotropic air resistant structure, and has a stiffening layer to make it self supporting.—CJR

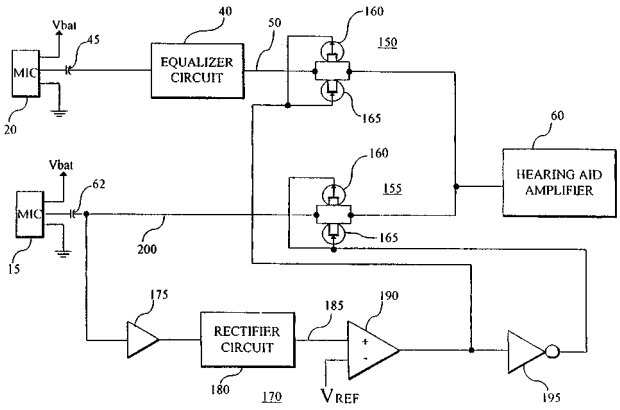
5,524,056

43.66.Ts HEARING AID HAVING PLURAL MICROPHONES AND A MICROPHONE SWITCHING SYSTEM

Mead Killion *et al.*, assignors to Etymotic Research, Incorporated

4 June 1996 (Class 381/68.2); filed 13 April 1993

An in-the-ear hearing aid having an omnidirectional microphone and at least one directional microphone of the first order is described. The purpose is to improve the signal-to-noise ratio for a hearing-impaired person by utilizing the directionality characteristics of the directional microphone. A



switching circuit accepts the signals from both microphones and connects the signal from the omnidirectional microphone to an input of a hearing aid amplifier automatically in response to sensed ambient noise levels.—SFL

5,524,150

43.66.Ts HEARING AID PROVIDING AN INFORMATION OUTPUT SIGNAL UPON SELECTION OF AN ELECTRONICALLY SET TRANSMISSION PARAMETER

Joseph Sauer, assignor to Siemens Audiologische Teknik GmbH

4 June 1996 (Class 381/68); filed in Germany 27 February 1992

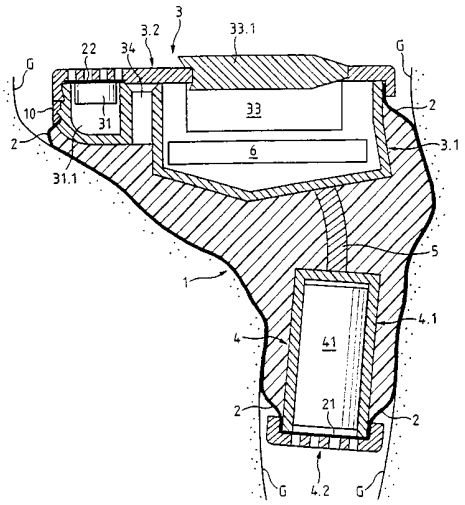
The patent describes a behind-the-ear hearing aid that has a number of adjustment possibilities. A single-easily operated switch is used to step sequentially from one characteristic to another. An audible signal is generated for each step, which tells the user, who cannot see the aid, which characteristic is being connected. The switch is operated until the characteristic desired by the user is provided.—SFL

5,530,763

43.66.Ts HEARING AID TO BE WORN IN THE EAR AND METHOD FOR ITS MANUFACTURE

Walter Aebi *et al.*, assignors to Ascom Audiosys AG
25 June 1996 (Class 381/69); filed in Switzerland 11 June 1993

The patent shows an in-the-ear hearing aid with the outer surface shaped by the individual ear canal. The aid has a skeleton which can be plastically deformed to a limited extent, and that carries the electronic components to which is fixed a flexible conical diaphragm. This assembly is



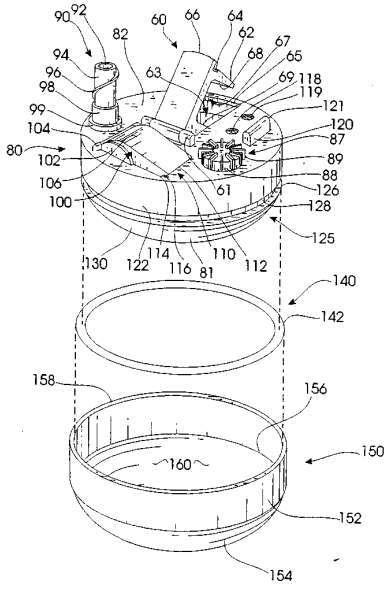
placed into the actual ear to be fitted and filled with a hardening plastic mass. The outer shape of the aid will then conform closely to that of the ear canal in which the aid is to be used.—SFL

5,533,130

43.66.Ts COSMETICALLY DISGUISED HEARING AID

Ed Staton, Hot Springs, AR
2 July 1996 (Class 381/68.5); filed 15 August 1994

A hearing aid is designed to be attached to eyeglasses or suspended from an earmold. The visible outer shell of the aid is hemispherically shaped



for cosmetic appearance. The volume control knob, battery door and sound outlet are on the flat side of the aid, facing the head.—SFL

5,535,282

43.66.Ts IN-THE-EAR HEARING AID

Racca Luca, assignor to Ermes S.r.l.
9 July 1996 (Class 381/68.6); filed in Italy 27 May 1994

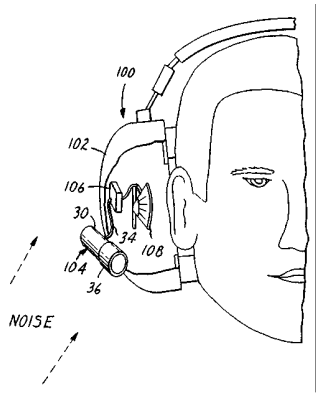
The patent shows an in-the-ear or in-the-canal hearing aid. The aid is provided with a vent tube that goes from its inner end to the outer face and that maintains an open connection between the ear canal and the outer face of the aid. The output of the hearing aid earphone is supplied to the vent at about its center. The arrangement is stated to protect the earphone from the effects of cerumen and other ear canal secretions. A means for closing the outer end of the vent tube when desirable is shown.—SFL

5,550,923

43.66.Vt DIRECTIONAL EAR DEVICE WITH ADAPTIVE BANDWIDTH AND GAIN CONTROL

David A. Hotvet, assignor to Minnesota Mining and Manufacturing Company
27 August 1996 (Class 381/72); filed 2 September 1994

This patent describes an ear protective device **100** (e.g., a cushioned earphone) that is equipped with a directional microphone **104**, an adaptive bandpass filter **106**, and an output transducer **108**. The purpose of the adaptive filter is to narrow the frequency range of the sound signal transmitted from the microphone to the output transducer, depending upon the level and



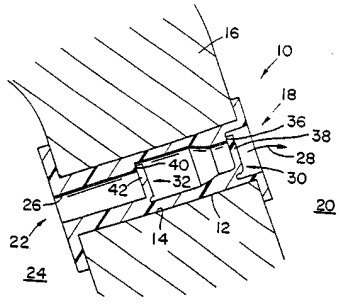
spectrum of the ambient noise exterior to the device. There is considerable emphasis on the role of the directional loudspeaker which the user orients toward the desired sound (i.e., a talker), and away from undesired ambient noise sources.—DWM

5,507,809

43.70.Dn MULTI-VALVED VOICE PROSTHESIS

Eric D. Blom, assignor to Hansa Medical Products, Incorporated
16 April 1996 (Class 623/9); filed 5 November 1993

This prosthesis is intended as an improvement for several problems related to the production of esophageal speech by patients with various



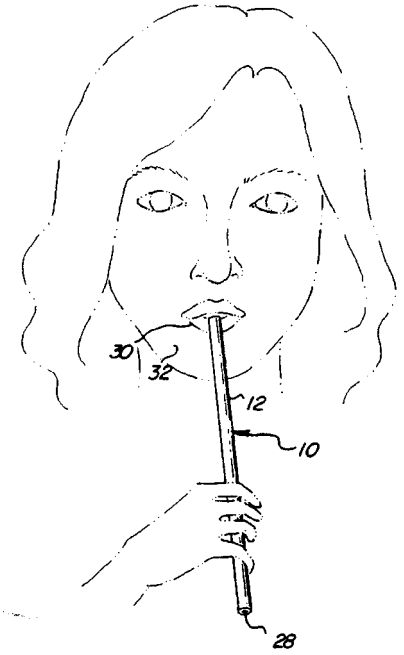
laryngeal disorders. The design allows a one-way flow of air from esophagus to trachea and blocks the flow of fluids with a double-check valve arrangement.—DLR

5,507,648

43.70.Dn HOLLOW TUBULAR DEVICE TO ASSIST PATIENTS DURING SPEECH THERAPY

Jonathan D. Knopf, Salina, KS
16 April 1996 (Class 434/185); filed 3 August 1994

This patent describes a hollow tube with capped ends having a small hole in each end cap. It would be held against the lower lip, as seen in the figure, to allow a person in speech therapy to have a better way of sensing the airflow from the lips when certain sounds are produced. The patent



implies, without justification, that the tube might be helpful somehow in sensing the vibrations from turbulence produced in the sibilant sounds.—DLR

5,504,834

43.72.Ar PITCH EPOCH SYNCHRONOUS LINEAR PREDICTIVE CODING VOCODER AND METHOD

Bruce A. Fette *et al.*, assignors to Motorola, Incorporated
2 April 1996 (Class 395/2.16); filed 28 May 1993

This variation on a linear prediction speech analyzer uses different techniques to analyze the voiced and unvoiced portions of the signal. An autocorrelation of the input forms the basis for pitch epoch extraction. Individual pitch intervals are then passed to a linear prediction analysis, resulting in a stable spectral estimate, avoiding glottal phase effects. An excitation codebook search is employed to determine a suitable representation for unvoiced intervals.—DLR

5,509,103

43.72.Bs METHOD OF TRAINING NEURAL NETWORKS USED FOR SPEECH RECOGNITION

Shay-Ping T. Wang, assignor to Motorola, Incorporated
16 April 1996 (Class 395/2.41); filed 3 June 1994

Presented as a speech recognizer based on neural network principles, the system described here seems to omit the usual hidden layer, and so essentially uses weighted sums of various powers of the input layer signals. Preprocessing consists of a customary speech analyzer, producing successive frames of filtered linear prediction cepstral coefficients. The sequence of frame vectors is then decimated, sending every fourth vector to each of four parallel network structures. No reason is given for the decimation. The coefficients are then squared, cubed, and weighted by an unspecified number of parallel adders to produce phoneme-specific output classification signals. The description of the training procedure is particularly unenlightening.—DLR

5,504,832

43.72.Gy REDUCTION OF PHASE INFORMATION IN CODING OF SPEECH

Tetsu Taguchi, assignor to NEC Corporation
2 April 1996 (Class 395/2.1); filed in Japan 24 December 1991

This patent describes adaptive transform coding (ATC) for speech transmission as a technique of selectively processing only the strongest spectral coefficients resulting from an analysis by any of several block orthogonal transform methods. In this variation of ATC coding, only the largest linear prediction spectral components retain their correct phase information. Weaker components are reconstructed using a pseudophase process, which seems to be a kind of extrapolation akin to phase unwrapping.—DLR

5,504,833

43.72.Gy SPEECH APPROXIMATION USING SUCCESSIVE SINUSOIDAL OVERLAP-ADD MODELS AND PITCH-SCALE MODIFICATIONS

E. Bryan George, Nashua, NH and Mark J. T. Smith, Atlanta, GA
2 April 1996 (Class 395/2.2); filed 22 August 1991

The overlap-add model for the representation of speech and other audio signals is capable of high-quality encoding, but typically involves heavy computational loads. A short-term FFT is used in an analysis-by-synthesis process to determine sets of windowed sinusoidal components with frequency, amplitude and phase parameters and grouped according to harmonic relationships. These sinusoids are added in overlapping time windows to approximate the input signal. The patent introduces an improved computational procedure and a new method of pitch information extraction.—DLR

5,506,899

43.72.Gy VOICE SUPPRESSOR

Koji Kimura, assignor to Sony Corporation
9 April 1996 (Class 379/387); filed in Japan 20 August 1993

This is a fairly typical code excited linear predictive (CELP) vocoder with the addition of a check circuit to prevent unnaturally abrupt changes in the speech output volume due to transmission errors. According to the background description, a common problem with cellular telephones is a sudden jump to a high output volume, causing severe annoyance to the listener. Here, a system of threshold checks prevents such an outburst.—DLR

5,506,934

43.72.Gy POST-FILTER FOR SPEECH SYNTHESIZING APPARATUS

Shuichi Kawama, assignor to Sharp Kabushiki Kaisha
9 April 1996 (Class 395/267); filed in Japan 28 June 1991

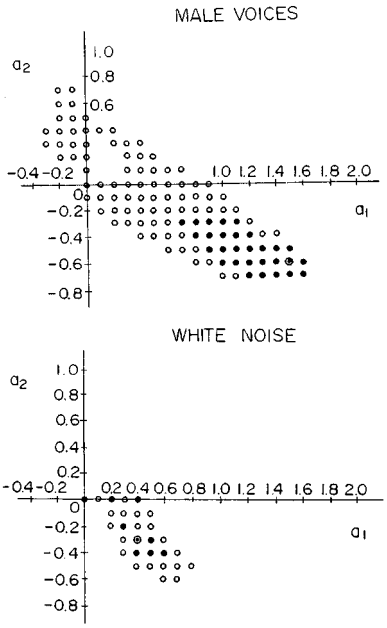
Linear prediction synthesis techniques, when used to code speech in a noisy environment, often lead to sudden changes in output amplitude and/or spectral characteristics. This patent describes a filter to be placed after the synthesis mechanism which is dynamically retuned so as to maintain a stable, gradually changing output amplitude and to enhance the spectral pattern by perceptual noise shaping.—DLR

5,509,102

43.72.Kb VOICE ENCODER USING A VOICE ACTIVITY DETECTOR

Seishi Sasaki, assignor to Kokusai Electric Company
16 April 1996 (Class 395/2.28); filed 1 July 1992

This voice presence detector is intended for use in a cordless or cellular telephone system in which the voice presence decision must be made quickly, on the order of 7 ms or less. Two different short-term predictors are



described having orders of two and six coefficients, respectively. The predictor coefficients are continuously updated using a ''simplified process of the gradient projection method,'' which is not further described. The figures show typical values from the two-coefficient predictor for male voice and white noise conditions.—DLR

5,506,933

43.72.Ne SPEECH RECOGNITION USING CONTINUOUS DENSITY HIDDEN MARKOV MODELS AND THE ORTHOGONALIZING KARHUNEN-LOEVE TRANSFORMATION

Tsuneo Nitta, assignor to Kabushiki Kaisha Toshiba
9 April 1996 (Class 395/2.65); filed in Japan 13 March 1992

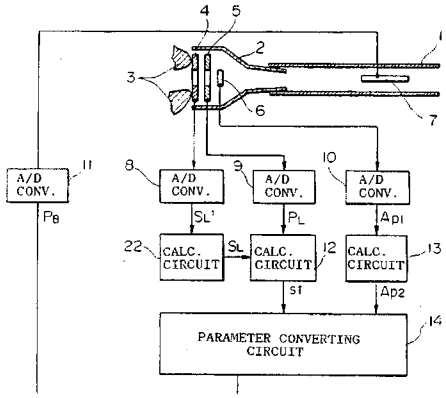
Hidden Markov models (HMMs) have been widely used as the most successful form of phonetic feature representation in large-vocabulary speech recognition systems. A recent variation is the continuous density HMM, which better models the phonetic patterns, but requires an even larger training corpus than the simple HMM. The method of this patent reduces the size of the required training corpus by modeling the HMM state output probability functions with Karhunen–Loeve-reduced orthogonal vectors.—DLR

5,543,580

43.75.Fg TONE SYNTHESIZER

Hideyuki Masuda, assignor to Yamaha Corporation
6 August 1996 (Class 84/723); filed in Japan 30 October 1990

This electronic synthesizer system is under the control of several types of sensors **4**, **5**, **6** installed within the mouthpiece **2** of a brass wind instrument. The sensors detect the contact area of lips **3**, the physical pressure of the lips on the mouthpiece and the area of opening between the lips. As a



long-time trumpeter and researcher into the physical performance of brass wind instruments, this reviewer has difficulty in understanding how these parameters can adequately control the synthesizer, but their presence in the mouthpiece would surely interfere with normal playing of a brass wind instrument.—DWM

5,537,908

43.75.Gh ACOUSTIC RESPONSE OF COMPONENTS OF MUSICAL INSTRUMENTS

Steven W. Rabe, Glendale, CA and Michael J. Tobias, Kingston, NY
23 July 1996 (Class 84/454); filed 8 February 1994

This patent describes a modified production procedure for musical instruments such as guitars, in which the strings and associated hardware of a fully assembled guitar are removed, and then replaced after the shell of the instrument is subjected ''for about 30 minutes at a plurality of frequencies over a frequency range of about 20 to about 4,000 vibrational cycles per second with a sufficiently high power density spectrum to cause a permanent change in a resonance spectrum of the partially assembled guitar.'' Presumably this ''artificial aging'' process would improve the tone producing quality of the instrument.—DWM

5,537,862

43.75.Mn QUANTITATIVE METHOD FOR EVALUATION OF THE STATE OF PIANO HAMMER FELT TONAL REGULATION

Henry A. Scarton *et al.*, assignors to Rensselaer Polytechnic Institute
23 July 1996 (Class 73/82); filed 1 November 1994

A recent Journal article by Harold A. Conklin, Jr. [J. Acoust. Soc. Am. **99**, 3286–3296 (1996)] emphasized the importance of piano hammer hardness, its measurement, and the effect upon piano tone. The present patent describes an alternative method for measuring the dynamic hardness of an elastic material such as piano hammer felt.—DWM

5,530,212

43.80.Qf MULTIPLE PHASE STETHOSCOPE

Frank A. Baffoni, East Greenwich, RI
25 June 1996 (Class 181/131); filed 12 May 1995

The stethoscope has a chest piece with a central diaphragm whose vibrations produce sound that is carried to the earpiece through the central channel of a connecting tube. The vibrations of an annular diaphragm, surrounding the central diaphragm produce sound that is carried through a separate surrounding channel and is transmitted through a flexible wall to the channel for the central diaphragm. Two diaphragm sections may be provided in the outer annular portion of the chestpiece and are stated to create a stereophonic effect.—SFL

ACOUSTICS 1996

The Executive Council decided in 1987 that several recent and newsworthy advances in acoustics should be selected each year for brief description in a format and in language which would be appropriate for publicizing progress in acoustics through science writers to the general public. Some of these short articles would also be submitted for possible publication in the annual Physics News published by the American Institute of Physics. All of the brief articles, to be solicited and selected by the President and Editor-in-Chief, would be published in our Journal. *The following two articles were selected in 1996.*

Acoustic waveguides as tools in fundamental nonlinear physics [43.20.Mv, 43.25.−x]

[S0001-4966(97)05301-0]

(Received 1 August 1996; accepted for publication 15 August 1996)
[DWM]

Since the days of Lord Rayleigh, acoustic waveguides have played a major role in the advancement of fundamental physics as well as technology. In linear as well as nonlinear[1,2] investigations, waveguides are frequently useful in restricting the dimensionality to one, which is accomplished by operating at frequencies below that of the first transverse mode so that only plane waves can propagate. Because acoustic media are nearly always nondispersive or only weakly dispersive, waveguides are also frequently employed to introduce dispersion, which arises geometrically for propagation involving transverse modes. Strong dispersion also occurs in periodic waveguides where propagation of linear and nonlinear Bloch waves have recently been investigated.[3]

Waveguides have also been used to further our understanding of *solitons*, which are exponentially localized waves of constant shape resulting from a stable balance between nonlinearity, which causes shocking, and dispersion, which causes spreading. Discovered in 1834 by Scott Russell, these waves have attracted interest in many fields of physics and in biology, and are dramatic examples of inherently nonlinear behavior. Envelope solitons, which are among the most well-known types, result from an instability in which an amplitude modulation superimposed on a wavetrain initially grows and eventually self-localizes. Azimuthal modes in a nondegenerate (e.g., elliptical) waveguide have been predicted to support these solitons.[4] Because a compression has higher temperature than a rarefaction, it may be possible to employ acoustic solitons as means to transport localized ''heat patches'' in heat pumps.

When an amplitude modulation superimposed on a wavetrain is stable, a modulation of finite extent has been predicted[5] and recently observed[4] to split into two disturbances which propagate with different velocities. This behavior is in contrast to linear theory, where such a disturbance undergoes distortion due to dispersion but does not split. As a consequence of the double group velocity, a signal that is amplitude-modulated at the source becomes frequency-modulated at periodic positions in space. Axially symmetric cross-modes in a cylindrical duct belong to this class. A possible application of these results is to broadband tunable lasers for light in the visible using fiber optics. For appropriate values of the parameter, an amplitude-modulated green light alternating between bright and dim at the source alternates between red and blue down the fiber. This mechanism thus allows the possibility of tunable coherent light from a single-frequency coherent source.[4]

Another recent application of waveguides is in the determination of stress–strain relations in rocks. Rock is extremely nonlinear, and its static and dynamic behavior exhibits hysteresis and discrete memory. Nonlinear pulse propagation experiments[6] in meter-long sandstone rods have detected 2nd and 3rd harmonic growth at strains of the order of only 10^{-7} (extrapolation to a 1-mm sample implies deformations of only one atomic diameter). The nonlinear behavior of rock has important consequences for processes such as earthquake slip[7] and stress fatigue damage in concrete.

Waveguides can also be used to probe the nature of systems driven far from thermodynamic equilibrium. We have recently measured a weak signal's attenuation due to high-intensity shockless noise in one dimension.[8] In agreement with theory,[9] the amplitude attenuates with distance as a *Gaussian* rather than an exponential, displaying a breakdown in translational invariance. Also in accord with theory,[10] we have observed[8] that the high-frequency spectrum follows an f^{-3} power law. Both the nonexponential relaxation and the power law spectrum are characteristics of one-dimensional shockless noise being far off equilibrium. The observed power law spectrum raises the possible existence of a collective mode analogous to zero sound in He^3 (Ref. 11). Also, because the density of states in three dimensions is proportional to f^2, a one-dimensional spectrum proportional to f^{-3} suggests the possibility of an $1/f$ spectrum in three dimensions.

ANDRÉS LARRAZA
Department of Physics, Code PH/La,
Naval Postgraduate School,
Monterey, California 93943

BRUCE DENARDO
Department of Physics and Astronomy, and National Center for Physical Acoustics,
University of Mississippi, University, Mississippi 38677

[1] M. Hamilton, ''Fundamentals and applications of nonlinear acoustics,'' in *Nonlinear Wave Propagation in Mechanics*—AMD Vol. 77, edited by T. W. Wright (The American Society of Mechanical Engineers, New York, 1986), and references therein.
[2] D. T. Blackstock, ''Nonlinear Acoustics (theoretical),'' in *American Institute of Physics Handbook*, edited by D. E. Gray (McGraw-Hill, New York, 1972), 3rd ed.
[3] C. E. Bradley, ''Time harmonic acoustics: Block wave propagation in periodic waveguides. Part III. Nonlinear effects,'' J. Acoust. Soc. Am. **98**, 2735–2744 (1995).

[4] A. Larraza and W. Coleman, ''Solitons, pulse-splitting and AM–FM conversion in cylindrical ducts,'' J. Acoust. Soc. Am. **100**, 139–147 (1996); W. F. Coleman, ''Pulse splitting and AM–FM conversion in a nonlinear dispersive medium,'' M.S. thesis, U.S. Naval Postgraduate School, 1993.
[5] G. B. Whitham, *Linear and Nonlinear Waves* (Wiley–Interscience, New York, 1974), Secs. 14.2 and 15.4.
[6] J. A. TenCate, K. E. A. Van Den Abeele, T. J. Shankland, and P. A. Johnson, ''Laboratory study of linear and nonlinear elastic pulse propagation in sandstone,'' J. Acoust. Soc. Am. **100**, 1383–1391 (1996).
[7] I. A. Beresnev and K.-L. Wen, ''The possibility of observing nonlinear path effect in earthquake-induced seismic wave propagation,'' Bull. Seism. Soc. Am. **86**(4) 1028–1041 (August 1996).
[8] A. Larraza, B. Denardo, and A. Atchley, ''Absorption of sound by noise in one dimension,'' J. Acoust. Soc. Am. **100**, 3554–3560 (1996).
[9] O. Rudenko and A. Chirkin, ''Theory of nonlinear interaction between monochromatic and noise waves in weakly dispersive media,'' Sov. Phys. JETP **40**, 945–949 (1975).
[10] V. P. Kuznetsov, ''On the spectra of high intensity noise,'' Sov. Phys. Acoust. **16**, 129–130 (1970).
[11] A. Larraza and G. Falkovich, ''Collective modes in open systems of nonlinear random waves,'' Phys. Rev. B **48**, 9855–9857 (1993).

Speech production parameters for automatic speech recognition [43.72.Ne, 43.70.Aj]

[S0001-4966(97)05401-5]

(Received 29 July 1996; accepted for publication 1 August 1996) [DWM]

What we hear as speech is produced by the continuous movement of the speech articulators, such as the tongue, lips, and larynx. These articulators modulate air flow in such a way that speech sounds reach our ear. Do we in any way perceive the movements of those articulators as part of our perception of human speech? One of the mysteries of speech production and perception is the transformation between the discrete units of linguistics and the continuous nature of speech production. For instance, the three distinct sounds, or phones, in the word *cop* are blended together in a continuous waveform, which is created by the continuous movement of the speech articulators. The human listener is able to decompose the continuous sound stream to recall the component sounds of the word. Does the listener do this strictly with the acoustic signal without reference to how the articulators move, or is that movement an object of perception?

A similar problem with decomposition exists in automatic speech recognition by machine, ASR, where a continuous signal must be decoded into a string of discrete units. Would it help statistical automatic speech recognition algorithms to incorporate constraints that are inherent in the speech production process, thus, at least partly, characterizing the sending channel? There appears to be a parallel between the difficulty in answering the question of whether humans perceive speech directly as gestures of the tongue and lips, and deciding how to include articulatory constraints in the statistical models used in ASR.

Much of the progress towards incorporating articulatory representations into ASR has been enabled by laboratories that measure and model speech articulatory movement and coordination. Models, such as the task-dynamic model, have been used recently. This is a control model of the coordinated movement of the speech articulators, where the control parameters are set for phone or diphone length intervals. These control parameters can be blended with parameters belonging to neighboring phones and diphones, and can provide a bridge from the discrete phone units to the continuous movement of articulators. The sequence of these parameters that are used to produce a word can be written into score form, called a gestural score. Something akin to gestural scores may provide part of the abstract representation of articulatory movement that several researchers are considering.

Recent work at The Ohio State University has shown how gestural scores can be derived from data on tongue and lip movement.[1] The articulatory movement data were obtained using the x-ray microbeam machine at the University of Wisconsin. This machine produces x-ray images of tongue and lip movement using very low dosages of x rays. The OSU group was further able to train a neural network to perform phone recognition from the gestural scores that they derived. To do ASR with gestural scores in the sense of recognizing the words of a spoken message from the acoustic waveform will require that gestural scores be derived from the waveform. Researchers at Haskins Laboratories are experimenting with recovering gestural scores from the speech waveform in computer simulation experiments.[2] Professor Li Deng of the University of Waterloo has been incorporating gestural scores into speech recognition systems.[3] Further progress in all of these areas can be expected.

RICHARD S. McGOWAN AND ALICE FABER
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511

[12] T-P. Jung, A. K. Krishnamurthy, S. C. Ahalt, M. E. Beckman, and S-H. Lee, ''Deriving gestural scores from articulatory-movement records using weighted temporal decomposition,'' IEEE Trans. Speech Audio Process. **14**, 2–18 (1996).
[13] R. S. McGowan and M. Lee, ''Task-dynamic and articulatory recovery of lip and velar approximations under model mismatch conditions,'' J. Acoust. Soc. Am. **99**, 595–608 (1996).
[14] L. Deng and D. X. Sun, ''A statistical approach to automatic speech recognition using the atomic speech units constructed from overlapping articulatory features,'' J. Acoust. Soc. Am. **95**, 2702–2719 (1994).

The detection of breast microcalcifications with medical ultrasound[a)]

Martin E. Anderson
Department of Biomedical Engineering, Duke University, Room 136 Engineering Building, Durham, North Carolina 27708

Mary S. Soo and Rex C. Bentley
Duke University Medical Center, Durham, North Carolina 27708

Gregg E. Trahey
Department of Biomedical Engineering, Duke University, Room 136 Engineering Building, Durham, North Carolina 27708

(Received 19 February 1996; accepted for publication 14 June 1996)

Microcalcifications are small crystals of calcium apatites which form in human tissue through a number of mechanisms. The size, morphology, and distribution of microcalcifications are important indicators in the mammographic screening for and diagnosis of various carcinomas in the breast. Although x-ray mammography is currently the only accepted method for detecting microcalcifications, its efficacy in this regard can be reduced in the presence of dense parenchyma. Current ultrasound scanners do not reliably detect microcalcifications in the size range of clinical interest. The results of theoretical, simulation, and experimental studies focused on the improvement of the ultrasonic visualization of microcalcifications are presented. Methods for estimating the changes in microcalcification detection performance which result from changes in aperture geometry or the presence of an aberrator are presented. An analysis of the relative efficacy of spatial compounding and synthetic receive aperture geometries in the detection of microcalcifications is described. The impact of log compression of the detected image on visualization is discussed. Registered high resolution ultrasound and digital spot mammography images of microcalcifications in excised breast carcinoma tissue and results from the imaging of suspected microcalcifications *in vivo* are presented. [S0001-4966(97)02612-X]

PACS numbers: 43.10.Ln, 43.80.Qf, 43.80.Jz, 43.80.Vj [FD]

INTRODUCTION

A. The significance of microcalcifications in mammography

The primary object of screening mammography is the early detection of breast cancer. Such detection can decrease the mortality and morbidity associated with breast cancer.[1] Microcalcifications (MCs) are small crystals of calcium apatites which form in human tissue through a number of mechanisms. Their size, morphology, and distribution are important indicators in the mammographic screening for and diagnosis of various carcinomas in the breast. MCs present in approximately 40% of cancers, and in some cases they are the only indication of malignancy at mammography, making their detection and interpretation critical.[1–4] MCs can present across a broad continuum of sizes, from several millimeters down to the resolution limit of mammography. Not all types are associated with cancer.[5] Currently, x-ray mammography is the gold standard for such screening mammography and the only accepted method for screening for MCs.[1]

B. The radiologically dense breast

The natural radiological density of certain types of glandular or fibrous breast tissue can reduce the sensitivity and specificity of mammography. (Note that in this context ''density'' refers to the attenuative character of the tissue, as opposed to the optical density of the mammography film.) One way in which it does this is by raising the local background density of the image, thus lowering the effective contrast of lesion(s) and/or microcalcifications against that background. Another is by extending the dynamic range of the image, which can then exceed that of the x-ray film used. As a consequence it can be difficult for the clinician to choose an exposure which optimally images all regions of the breast. Dense parenchyma also increases scattering of the x rays, further reducing the image contrast. Finally, the longer exposure time which dense parenchyma necessitates increases the likelihood that the mammography image will be degraded by patient motion artifact. It has been suggested that one of the primary causes of false negatives in the early detection of cancer is inadequate imaging of the dense breast. Approximately one in four women have dense breasts, which gives a measure of the magnitude of the problem.[6,7]

It has been shown that breast density is inversely correlated with age, such that the efficacy of mammography in the young breast is reduced.[8–10] Brekelmans *et al.* propose this

[a)]''Selected research articles'' are ones chosen occasionally by the Editor-in-Chief, that are judged (a) to have a subject of wide acoustical interest, and (b) to be written for understanding by broad acoustical readership.

as one probable cause of reduced sensitivity in detecting early cancers.[7]

C. The current role of ultrasound in mammography

Medical ultrasound plays an important role in the breast clinic as an adjunct to conventional x-ray mammography. Its primary uses include the differentiation of solid lesions from benign cysts, the examination of the matrix of solid lesions, and guiding needle biopsy.[6] Other important applications are in the examinations of young women, women with dense breasts, and women with breast implants, which are radio-opaque.[11,12] It is important to note that a radiologically dense breast may image well under ultrasound.

D. Microcalcifications under ultrasound

Current ultrasound technology and protocol does not reliably detect MCs.[13,14] Microcalcifications which lie within a hypoechoic region, such as the interior of a hypoechoic lesion, are more easily detected.[11,15] Their visualization is limited by a number of factors which may include speckle noise, phase aberration, the system spatial resolution, attenuation, display parameters, and human perception of the displayed image. Estimates of the relative impact of these effects could potentially guide attempts to improve detection performance.

The analysis of MC detectability is limited by the poor characterization of their acoustic properties *in vivo*. Filipczynski *et al.* present an analysis of MC detectability based on the theoretical radiation patterns of rigid and elastic spheres and clinically measured speckle noise levels.[16–18] This analysis was based on a 5-MHz system center frequency and assumed that the acoustic properties of MCs are similar to those for bone. As a consequence, this analysis does not reflect the capabilities of modern 7.5- and 10-MHz transducers and may underestimate the reflectivity of MCs. Based on the acoustic impedance of hydroxyapatite, the most common constituent of MCs, their amplitude reflection coefficient in tissue is close to 0.9.[5,19] For this reason, a subresolution MC is modeled below as a bright point reflector under ultrasound.

I. THEORETICAL FRAMEWORK FOR DETECTION PERFORMANCE OPTIMIZATION

The detectability of MCs under ultrasound is most likely affected by a number of factors, including, but not limited to, the spatial resolution of the imaging system, speckle noise in the image, and phase aberration. An analysis of the impact of these factors on MC visualization could guide the optimization of imaging systems for this task. We present a discussion of these factors and a theoretical framework for this analysis below. The ultimate goal of such investigation is the improved visualization of MCs in the clinic.

A. Spatial resolution

While MCs are most likely to be bright reflectors, for detection their small size would require the use of a system with a small resolution volume focused on the MCs in order for them to return sufficient signal for detection relative to the surrounding diffuse scatterers. The design factors affecting the resolution of an ultrasound system are well understood. Primary among these are the center frequency, bandwidth, and aperture size of the transducer used to transmit and receive the ultrasound signal. System resolution must be traded against depth of penetration as the attenuation coefficient of tissue also increases with frequency. The improved sensitivity of ultrasound in cancerous lesion detection which accrues with an increase of system resolution has been demonstrated.[11,20–22] In a study of 14 lesions presenting microcalcifications at mammography, Jackson *et al.* found that the visualization of microcalcifications was improved in 57% of the patients on changing from a 4-MHz transducer to a 7.5-MHz transducer. In four patients the microcalcifications were visible only under the higher-frequency transducer. In the two patients with no accompanying mass, neither transducer allowed them to be imaged.[22]

B. Synthetic receive aperture imaging

One means to improve the resolution of the system is to increase its aperture size. The additional imaging system complexity associated with a larger aperture can be reduced through synthetic receive aperture (SRA) imaging. An SRA system transmits into the same region of interest several times from a single transmit aperture. After each transmit, the echo signals are received on a different receive subaperture. These signals are then coherently summed to form a large effective receive aperture. Such a system requires only sufficient channels to populate each receive subaperture.[23]

C. Speckle reduction

Another probable cause of the failure of clinical systems to detect MCs is that their bright signals are obscured by speckle noise. One means to reduce speckle noise is the technique of spatial compounding, through which the speckle patterns received on discrete apertures from the region of interest are averaged, reducing the variance of the speckle. This in turn increases the effective signal-to-noise ratio of the coherent MC echo to the speckle noise.[24]

D. Phase aberration

The spatial and contrast resolution of medical ultrasound can be severely limited by a phenomenon known as phase aberration. The steering and focusing of an ultrasound beam using a phased array relies on an approximation of the velocity of sound in tissue, usually 1540 m/s. In fact, the velocity of sound through different tissues can vary greatly. As an acoustic wavefront passes through inhomogeneous tissues, it can become distorted as portions of its surface are advanced or retarded. On transmit this phenomenon affects the focusing and steering of the system point spread function (PSF). The returning echoes incident on the elements of the transducer array are also misaligned such that when these signals are summed to form a single echo line they no longer sum coherently. In a comprehensive study of the acoustic properties of both healthy and cancerous breast tissues, Edmonds *et al.* measured a range sound velocities from 1400 to over 1600 m/s.[25] This suggests that imaging in the breast will very likely be affected by phase aberration.

E. Visualization as a detection problem

It is assumed that a subresolution MC can be modeled as a point target, and is to be detected on the basis of amplitude by an ideal observer. In detection theory, the probability of a correct decision is described in terms of the probabilities of the H_1 and H_0 hypotheses, which in this context corresponds to the ''MC present'' and ''MC absent'' hypotheses, respectively. If an amplitude threshold is applied to the ultrasound signal to decide whether a MC is present, the probabilities of a correct detection (Pd) and that of a false alarm (Pf) are predicted by the integrals above that threshold of the probability density functions of amplitude corresponding to the H_1 and H_0 hypotheses. Receiver operating characteristic (ROC) curves plot Pd versus Pf as the amplitude threshold is allowed to vary. Such curves thus indicate detection performance as the stringency of the decision criterion ranges from low sensitivity and high specificity (low Pd, low Pf) to high sensitivity and low specificity (high Pd, high Pf). The relative performance of different systems can be compared using the corresponding ROC curves.

A ROC analysis comparing the relative detection performance of four imaging systems is presented below. These include a spatial compounding system, an SRA system, and two conventional systems having different spatial resolutions. An ROC analysis demonstrating the impact on detection of phase aberration modeled as a thin phase screen is also presented below. It is important to note that these analyses are presented solely as a means to compare system performance and do not reflect the performance of a human observer.

F. Theoretical probability density functions of amplitude

The statistics used here to describe ultrasound speckle are drawn from the literature of laser optics.[26] In fully developed speckle, the complex radio-frequency echo signal from diffuse scatterers alone has a zero mean, two-dimensional Gaussian probability density function (PDF) in the complex plane. Envelope detection removes the phase component, creating a signal with a Rayleigh amplitude distribution. When a bright target, such as a subresolution microcalcification, is introduced to the speckle, it adds a constant strong phasor to the diffuse scatterers echoes and shifts the mean of the complex echo signal away from the origin in the complex plane. Upon detection, this has the effect of changing the Rayleigh distribution into a Rician distribution. The Rician PDF is defined by the following equation:

$$p_A(a) = \frac{a}{\sigma^2} \exp\left(- \frac{a^2+s^2}{2\sigma^2} \right) I_0\left(\frac{as}{\sigma^2} \right). \quad (1)$$

This distribution is nonzero for $a>0$ only. The parameter s is the strength of the bright scatterer, while σ is the standard deviation of the complex Gaussian described above. I_0 is the incomplete Bessel function of zero order. The Rician distribution is parameterized by the variable k, which is defined as s/σ.[26] The Rician distribution reduces to the Rayleigh distribution for the special case $s=0$. A family of Rician distributions for various values of k is shown in Fig. 1. Note that the curve for $k=0$ corresponds to the Rayleigh distribution. In this discussion it is assumed that a subresolution microcalcification contributes a constant strong phasor to the echo signal, although it is not possible at this time to predict the k parameter corresponding to a particular size MC in breast tissue.

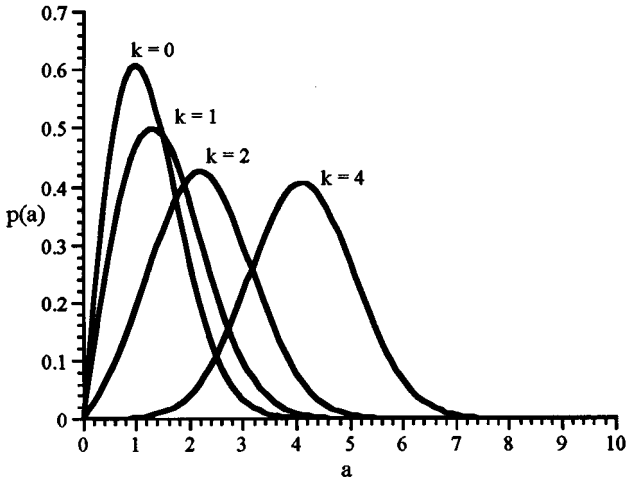


FIG. 1. Family of Rician probability density functions, parameterized by **k** value for a σ value of 1.[26]

While the development of the Rician statistic in optics implies monochromicity, one can generalize the Rayleigh and the Rician statistics to the broadband case, such as an ultrasound system. If acoustic propagation is limited to the linear regime, the broadband signal of an ultrasound system can be represented as the linear combination of a series of monochromatic systems, each at a different frequency. The insonification of diffuse scatterers with each of these systems produces a Gaussian echo signal as described above. By the principle of the orthogonality, these signals are statistically independent. The broadband signal produced by their superposition is the summation of independent Gaussian random variables, and is thus also Gaussian. For a broadband system the constant phasor which distinguishes the Rayleigh from the Rician correctly describes the signal from a strong scatterer provided this scatterer is at the focus.

G. ROC comparison of aperture geometries

In this context the differences between the imaging systems described manifest themselves in the parameters of the Rician [Eq. (1)]. For the analysis of aperture geometry, the s value was constant for the H_1 case and equal to zero for the H_0 case. The geometries used are shown in Fig. 2, where the f number (f/#) refers to the ratio of the focal range to aperture size. The $f/2$ control used the center half of the aperture elements to transmit and receive, while the $f/1$ control used the entire aperture to transmit and receive. The SRA system transmitted twice from the center of the array, receiving first on the center elements and then on the flanking elements, forming an effective $f/2$ on transmit, $f/1$ on receive system. The spatial compounding geometry transmitted and received on each half of the array separately, summing the echoes from the two halves after they had been envelope detected.

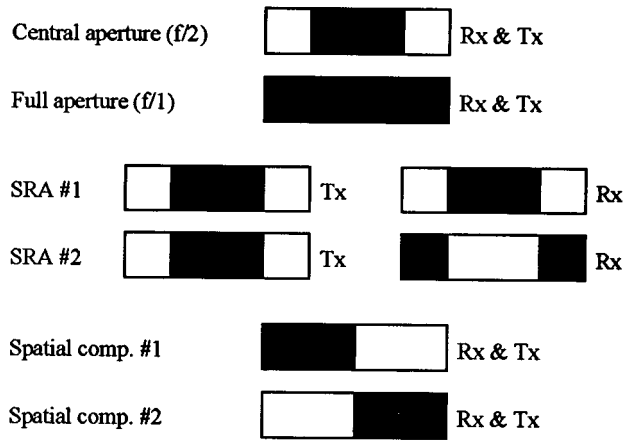


FIG. 2. Geometries of apertures of simulated imaging systems used, with active portions of aperture shown in black.

The PDFs used to generate the theoretical ROC curves are summarized in Table I, where $p_A(\alpha,s,\sigma)$ is the Rician distribution shown above with parameters α, s, and σ, and $*$ represents the convolution operation. The s parameter was a constant in the H_1 case and zero in the H_0 case. The derivation of these σ values is described in the Appendix.

H. Methods

To simulate the imaging process described above for the purpose of creating data for ROC analysis, fields of random numbers with Gaussian amplitude distribution of zero mean and unit variance were created. A central point in each field was set to either zero (H_0 case) or ten (H_1 case). This strong point scatterer surrounded by a field of randomly weighted scatterers models the presence of a MC in an environment of weaker diffuse scatterers. The PSFs of the various aperture configurations were also created using an acoustic field simulation program.[27] The input field was convolved with the PSF of each system under evaluation for both H_1 and H_0 cases, including the controls. This convolution defined a scattering grid of 15-μm, spacing laterally by 15.4-μm spacing axially, resulting in a scatterer density of over 100 scatterers per resolution cell for all the systems simulated. Finally, the resulting echo patterns were envelope detected using the Hilbert transform and the amplitude was recorded at the target location for each imaging system and the controls. After a series of 1000 trials, histogram PDFs for each system were created from which ROC curves were calculated.

TABLE I. Summary of PDFs used to generate theoretical ROC curves.

System	PDF
$f/2$ control	$p_{A\|H_1}(a,s,\sigma)$
$f/1$ control	$p_{A\|H_1}(a,s,0.707\sigma)$
SRA	$p_{A\|H_1}(a,s,0.791\sigma)$
Spatial compounding	$p_{A\|H_1}(a,s,1.03\sigma)*p_{A\|H_1}(a,s,1.03\sigma)$

I. Theoretical/simulation results for aperture geometries

The theoretical results are shown as solid curves in Fig. 3(a)–(d). These curves represent the system performance under relatively difficult detection conditions, i.e., for the case where the MC has a relatively low strength of $\mathbf{k}=1.265$. The $\mathbf{k}$ value is a function both of the original scatterer field and the system point spread function. For these simulations $\mathbf{k}$ was estimated from the first order statistics of the H_1 and H_0 histograms for the $f/2$ control, following the theoretical expressions given by Goodman.[26] This $\mathbf{k}$ value was also calculated using the simulated point spread function following the method described in the Appendix, giving a value of 1.280. The curves for the SRA system and the spatial compounding system are almost identical. Both systems perform better than the $f/2$ control case, while the $f/1$ control performs best of all. If based on these results alone, the better choice of imaging method between SRA and spatial compounding is not indicated.

The simulation results for the same set of parameters are superimposed on the respective theoretical curves in Fig. 3, (a)–(d). These results are in good agreement with the theoretical results. A common method of reducing a ROC curve to a single index of performance is to integrate the area under the curve.[28] This parameter ranges from 0.5 to 1, with a greater area indicating better performance. In this context it is also the expected fraction of correct diagnosis by an ideal observer. The areas under the theoretical curves are included for comparison in Fig. 3(a)–(d).

J. ROC analysis of the impact of aberration

The model of phase aberration as a thin phase screen at the aperture can be used to compare the performance of a system with an aberrator present to that of the unaberrated control. In the simulations discussed below, the aberrator is applied as a random phase error on the elements of the array. This random error is described in terms of its standard deviation (rms phase error) and its spatial autocorrelation function across the aperture. This spatial autocorrelation function is assumed to be Gaussian with a known full-width-half-maximum (FWHM). The choice of appropriate first- and second-order statistics to describe aberration in the breast is hindered by the lack of comprehensive measurements of such aberrators *in vivo*.

Assuming aberrators of random structure, the authors find it most meaningful to characterize system performance for a statistically defined class of aberrators, rather than for a single realization. Over an ensemble of aberrators of particular statistics, the s and σ parameters for each realization can vary considerably, defining a family of ROC curves for a particular target strength. In order to compare the families of ROC curves produced over different classes of aberrators to each other and to the control, the area under every curve in the ensemble was calculated and the statistics of area for each class of aberrator is reported.

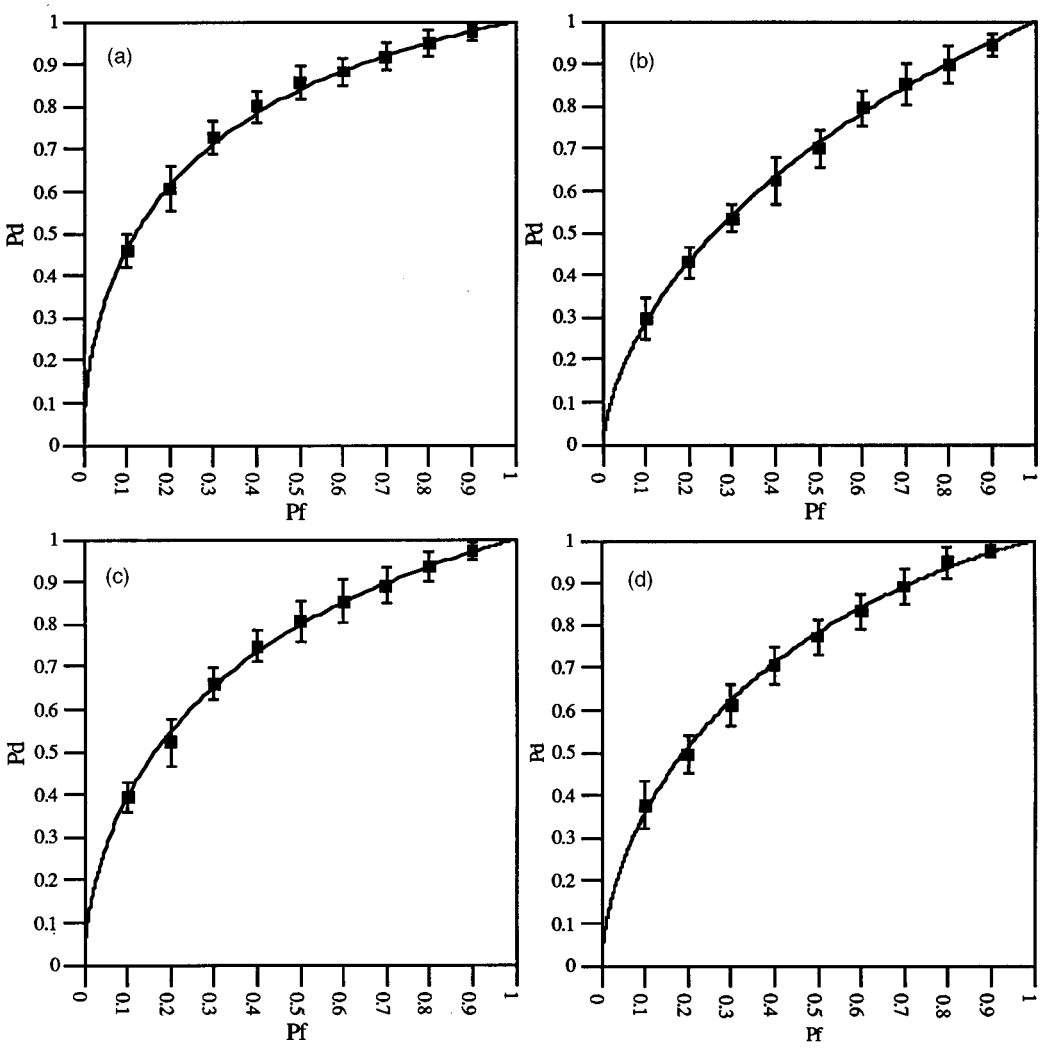


FIG. 3. (a) $f/1$ control, ROC area=0.775. (b) $f/2$ control, ROC area=0.665. (c) Synthetic receive aperture, ROC area=0.736. (d) Spatial compounding, ROC area=0.720. Given the model of microcalcification detection described in the text, these ROC curves show the relative performance of four different aperture geometries, described in the text and Fig. 2. Simulation results with error bars (± 1 standard deviation) are plotted on theoretical curves.

K. Methods

Random aberrator profiles having the first- and second-order statistics of interest were created and applied to a simulated ultrasound transducer in the form of a timing error on each element on both transmit and receive. For each aberrator and for the unaberrated control the corresponding PSF at the focus was created using an acoustic field simulation program.[27] The energy of each PSF over the region of support was calculated and used to estimate the Rician σ parameter for each realization. Each PSF was then envelope detected using the Hilbert transform, and the peak envelope amplitude used to estimate the corresponding Rician s parameter. Specifically, to form each $\mathbf{k}$ estimate the envelope peak value was divided by the square root of the PSF energy for each trial. This ratio was then scaled by a constant chosen to produce the desired $\mathbf{k}(\mathbf{k}=s/\sigma)$ parameter for the unaberrated control. This constant is equivalent to the original scatterer strength, as opposed to the strength s of the echo signal returned from it. The justification for this method is presented in the Appendix.

For each realization, the area under the ROC curve generated using the corresponding $\mathbf{k}$ estimate was calculated by numerical integration. After 1000 trials the mean and standard deviation of the area was found for each ensemble of ROC curves, and hence for each class of aberrator.

L. Simulation results

The mean ROC areas and associated standard deviations for three classes of aberrators are listed in Table II. The $\mathbf{k}$ value for the control was 1.5. The aberrators used all had a Gaussian spatial autocorrelation function with a 6-mm FWHM. The severity of the aberrators were 10-, 20-, and

TABLE II. Mean ROC areas and standard deviations for three classes of aberrators.

rms phase error (ns)	Mean ROC area	Standard dev.
0 (control)	0.714	NA
10	0.690	0.019
20	0.632	0.045
30	0.594	0.045

30-ns rms phase error. The simulated system was a 10-MHz $f/1$ system with 100% bandwidth. For this system these aberrators are weak relative to aberrators measured in the breast.[24,31] The results show a significant decrease in performance with the increase in aberrator severity.

II. EMPIRICAL DATA ACQUISITION

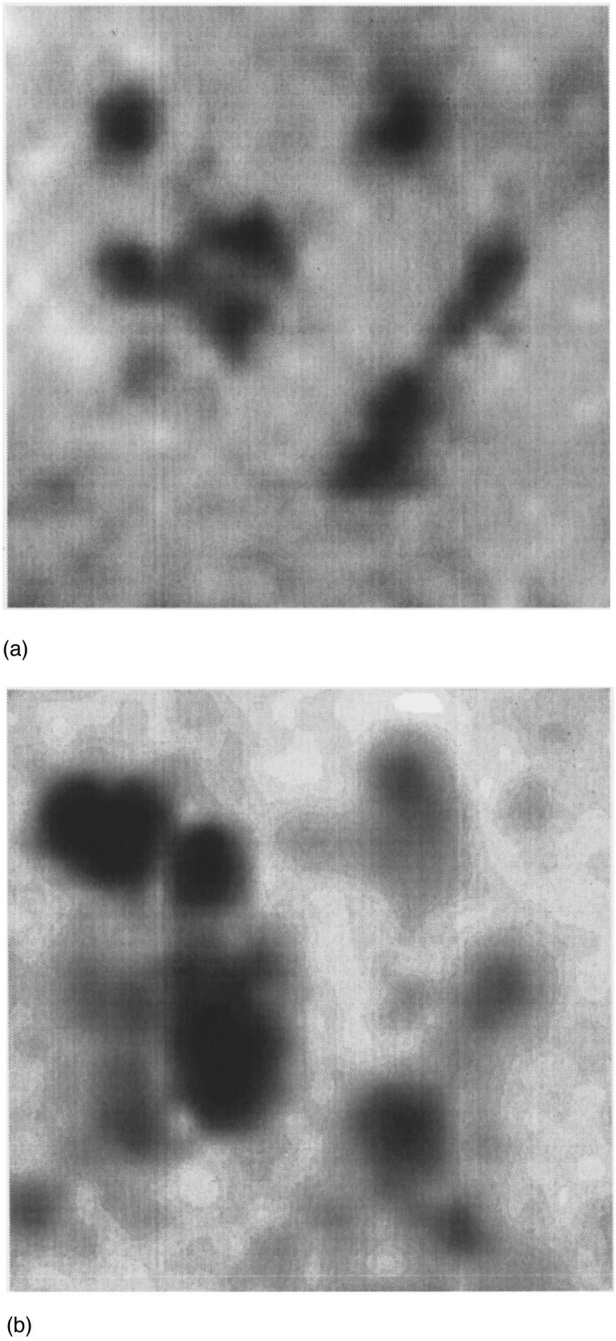
Ultrasound data from excised tissue samples of breast carcinoma which contain MCs have also been collected. These samples were obtained from excisional biopsy and mastectomy specimens. For some samples digital spot mammography was also used to test whether MCs were present and to determine their location if found. Suspected microcalcifications have also been imaged *in vivo*.

A. Methods for and images of excised tissues

In this procedure, upon excision and transport to the ultrasound laboratory on ice the unfixed tissue sample was immobilized in the center of a polystyrene Petri dish with a thin layer of 10% gelatin in lactated Ringer's solution, an iso-osmotic buffer. Two small lead beads placed in the gel served as position markers. The specimen underwent specimen digital spot mammography. All views were acquired on a LORAD StereoGuide digital spot mammography system with a CCD device having just over 10 lines/mm resolution (512 lines/5 cm on each axis).

The disk of gelatin holding the specimen was immersed in lactated Ringer's solution at room temperature and scanned using ultrasound. The lead beads reflected ultrasound well and served as reference points for specimen registration. These scans were carried out using an $f/1.3$ 10-MHz Panametrics piston transducer. The transducer was driven using a Tektronix PG501 pulse generator triggering a Metrotek MP215 ultrasound pulser. The echo signals were received using a Metrotek MR101 receiver and digitized at 100 MHz at 8-bit resolution using a Lecroy 9424E digital oscilloscope and stored on a computer. The transducer was translated using a computer-controlled NTR Systems 3-D positioning system. By digitizing echo lines at 100-μm increments in the two dimensions perpendicular to the beam, a volume of echo data was acquired over each region of interest.

A pair of images of a cluster of MCs scanned in this manner are shown in Fig. 4(a) and (b). Each image has been interpolated to a finer grid using bicubic interpolation, and is shown in inverted grayscale such that MCs appear as dark regions. Each image represents a region 2.5 mm^2 in area. Figure 4(a) is the digital spot mammography image of the cluster. This image has been enhanced by the removal of a planar intensity component which was a consequence of the local variation in specimen thickness. Figure 4(b) is the detected ultrasound image, here presented as the maximum value projection along the axis of acoustic propagation for the sake of comparison. The peak amplitude of their echoes is a function of their size and their axial position relative to the focus of the transducer, and is on the order of 20 dB higher than the mean amplitude of the surrounding speckle. Some differences between the images can be seen, and is to be expected considering the differences between the mammography and ultrasound systems as well as the tissue parameters imaged by them. The theoretical FWHM lateral resolution of the transducer used is just under 200 μm, approximately the size of the MCs in the ultrasound image. Thus, the MCs shown in Fig. 4 appear to be subresolution. This high resolution, highly focused imaging system provides good visualization of MCs under what are ideal conditions relative to *in vivo* imaging.



(a)

(b)

FIG. 4. Cluster of microcalcifications in excised breast carcinoma under (a) digital spot mammography and (b) high-resolution ultrasound, shown in inverted grayscale.

B. *In vivo* methods and observations

A 66-year-old volunteer presenting at mammography with a cluster of MCs approximately 1.5 cm from the skin

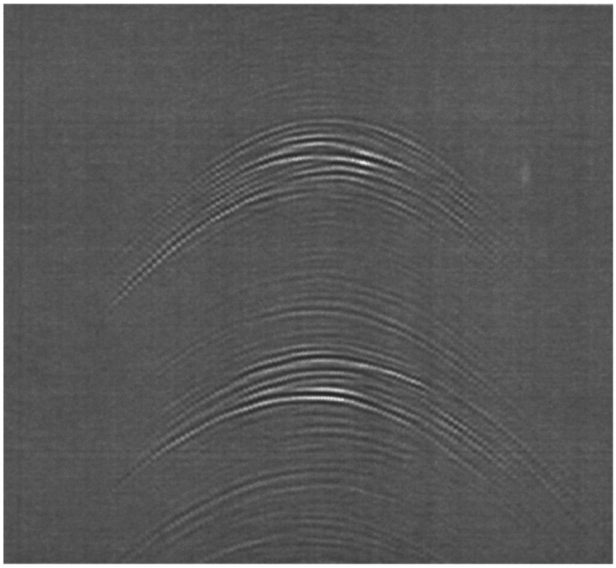


FIG. 5. rf echoes from a pair of suspected microcalcifications *in vivo*.

surface was recruited and underwent a localized ultrasound exam. The MC cluster had been the object of a previous needle core biopsy which did not remove the MCs and which left a small scar on the skin. An Elegra scanner manufactured by the Siemens Medical Systems Ultrasound Group equipped with a 7.5-MHz linear array transducer was used to scan the breast in the region of this scar. In this imaging mode the scanner provided FWHM spatial resolution of approximately 200 μm axially and 220 μm laterally. After adjustment of the B-mode image gain and logarithmic compression, a pair of bright targets were identified at approximately 13 and 16 mm depth. These targets appeared as isolated points rather than extended structures under dynamic scanning. The unfocused radio frequency (rf) data on each channel of the transducer were simultaneously captured for the transmit scan line passing through the centers of the targets. These data were captured several times over a range of system gain settings and stored on a computer. The data set found to have the maximum gain without saturation was selected for further analysis. One data set was also captured at the same gain settings within the same scan plane through the tissue approximately 6 mm away from the targets laterally. This data set was used to calculate a rough estimate of the echogenicity of the surrounding tissue.

The mammogram films on which this cluster were localized were also examined using a binocular microscope equipped with a measurement reticule. The cluster consisted of 6 MCs ranging approximately 200–550 μm in diameter. The poor contrast of some of the targets prevented exact measurement.

One rf data set is shown in Fig. 5. The deviation in the echo arrival time profile in this data from the geometric delay profile is used below as an estimate of phase aberration due to tissue inhomogenaities. In order to estimate the arrival time profile across the array, segments of the echo data surrounding each target were upsampled by a factor of 8 using interpolation and aligned on a channel-to-channel basis using normalized cross-correlation. This method of alignment is a refinement of that described by Flax and O'Donnell.[29] The signal amplitude is progressively diminished away from the center of the aperture due to the limited angular response of the array elements, which each have a lateral FWHM beamwidth of approximately $\pm 23°$. Further analysis was restricted to the group of channels at the center of the aperture which had interelement correlation values of $\geqslant 0.5$. As the exact location of the targets relative to the transducer was unknown, the respective local peaks of the envelope-detected signal were used as range estimates.

For a linear array of transducer elements receiving echoes from a point target, the geometric delay at each element τ_e is defined by the equation

$$\tau_e = \frac{\sqrt{(x_e - x_t)^2 + y_t^2 + z_t^2}}{c}, \tag{2}$$

where $(x_e,0,0)$ is the element location, (x_t,y_t,z_t) is the target location, and c is the speed of sound. Once this quantity is squared it becomes a second-order polynomial. Thus, to find the best-fit geometric delays for the *in vivo* data, a second-order polynomial was fit (by least-mean-squared) to the square of the measured arrival time profiles, and the square root of this best-fit curve was taken to find the geometric delays. These were then subtracted from the measured arrival time profiles to achieve focusing. After focusing, the arrival time profiles have a residual phase error of 7.1 ns for the proximal target and 8.5 ns for the distal target. It should be noted that the distal target may be subject to acoustic shadowing by the proximal target. The phase profiles before and after focusing for the proximal target are shown in Fig. 6(a) and (b).

The patterns of echoes seen in Fig. 5 are similar to those observed from wire and point targets in water tank experiments. Some pulse distortion was evident. If the echoes were in fact from the MCs observed at mammography, it should be noted that these were not truly subresolution targets. Diffraction, resonance, and ''creeping-wave'' effects associated with reflection from elastic targets on the order of the insonification wavelength may be contributing to this distortion.[30] As a control for phase error, a wire target at a depth of 19.7 mm in a Radiation Measurements, Inc. tissue mimicking phantom was also imaged and the data analyzed in the manner described above. After geometric focusing the wire target echoes had a residual r.m.s phase error of 5.6 ns.

The targets are also highly echogenic, making it unlikely that these targets were merely bright speckles. To obtain a crude estimate of the background speckle echogenicity, a control data set for an rf line through the surrounding tissue (described above) was used. Both the target data set and the control data set were synthetically focused and summed for each target in turn. The local peak of the envelope detected signal was used as an estimate of the Rician s parameter for each target. The standard deviation of the control rf over a 2.5-mm window centered on these peak values was used as an estimate of the Rician σ parameter. The **k** parameter estimates for the targets was found by calculating the corresponding s/σ ratio. The approximate **k** values found by this method were 17 for the proximal target and 44 for the distal target. Technical limitations of the data acquisition process

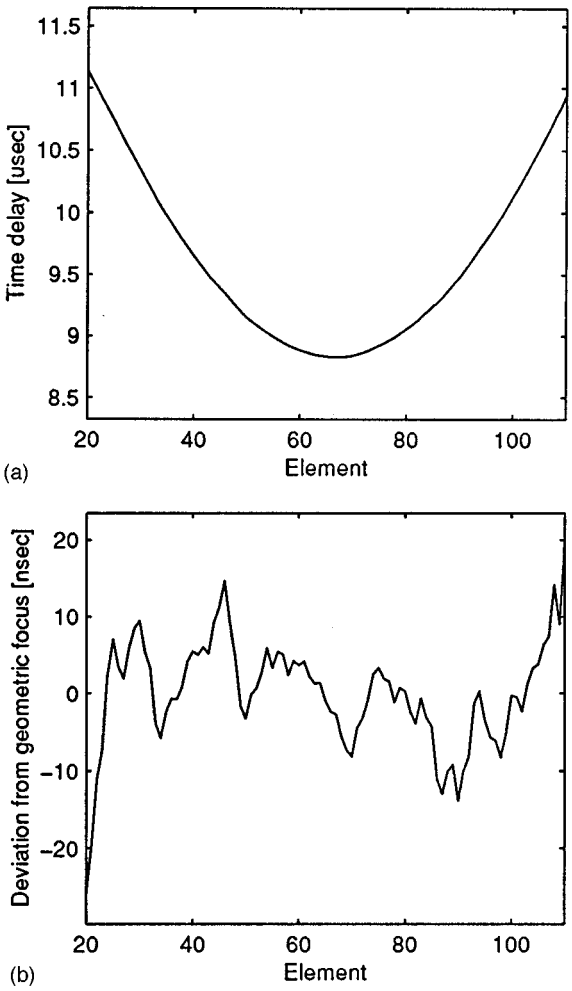


FIG. 6. Arrival time profile of echoes from proximal suspected *in vivo* microcalcification (a) before and (b) after geometric focusing. The aperture used has been limited to those elements with nearest-neighbor cross-correlation coefficients of 0.5 or greater.

prevent estimation of the error in these measurements. These high **k** values suggest that these targets could be easily detected by an ideal observer in uniform tissue. However, the clinical detection problem is complicated by the limitations of the human observer, the nonuniformity of breast tissue, and the presence of other echogenic structures within the breast, among other effects.

The authors are also interested in the potential of subresolution MCs as echogenic point targets to aid in the measurement and characterization of phase aberration in the breast. The authors expect the phase profile of echoes from such targets to reflect the presence of an aberrator. The correlation between the measured profile and the actual aberrator will be limited by system parameters such as noise and aperture geometry as well as assumptions made about the structure of the aberrator. No significant phase error was found in the profiles after geometric focusing for the targets described above. While conclusions regarding aberration in the breast cannot be drawn from this single case, the authors find it interesting that this result differs significantly from other reported breast aberrator measurements which describe rms phase errors on the order of 60 ns.[24,31] However, the experimental procedure described above is also different from those used to make these other aberrator measurements.

As described in the theoretical discussion above, one factor which prevents the reliable visualization of *in vivo* MCs is speckle noise in the detected image. We expect a subresolution MC to appear as a bright speckle. The ability of the clinician to discern such a target from the background speckle noise will be profoundly affected by the degree of brightness compression applied to the B-mode image. In current ultrasound scanners, the dynamic range of the B-mode signal can be compressed, often using a logarithmic function controlled by the operator, to fit within the dynamic range of the display. As MC detection is not the object of typical clinical scanning protocols, it is unlikely the imaging parameters used are optimized in this regard. For example, the targets described above were not visible on the initial scan with the scanner configured with typical logarithmic compression and mean brightness. The targets became visible only after the compression was reduced to its minimum setting and the gain adjusted to return the image to the original mean brightness.

III. CONCLUSION

Medical ultrasound is not currently considered a reliable means to visualize MCs in the breast. While such visualization has been discussed in the literature, closer examination of this issue is necessary in light of the rapid advancement of imaging systems. Improvements in MC visualization would extend the capability of medical ultrasound and be of potential clinical benefit, particularly to the young and/or radiographically challenging patient. The work described in this paper is ultimately directed towards quantifying both the physical factors which we believe currently limit visualization and the relative impact of system design parameters on visualization.

We have begun this examination by posing the task of MC visualization as a detection problem. The methods are based on theory and simulations for analyzing the changes in the Rician statistic which result from changes in aperture geometry or the presence of an aberrator modeled as a thin phase screen. Examples of the application of these methods are presented to show the expected changes in detection performance due to changes in aperture geometry and the presence of an aberrator. These methods can be used to assess the relative performance of ultrasound systems in the detection of subresolution MCs modeled as point targets in an environment of diffuse scatterers. Initial results indicate that a large imaging aperture is best used coherently rather than in the spatial compounding configuration considered. We present observations from the imaging of MCs in excised tissue and suspected MCs *in vivo*, demonstrating the high echogenicity of MCs and their potential to serve as *in vivo* point targets. In order to improve modeling of this imaging task in the interest of improving visualization, considerable experimental and clinical work is still required to characterize the typical scattering properties of MCs and breast tissue.

ACKNOWLEDGMENTS

This work was supported by NIH Grant No. 7R01CA43334 and NSF Engineering Research Center Grant No. CDR-8622201. Technical support for this work was provided by the Ultrasound Group of Siemens Medical Systems, Inc.

APPENDIX

For a given s value the Rician distribution is parametrized by the σ parameter. In comparing the detection performance of different apertures or of the same aperture with different aberrators, it has been assumed that the differences among them can be described solely in terms of scaling these parameters appropriately. The resulting speckle statistics are then used to describe relative performance at detecting a strong coherent scatterer in a volume of weaker diffuse scatterers.

A method is required to calculate the variance of the echo signal using an arbitrary aperture geometry or aperture aberrator. The output of the ultrasound system at the focus can be described in terms of the convolution of the scattering function with the point spread function of the system at the focus. We model the scattering function as a field or volume of random numbers with zero-mean Gaussian amplitude distribution. The PSF has zero mean and is deterministic. For clarity a vector notation to represent locations in space is adopted. Hence the echo signal received from a point at the focus is written as a convolution integral:

$$e(\mathbf{x}) = \int_{\text{All}\,\boldsymbol{\eta}} s(\boldsymbol{\eta}) p(\mathbf{x}-\boldsymbol{\eta})\; d\boldsymbol{\eta}, \tag{A1}$$

where $e(\mathbf{x})$ is the echo signal, $s(\mathbf{x})$ is the scattering function, $p(\mathbf{x})$ is the PSF, and $\mathbf{x}$ is a location in space. For a particular scattering function this will be a constant. To find the variance of the echo signal one must find

$$\sigma_e^2 = \langle (e(\mathbf{x}))^2 \rangle - \langle e(\mathbf{x}) \rangle^2, \tag{A2}$$

where $\langle\ \rangle$ represents the expectation operator over many different scattering functions. Noting that the echo signal has zero mean, the second term can be dropped. We substitute (A1) into the first term:

$$\langle (e(\mathbf{x}))^2 \rangle = \left\langle \left(\int_{\text{All}\,\boldsymbol{\eta}} s(\mathbf{h}) p(\mathbf{x}-\mathbf{h})\; d\mathbf{h} \right)^2 \right\rangle. \tag{A3}$$

Introduce dummy variables to simplify the product of integrals:

$$\begin{aligned} \langle (e(\mathbf{x}))^2 \rangle &= \left\langle \int_{\text{All}\,\boldsymbol{\eta}_1} s(\boldsymbol{\eta}_1) p(\mathbf{x}-\boldsymbol{\eta}_1) d\boldsymbol{\eta}_1 \right. \\ &\quad \left. \times \int_{\text{All}\,\boldsymbol{\eta}_2} s(\boldsymbol{\eta}_2) p(\mathbf{x}-\boldsymbol{\eta}_2) d\boldsymbol{\eta}_2 \right\rangle, \\ &= \left\langle \int_{\text{All}\,\boldsymbol{\eta}_1} \int_{\text{All}\,\boldsymbol{\eta}_2} s(\boldsymbol{\eta}_1) s(\boldsymbol{\eta}_2) p(\mathbf{x}-\boldsymbol{\eta}_1) \right. \\ &\quad \left. \times p(\mathbf{x}-\boldsymbol{\eta}_2) d\boldsymbol{\eta}_1\; d\boldsymbol{\eta}_2 \right\rangle. \end{aligned} \tag{A4}$$

The expectation operator can be moved within the integral:

$$\begin{aligned} \langle (e(\mathbf{x}))^2 \rangle &= \int_{\text{All}\,\boldsymbol{\eta}_1} \int_{\text{All}\,\boldsymbol{\eta}_2} \langle s(\boldsymbol{\eta}_1) s(\boldsymbol{\eta}_2) \rangle \langle p(\mathbf{x}-\boldsymbol{\eta}_1) \\ &\quad \times p(\mathbf{x}-\boldsymbol{\eta}_2) \rangle d\boldsymbol{\eta}_1\; d\boldsymbol{\eta}_2. \end{aligned} \tag{A5}$$

The term $\langle s(\boldsymbol{\eta}_1) s(\boldsymbol{\eta}_2) \rangle$ is recognized as the autocorrelation function. For a scattering function modeled as a Gaussian white random process with variance σ_s^2, this simplifies to a delta function at the origin. Also, the function $p(\mathbf{x})$ is deterministic. This allows the simplification of the integral:

$$\begin{aligned} \langle (e(\mathbf{x}))^2 \rangle &= \int_{\text{All}\,\boldsymbol{\eta}_1} \int_{\text{All}\,\boldsymbol{\eta}_2} \delta(\boldsymbol{\eta}_1 - \boldsymbol{\eta}_2) p(\mathbf{x}-\boldsymbol{\eta}_1) \\ &\quad \times p(\mathbf{x}-\boldsymbol{\eta}_2) d\boldsymbol{\eta}_1\; d\boldsymbol{\eta}_2, \\ &= \sigma_s^2 \int_{\text{All}\,\boldsymbol{\eta}_1} p(\mathbf{x}-\boldsymbol{\eta}_1)^2\; d\boldsymbol{\eta}_1. \end{aligned} \tag{A6}$$

For a given σ_s^2, the difference between the speckle variance of different imaging systems can be described solely in terms of their respective point spread functions. The integral above represents the energy of the PSF, considering the function $p(\mathbf{x})$ as solely real. This can also be determined in the k-space domain by the application of Parseval's theorem:

$$\sigma_e^2 = \sigma_s^2 \int_{-\infty}^{\infty} |P(k)|^2\; dk. \tag{A7}$$

For the ROC analysis comparing aperture geometries, σ_e^2 for each system was both calculated in the k-space domain and estimated by directly integrating the energy of the simulated PSF. These two methods gave results which agree to within one percent. The Rician s parameter in this analysis was also calculated for the $f/2$ control from the simulated PSF by setting s equal to the product of the target strength and the peak amplitude of the detected PSF. In these simulations the target strength was 10. For the ROC analysis with phase aberration, σ_e^2 was estimated by directly integrating the energy of the simulated PSF.

The Fraunhofer approximation states that under certain conditions the lateral and elevational components of the PSF can be approximated by the spatial Fourier transform of the aperture times a quadratic phase term.[32] The transform of the PSF is the system response in k space, which through the application of the Fraunhofer approximation amounts to the convolution of the transmit and receive aperture functions. For the comparison of the geometries, the systems differed only in the lateral dimension, and thus the integration of PSF energy in the k-space domain was reduced to the integral in the lateral dimension only:

$$\sigma_e^2 = \sigma_s^2 \int_{-My}^{\infty} |P(k_x)|^2 dk_x. \tag{A8}$$

The normalized apertures used to represent the different systems are defined as simple rectangle functions, described in arbitrary units of amplitude (A) and space (x), and are shown schematically in Fig. A1(a). The lateral transmit–receive response of these systems in the spatial frequency domain in units of magnitude ($|A|$) and spatial frequency

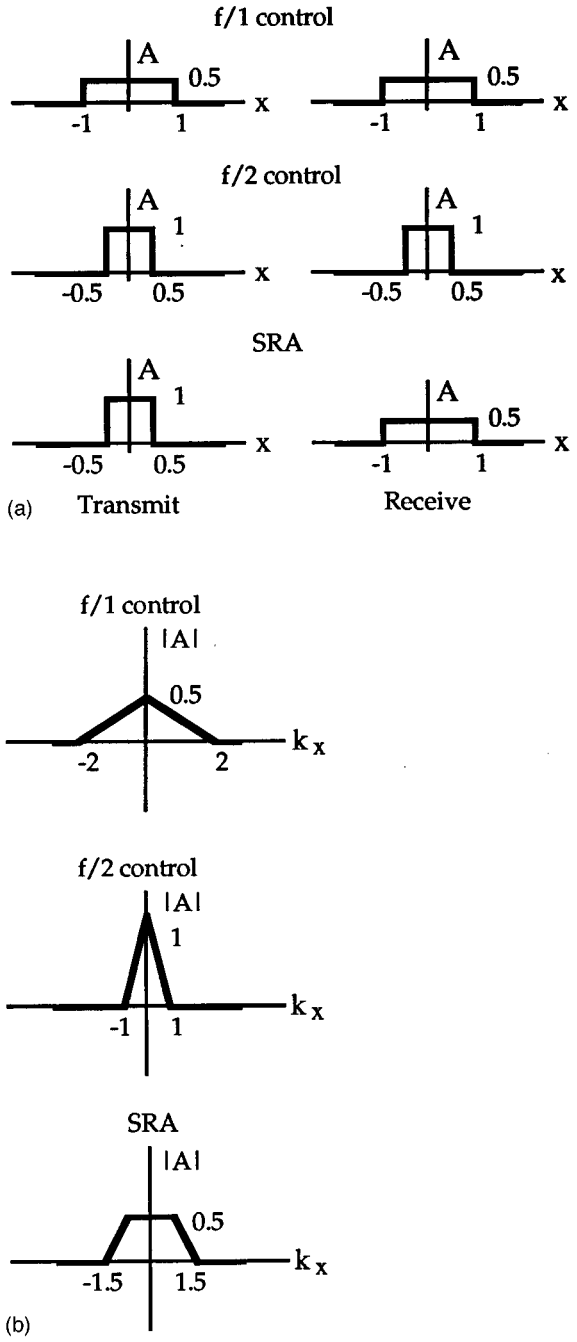


FIG. A1. Aperture functions used to represent imaging systems, shown in arbitrary units of space. (b) Lateral transmit–receive response in the spatial frequency domain, shown in arbitrary units of lateral spatial frequency.

(k_x) are shown in Fig. A1(b). Their respective σ values are shown in Table AI. These σ values were calculated by finding the integral [Eq. (A8)] of each of the respective transmit–receive responses, and normalizing these to the $f/2$ control.

For the spatial compounding case the two speckle patterns which are summed to form an average are statistically independent, and the PDF of their sum equals the convolution of the PDFs of the two images. The PDFs for both subapertures were defined, and these were then convolved. The σ value used for the right and left subapertures was that of the $f/2$ control with a slight correction reflecting the decrease in their effective size due to look angle.

TABLE AI. σ values for simulated systems calculated in the frequency domain.

System	$f/1$ control	$f/2$ control	SRA
σ values	$\sqrt{\frac{1}{3}}$	$\sqrt{\frac{2}{3}}$	$\frac{1}{2}\sqrt{\frac{5}{3}}$
Normalized	$\frac{1}{\sqrt{2}}$	1	$\frac{1}{\sqrt{\frac{8}{5}}}$

The scaling of the σ parameter must also be taken into account. Changing the system aperture size will affect the absolute σ value as the sensitivity of the system is changed. However, this scaling affects the diffuse scatterers equally, such that this sensitivity change does not affect the **k** parameter, which determines detectability. In the aberrated case, the PSF is distorted and its peak is often shifted away from the focus. In order to include such cases in our estimate of detection performance, the authors adopt the peak value of the PSF envelope as an estimate of the scaling of s, regardless of the peak's location. This approach regards a detection successful even if the target visualization is misregistered. Consider the introduction of a point scatterer of strength A at location **x** in Eq. (A1):

$$e(\mathbf{x}) = \int_{\mathrm{All}\,\boldsymbol{\eta}} [s(\boldsymbol{\eta}) + A\,\delta(\boldsymbol{\eta}-\mathbf{x}')]p(\mathbf{x}-\boldsymbol{\eta})\,d\boldsymbol{\eta},$$

$$= Ap(\mathbf{x}-\mathbf{x}') + \int_{\mathrm{All}\,\boldsymbol{\eta}} s(\boldsymbol{\eta})p(\mathbf{x}-\boldsymbol{\eta})\,d\boldsymbol{\eta}. \tag{A9}$$

We wish to maximize $\langle|e(\mathbf{x})|\rangle$ over many realizations of the scattering function. We apply the expectation operator over many realizations of $s(\mathbf{x})$:

$$\langle|e(\mathbf{x})|\rangle = \left\langle \left| Ap(\mathbf{x}-\mathbf{x}') + \int_{\mathrm{All}\,\boldsymbol{\eta}} s(\boldsymbol{\eta})p(x-\boldsymbol{\eta})\,d\boldsymbol{\eta} \right| \right\rangle. \tag{A10}$$

Finally, the triangle inequality is applied. Note that the PSF $p(\mathbf{x})$ is deterministic:

$$\langle|e(\mathbf{x})|\rangle \leqslant A|p(\mathbf{x}-\mathbf{x}')| + \left\langle \left| \int_{\mathrm{All}\,\boldsymbol{\eta}} s(\boldsymbol{\eta})p(\mathbf{x}-\boldsymbol{\eta})\,d\boldsymbol{\eta} \right| \right\rangle. \tag{A11}$$

The second term simplifies to a constant in the expectation, thus $\langle|e(\mathbf{x})|\rangle$ can only be maximized by choosing **x** to coincide with the peak of the envelope of $p(\mathbf{x})$. This supports the choice of the envelope peak as an estimate of the scaling of s.

[1] L. W. Bassett and D. L. Butler, ''Mammography and early breast cancer detection,'' Am. Fam. Physician **43**, 547–557 (1991).
[2] D. M. Hansell, J. C. Cooke, C. A. Parsons, S. H. Evans, D. R. Dance, J. M. Bliss, and I. Ilesley, ''A quantitative analysis of the spatial relationships of grouped microcalcifications demonstrated on xeromammography in benign and malignant breast disease,'' Br. J. Radiol. **61**, 21–25 (1988).
[3] B. de Lafontan, J. P. Daures, B. Salicru, F. Eynius, J. Mihura, P. Rouanet, J. L. Lamarque, A. Naja, and H. Pujol, ''Isolated clustered microcalcifications: diagnostic value of mammography—series of 400 cases with surgical verification,'' Radiology **190**(2), 479–483 (1994).
[4] E. S. de Paredes, P. L. Abbitt, S. Tabbarah, M. A. Bickers, and D. C.

Smith, ''Mammographic and histologic correlations of microcalcifications,'' RadioGraphics **10**(4), 577–589 (1990).
[5] A. Fandos-Morera, M. Prats-Esteve, J. Tura-Soteras, and A. Traveria-Cros, ''Breast tumors: composition of microcalcifications,'' Radiology **169**(2), 325–327 (1988).
[6] V. P. Jackson, R. E. Hendrick, S. A. Feig, and D. B. Kopans, ''Imaging of the radiographically dense breast,'' Radiology **188**, 297–301 (1993).
[7] C. T. M. Brekelmans, H. J. A. Collette, C. Collette, J. Fracheboud, and F. de Waard, ''Breast cancer after a negative screen: follow-up of women participating in the DOM screening programme,'' Eur. J. Cancer **28A**, 893–895 (1992).
[8] I. C. Bennett, R. Freitas, Jr., and I. S. Fentiman, ''Diagnosis of breast cancer in young women,'' Aust. New Zealand J. Surg. **61**, 284–289 (1991).
[9] M. L. Palmer and T. N. Tsangaris, ''Breast biopsy in women 30 years old or less,'' Am. J. Surg. **165**, 708–712 (1993).
[10] I. R. Brand, D. A. Sapherson, and T. S. Brown, ''Breast imaging in women under 35 with symptomatic breast disease,'' Br. J. Radiol. **66**, 394–397 (1993).
[11] V. P. Jackson, ''Sonography of malignant breast disease,'' Semin. Ultrasound, CT, MR **10**(2), 119–131 (1989).
[12] P. B. Guyer and K. C. Dewbury, ''Ultrasound of the breast in the symptomatic and X-ray dense breast,'' Clin. Radiol. **36**, 69–76 (1985).
[13] A. J. Potterton, D. J. Peakman, and J. R. Young, ''Ultrasound demonstration of small breast cancers detected by mammographic screening,'' Clin. Radiol. **49**, 808–813 (1994).
[14] P. B. Guyer, K. C. Dewbury, D. Warwick, J. Smallwood, and I. Taylor, ''Direct contact B-scan ultrasound in the diagnosis of solid breast masses,'' Clin. Radiol. **37**, 451–458 (1986).
[15] F. Kasumi, ''Can microcalcifications located within breast carcinomas be detected by ultrasound imaging?,'' Ultrasound Med. Biol. Suppl. 1 **14**, 175–182 (1988).
[16] L. Filipczynski, ''Detectability of calcifications in breast tissues by the ultrasonic echo method,'' Arch. Acoust. **8**(3), 205–222 (1983).
[17] L. Filipczynski, and G. Lypacewicz, ''Estimation of calcification in breast tissues by means of the ultrasonic echo and shadow methods,'' Arch. Acoust. **9**(1–2), 41–50 (1984).
[18] L. Filipczynski, T. Kujawska, and G. Lypacewicz, ''Ultrasonic echo method in detection breast calcifications. Transient analysis,'' Arch. Acoust. **11**(3), 287–298 (1986).
[19] D. E. Grenoble, J. L. Katz, K. L. Dunn, R. S. Gilmore, and K. L. Murty, ''The elastic properties of hard tissues and apatites,'' J. Biomed. Mater. Res. **6**, 221–233 (1972).
[20] E. Kelly-Fry, S. T. Morris, V. P. Jackson, R. W. Holder, and N. T. Sanghvi, ''Variation of transducer frequency output and receiver band-pass characteristics for improved detection and image characterization of solid breast masses,'' Ultrasound Med. Biol. Suppl. 1 **14**, 143–161 (1988).
[21] P. J. Dempsey, ''The importance of resolution in the clinical application of breast sonography,'' Ultrasound Med. Biol. Suppl. 1 **14**, 43–48 (1988).
[22] V. P. Jackson, E. Kelly-Fry, P. A. Rothschild, H. R. W. , and S. A. Clark, ''Automated breast sonography using a 7.5 MHz PVDF transducer: preliminary clinical evaluation,'' Radiology **159**(3), 679–684 (1986).
[23] L. F. Nock, and G. E. Trahey, ''Synthetic receive aperture imaging with phase correction for motion and for tissue inhomogeneities, part I: basic principles,'' IEEE Trans. Ultrason. Ferroelec. Freq. Contr. **39**(4), 489–495 (1992).
[24] P. D. Freiburger, D. C. Sullivan, B. H. LeBlanc, S. W. Smith, and G. E. Trahey, ''Two-dimensional ultrasonic beam distortion in the breast: *in vivo* measurements and effects,'' Ultrason. Imag. **14**, 398–414 (1992).
[25] P. D. Edmonds, C. L. Mortensen, J. R. Hill, S. K. Holland, J. F. Jensen, P. Schattner, and A. D. Valdes, ''Ultrasound tissue characterization of breast biopsy specimens,'' Ultrason. Imag. **13**, 162–185 (1991).
[26] J. W. Goodman, *Statistical Optics* (Wiley–Interscience, New York, 1985).
[27] J. A. Jensen, and N. B. Svendsen, ''Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers,'' IEEE Trans. Ultrason. Ferroelec. Freq. Control **39**(2), 262–267 (1992).
[28] J. A. Swets, ''ROC analysis applied to the evaluation of medical imaging techniques,'' Invest. Radiol. **<dbxNULL>14**(2), 109–121 (1979).
[29] S. W. Flax, and M. O'Donnell, ''Phase aberration correction using signals from point reflectors and diffuse scatters: basic principles,'' IEEE Trans. Ultrason. Ferroelec. Freq. Control **<dbxNULL>35**(6), 758–767 (1988).
[30] A. J. Rudgers, ''Acoustic pulses scattered by a rigid sphere immersed in a fluid,'' J. Acoust. Soc. Am. **45**(4), 900–910 (1968).
[31] L. M. Hinkelman, D. L. Liu, and R. C. Waag, ''Measurement and correction of ultrasonic pulse distortion produced by the human breast,'' J. Acoust. Soc. Am. **97**, 1958–1969 (1995).
[32] J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, San Francisco, 1986).

Circular concentric Helmholtz resonators[a)]

A. Selamet and P. M. Radavich
Department of Mechanical Engineering, The Ohio State University, Columbus, Ohio 43210

N. S. Dickey
The University of Michigan, Ann Arbor, Michigan 48109

J. M. Novak
Powertrain Operations, Ford Motor Company, Dearborn, Michigan 48121

(Received 27 December 1995; accepted for publication 30 July 1996)

The effect of specific cavity dimensions of circular concentric Helmholtz resonators is investigated theoretically, computationally, and experimentally. Three analytical models are employed in this study: (1) A two-dimensional model developed to account for the nonplanar wave propagation in both the neck and the cavity; (2) a one-dimensional solution developed for the limit of small cavity length-to-diameter ratio, l/d, representing a radial propagation in the cavity; and (3) a one-dimensional closed-form solution for configurations with large l/d ratios which considers purely axial wave propagation in the neck and the cavity. For low and high l/d, the resonance frequencies determined from the two-dimensional approach are shown to match the one-dimensional predictions. For cavity volumes with $l/d>0.1$, the resonance frequencies predicted by combining Ingard's end correction with one-dimensional axial wave propagation are also shown to agree closely with the results of the two-dimensional model. The results from the analytical methods are then compared with the numerical predictions from a three-dimensional boundary element method and with experiments. Finally, these approaches are employed to determine the wave suppression performance of circular Helmholtz resonators in the frequency domain. [S0001-4966(97)05312-5]

PACS numbers: 43.20.Ks, 43.20.Mv, 43.50.Gf [JEG]

INTRODUCTION

Helmholtz resonators, which consist of a volume communicating through an orifice or neck to some external excitation, produce narrow bands of high wave attenuation. The classical approach in modeling these resonators is to neglect the spatial distribution leading to an equivalent spring–mass system where the mass of air in the neck, $m=\rho_0 A_c l_c$, is driven by an external force and the volume acts as a spring with stiffness $s=\rho_0 c_0^2 A_c^2/V$, A_c and l_c being the neck area and length, respectively, and V the resonator volume (Rayleigh, 1945, Art. 303; Kinsler *et al.*, 1982, Chap. 10). For this one degree of freedom system, $\omega_r=\sqrt{s/m}$, leading to a single resonance frequency of $f_r=(c_0/2\pi)\sqrt{A_c/l_c V}$, which is a function of the cavity volume, but independent of the volume dimensions. Experimental observations, however, have deviated from this frequency, which is attributed to the motion of some additional mass on both sides of the neck. To improve the accuracy, the neck length is usually ''corrected'' by adding a term for each end in order to account for this fluid motion, thereby modifying l_c in the foregoing expression for f_r by $l_c' = l_c + \delta_v + \delta_p$. A number of analytical treatments based on somewhat simplified physics have been introduced to develop these end correction factors. Rayleigh (1945) derived a length correction for resonators mounted in a baffle. Ingard (1953) developed an end correction to account for multidimensional wave propagation excited at the area discontinuity from the neck to the volume, but it is only effective for a limited range of geometries when used with the classical model above. In their extensive work, Miles (1971) and Miles and Lee (1975) investigated the Helmholtz resonance behavior of harbors by employing the electrical analogy; the former treats the simple shapes including circular and rectangular harbors with constant depth, the latter develops analytical approaches for irregular geometries with variable depth. Alster (1972) extended the spring-mass analogy by including the mass of the spring and incorporating a spring with varying stiffness. His experimental results for a number of different resonator shapes showed a significant improvement over the lumped model. Several other simple end corrections are also listed by Chuka (1973). To predict the resonance frequency, Tang and Sirignano (1973) assumed one-dimensional wave propagation in both the resonator neck and cavity volume. Panton and Miller (1975) developed a relationship for the resonance frequency which matched experimental results well when used with Ingard's end correction. Their effort was based on a one-dimensional wave motion in the cavity alone and a spatially lumped short neck length. Monkewitz and Nguyen-Vo (1985) studied nonplanar effects in two-dimensional resonators with a semicylindrical cavity and three-dimensional resonators with a hemispherical cavity. By asymptotically matching the solutions of the linearized inviscid equations in terms of low-frequency expansions in the exterior, neck, and cavity, they proposed volume and length corrections for these configurations. Selamet *et al.* (1993, 1995a) developed an expression for the transmission loss of a Helmholtz resonator with wave

[a)]A preliminary version of this study has been presented at the SAE Noise and Vibration Conference as SAE 951263, Traverse City, MI.

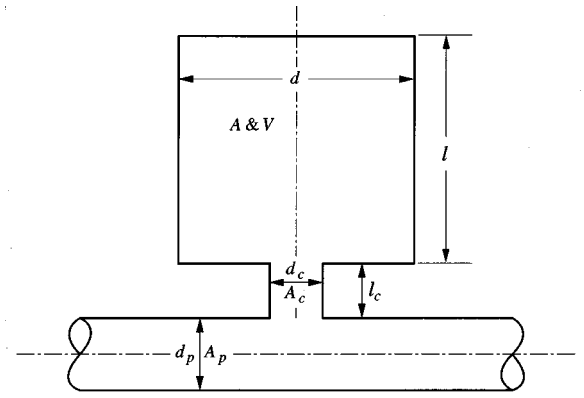


FIG. 1. Helmholtz resonator dimensions (d_p=4.859 cm, d_c=4.044 cm, l_c=8.5 cm, and V=4500 cm^3).

motion. They also provided experimental results illustrating that the relationship for resonance frequency given by Tang and Sirignano (without end correction factors) worked well for volumes with large length-to-diameter ratio, while showing some deviation at low l_v/d_v ratios (hereafter, l/d, for brevity). In a preliminary investigation, Selamet *et al.* (1994a) studied the effect of multidimensional propagation and showed deviations in resonance frequency from the lumped parameter analysis, particularly at low l/d ratios. Recently, Chanaud (1994) developed a relationship for the resonance frequency of configurations with a rectangular parallelepiped cavity volume. Finally, by curve fitting finite element results, Sahasrabudhe *et al.* (1995) provided polynomial expressions for the end correction as a function of expansion ratio and frequency.

The present study considers the circular concentric Helmholtz configurations with constant cavity volume and neck length, as shown in Fig. 1. The objective is to investigate, as a function of the l/d ratio: (1) the discrepancy in the resonance frequency from classical approaches, and (2) the acoustic attenuation behavior. To estimate the end correction accurately at the neck–volume interface, a two-dimensional axisymmetric analytical model is developed for these resonators. The study also employs two different one-dimensional analytical solutions based on (1) radial propagation for small l/d ratios, and (2) axial propagation for large l/d ratios, to be referred hereafter as one-dimensional radial and axial solutions, respectively. Both closed-form solutions are used in the study to examine their corresponding limits. The pressure field inside the resonators is determined by the boundary element method and used to evaluate the degree of multidimensionality of the sound-pressure field at both ends of the resonator neck. Experimental results are obtained by installing Helmholtz resonators in an impedance tube setup. Results from the boundary element method and experiments are then compared with the resonance frequency relationship of the one-dimensional axial model modified by Ingard's end correction for multidimensional propagation. The analytical predictions for the acoustic attenuation of the resonators are then compared with the computations and the experiments for the case of zero mean flow.

Following the Introduction, Sec. I describes the one-dimensional methods, Sec. II the two-dimensional analytical approach, Sec. III the boundary element method, and Sec. IV the experimental setup. The results from the analytical approaches and the boundary element method are compared with experiments and used to evaluate the end corrections in Sec. V. The study is concluded with final remarks in Sec. VI.

I. ONE-DIMENSIONAL ANALYTICAL APPROACHES

The generation and propagation of multidimensional waves in the resonator volume is clearly dependent on the relative magnitudes of the wavelength, the neck and duct diameters and the volume dimensions. Provided that the incident wave from the neck to the volume is planar, there are two limiting configurations for which the wave propagation in the volume can be considered one-dimensional, thereby allowing a relatively simple closed-form solution for transmission loss. This section presents the expressions for the wave attenuation properties for these one-dimensional limits.

A. Radial propagation limit

The discrepancy between the analytical results and the experiment is largest for l/d of the order 0.1 to 1.0. An accurate prediction of the transmission loss and resonance frequency in this region then requires a multidimensional analysis. For smaller l/d ratios as the volume approaches a ''pancake'' geometry, a one-dimensional solution in the radial direction becomes possible as discussed next.

Consider planar propagation in the main duct and connector and purely radial wave motion in the cavity volume. For circularly symmetric propagation of acoustic waves in a hollow disk of constant width, the linearized, inviscid wave equation may be written as

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial p}{\partial r}\right)=\frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2}. \tag{1}$$

Introducing the harmonic dependence

$$p(r,t)=P(r)e^{i\omega t}, \tag{2}$$

and using the separation of variables, the solution for complex pressure amplitude may be determined as

$$P(r)=C_1H_0^{(1)}(kr)+C_2H_0^{(2)}(kr), \tag{3}$$

where C_1 and C_2 are complex constants related to the amplitude of oscillation for the inward and outward traveling waves, respectively, and $H_\beta^{(\eta)}$ is the Hankel function (the Bessel function of the third kind) of order β and type η. As a function of the latter variable, η, Hankel functions may readily be expressed as $H^{(1)}=J+iY$ and $H^{(2)}=J-iY$ with J and Y being the Bessel functions of the first and second kind, respectively. The momentum equation

$$\frac{\partial \mathbf{u}}{\partial t}=-\frac{1}{\rho_0}\nabla p \tag{4}$$

combined with the harmonic relationship

$$u(r,t)=U(r)e^{i\omega t} \tag{5}$$

and Eqs. (2) and (3) yields the complex velocity amplitude as

$$U(r)=-\frac{i}{\rho_0 c_0}[C_1H_1^{(1)}(kr)+C_2H_1^{(2)}(kr)]. \tag{6}$$

At the intersection of the neck and cavity, a constant pressure cylindrical junction is assumed, where the boundary conditions of continuity of pressure and conservation of flow volume are applied. To satisfy these requirements, the pressure and flow volume at $r=r_c$ are matched with those at the cavity end of the neck. Furthermore, the rigid wall requires that $u=0$ at $r=r_v$. Combining the foregoing relationships with those for planar propagation in the main duct and neck yields the transmission loss for a Helmholtz resonator in the low l/d limit as

$$\mathrm{TL}=10\log_{10}\left|1+\frac{A_c}{2A_p}\left[\frac{1+iX\tan kl_c}{X+i\tan kl_c}\right]\right|^2, \tag{7}$$

where

$$X=\frac{Z_v}{\rho_0 c_0}\left(\frac{r_c}{2l_v}\right) \tag{8}$$

is introduced for convenience, and

$$Z_v=\frac{p|_{r=r_c}}{u|_{r=r_c}} =i\rho_0 c_0\left[\frac{H_0^{(1)}(kr_c)-\left(\frac{H_1^{(1)}(kr_0)}{H_1^{(2)}(kr_0)}\right)H_0^{(2)}(kr_c)}{H_1^{(1)}(kr_c)-\left(\frac{H_1^{(1)}(kr_0)}{H_1^{(2)}(kr_0)}\right)H_1^{(2)}(kr_c)}\right] \tag{9}$$

is the volume impedance at $r=r_c$.

B. Axial propagation limit

The effect of nonplanar wave propagation is expected to diminish as the cavity diameter approaches that of the neck, which is equivalent to increasing the l/d ratio of the volume for a specified volume. Considering only one-dimensional propagation in the axial direction in the neck and cavity volume leads to the closed-form relationship

$$\mathrm{TL}=10\log_{10}\left[1+\left(\frac{A_c}{2A_p}\frac{\tan kl_c+(A_v/A_c)\tan kl_v}{1-(A_v/A_c)\tan kl_c\tan kl_v}\right)^2\right] \tag{10}$$

for transmission loss (Selamet *et al.*, 1993, 1995a). The denominator of Eq. (10) provides a relationship for the resonance frequency as

$$\tan kl_c\tan kl_v=\frac{A_c}{A_v}, \tag{11}$$

(the form derived by Tang and Sirignano, 1973) which implies that the frequency is a function of the cavity dimensions.

II. TWO-DIMENSIONAL ANALYTICAL APPROACH

Nonplanar wave propagation in the vicinity of duct discontinuities has been studied by Miles (1944, 1946, 1948). Following his works, the present study considers axisymmetric wave propagation in concentric circular ducts and develops a two-dimensional analytical solution to account for the wave motion in the resonator neck and volume in terms of the resulting pressure waves. For propagation in a circular and concentric configuration, the solution to the linearized, inviscid wave equation

$$\nabla^2 p=\frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2} \tag{12}$$

can be written (Munjal, 1987), in view of $p(r,x,t)=P(r,x)e^{i\omega t}$, as a combination of planar and radial waves as

$$P_A(r,x)=A_0e^{-ikx}+\sum_{n=1}^{\infty}A_nJ_0(\gamma_{j,0n}r)e^{ik_{j,0n}x} \tag{13}$$

for a wave traveling in the positive x direction, and

$$P_B(r,x)=B_0e^{ikx}+\sum_{n=1}^{\infty}B_nJ_0(\gamma_{j,0n}r)e^{-ik_{j,0n}x} \tag{14}$$

for a wave traveling in the negative x direction. Here, P is the complex amplitude of p, J_0 is the Bessel function of the first kind and order zero, $k=\omega/c_0$ is the planar wave number, and $k_{j,0n}$ is the wave number in the x direction given by

$$k_{j,0n}=\sqrt{k^2-\gamma_{j,0n}^2}, \tag{15}$$

where $\gamma_{j,0n}=\alpha_{0n}/r_j$ is the radial wave number in a pipe of radius r_j, with α_{0n} being the roots of the Bessel function $J_0'(\alpha_{0n})=0$. In terms of the momentum equation,

$$\frac{\partial\mathbf{u}}{\partial t}=-\frac{1}{\rho_0}\nabla p, \tag{16}$$

the velocities of these waves can be obtained as

$$U_A(r,x)=\frac{1}{\rho_0 c_0}A_0e^{-ikx} -\frac{1}{\rho_0\omega}\sum_{n=1}^{\infty}A_nk_{j,0n}J_0(\gamma_{j,0n}r)e^{ik_{j,0n}x}, \tag{17}$$

$$U_B(r,x)=-\frac{1}{\rho_0 c_0}B_0e^{ikx} +\frac{1}{\rho_0\omega}\sum_{n=1}^{\infty}B_nk_{j,0n}J_0(\gamma_{j,0n}r)e^{-ik_{j,0n}x}. \tag{18}$$

Driving the resonator via a piston with an oscillating velocity amplitude of U_P allows the interface between the neck and volume to be isolated. For the piston-excited resonator shown in Fig. 2, matching the velocity boundary condition at the piston

$$U_Pe^{i\omega t}=(u_{xA}+u_{xB})|_{x_c=0}, \tag{19}$$

gives $s=0,1,...,\infty$ equations; for $s=0$,

$$U_P=\frac{1}{\rho_0 c_0}(A_0-B_0), \tag{20}$$

and for $s=1,2,...,\infty$,

$$0=A_s-B_s. \tag{21}$$

At the expansion from the neck to the volume, $x_c=l_c$, the pressure boundary condition

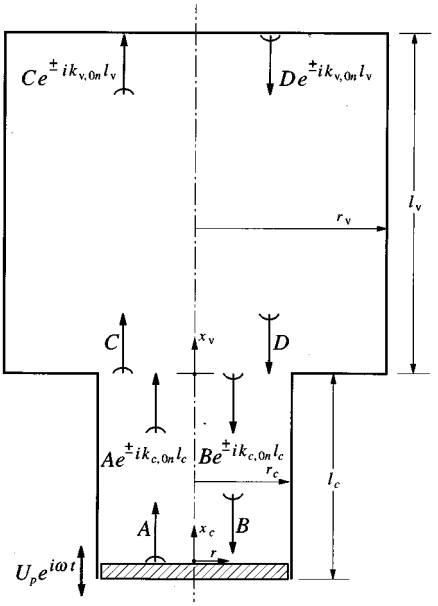


FIG. 2. Geometry for piston-driven Helmholtz resonator (semicircles at the arrow bases represent two-dimensional propagation).

$$(P_A+P_B)|_{x_c=l_c}=(P_C+P_D)|_{x_{v=0}}, \quad \text{for } 0\leqslant r\leqslant r_c \tag{22}$$

(Miles, 1944) gives, for $s=0$,

$$A_0\left(\frac{r_c^2}{2}\right)e^{-ikl_c}+B_0\left(\frac{r_c^2}{2}\right)e^{ikl_c}$$
$$=C_0\left(\frac{r_c^2}{2}\right)+\sum_{n=1}^{\infty}C_n\left[\frac{r_cJ_1(\gamma_{v,0n}r_c)}{\gamma_{v,0n}}\right]+D_0\left(\frac{r_c^2}{2}\right)$$
$$+\sum_{n=1}^{\infty}D_n\left[\frac{r_cJ_1(\gamma_{v,0n}r_c)}{\gamma_{v,0n}}\right], \tag{23}$$

and for $s=1,2,...,\infty$

$$A_se^{ik_{c,0s}l_c}\left[\frac{r_c^2}{2}J_0^2(\gamma_{c,0s}r_c)\right]+B_se^{-ik_{c,0s}l_c}\left[\frac{r_c^2}{2}J_0^2(\gamma_{c,0s}r_c)\right]$$
$$=\sum_{n=1}^{\infty}C_n\left[\frac{\gamma_{v,0n}r_cJ_1(\gamma_{v,0n}r_c)J_0(\gamma_{c,0s}r_c)}{\gamma_{v,0n}^2-\gamma_{c,0s}^2}\right]$$
$$+\sum_{n=1}^{\infty}D_n\left[\frac{\gamma_{v,0n}r_cJ_1(\gamma_{v,0n}r_c)J_0(\gamma_{c,0s}r_c)}{\gamma_{v,0n}^2-\gamma_{c,0s}^2}\right]. \tag{24}$$

Similarly, the velocity boundary conditions

$$(u_{xA}+u_{xB})|_{x_c=l_c}=(u_{xC}+u_{xD})|_{x_{v=0}}, \quad \text{for } 0\leqslant r\leqslant r_c, \tag{25}$$

$$(u_{xC}+u_{xD})|_{x_{v=0}}=0, \quad \text{for } r_c\leqslant r\leqslant r_v, \tag{26}$$

give, for $s=0$,

$$A_0r_c^2e^{-ikl_c}-B_0r_c^2e^{ikl_c}=C_0r_v^2-D_0r_v^2 \tag{27}$$

and for $s=1,2,...,\infty$

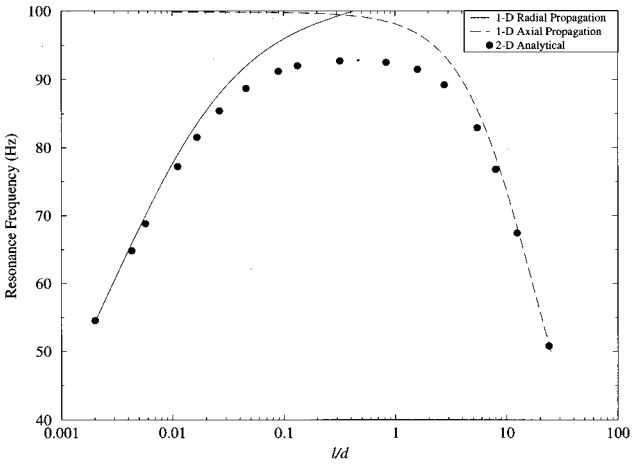


FIG. 3. Resonance frequency versus l/d ratio comparison between axial and radial one-dimensional methods and the two-dimensional analytical approach.

$$kA_0\left[\frac{r_cJ_1(\gamma_{v,0s}r_c)}{\gamma_{v,0s}}\right]e^{-ikl_c}$$
$$-\sum_{n=1}^{\infty}A_nk_{c,0n}\left[\frac{\gamma_{v,0s}r_cJ_0(\gamma_{c,0n}r_c)J_1(\gamma_{v,0s}r_c)}{\gamma_{v,0s}^2-\gamma_{c,0n}^2}\right]e^{ik_{c,0n}l_c}$$
$$-kB_0\left[\frac{r_cJ_1(\gamma_{v,0s}r_c)}{\gamma_{v,0s}}\right]e^{ikl_c}$$
$$+\sum_{n=1}^{\infty}B_nk_{c,0n}\left[\frac{\gamma_{v,0s}r_cJ_0(\gamma_{c,0n}r_c)J_1(\gamma_{v,0s}r_c)}{\gamma_{v,0s}^2-\gamma_{c,0n}^2}\right]e^{-ik_{c,0n}l_c}$$
$$=-k_{v,0s}C_s\left[\frac{r_v^2}{2}J_0^2(\gamma_{v,0s}r_v)\right]+k_{v,0s}D_s\left[\frac{r_v^2}{2}J_0^2(\gamma_{v,0s}r_v)\right]. \tag{28}$$

For the reflection from the rigid wall at the end of the volume, Eq. (20) with $U_p=0$ evaluated at $x_v=l_v$ gives, for $s=0$,

$$0=C_0e^{-ikl_v}-D_0e^{ikl_v}, \tag{29}$$

and for $s=1,2,...,\infty$, Eq. (21) gives

$$0=C_se^{ik_{v,0s}l_v}-D_se^{-ik_{v,0s}l_v}. \tag{30}$$

Equations (20), (21), (23), (24), and (27)–(30) provide a set of simultaneous equations to determine the pressure amplitudes in the neck and volume A_n, B_n, C_n, and D_n ($n=0,1,...$) (Radavich, 1995; Selamet and Radavich, 1995b, 1995c). Higher-order radial terms have a diminishing effect on the solution, thereby allowing the truncation of the infinite series to a finite number of terms sufficient to provide an accurate solution for the pressure variation in the resonator. In this closed system, the resonance occurs when the pressure or the resistance on the piston is a minimum (Kinsler *et al.*, 1982, Chap. 9). This minimum pressure at the piston can be evaluated by substituting A_n and B_n into Eqs. (13) and (14).

Figure 3 provides a comparison of the resonance frequency versus the l/d ratio predicted by the one-dimensional methods and the two-dimensional analytical approach. There is a good agreement between the one-dimensional and the

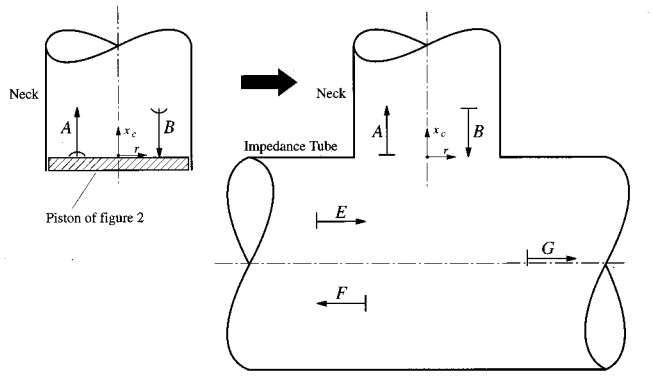


FIG. 4. Geometry for impedance-tube-mounted Helmholtz resonator (semicircles at the arrow bases represent two-dimensional propagation and lines represent planar propagation).

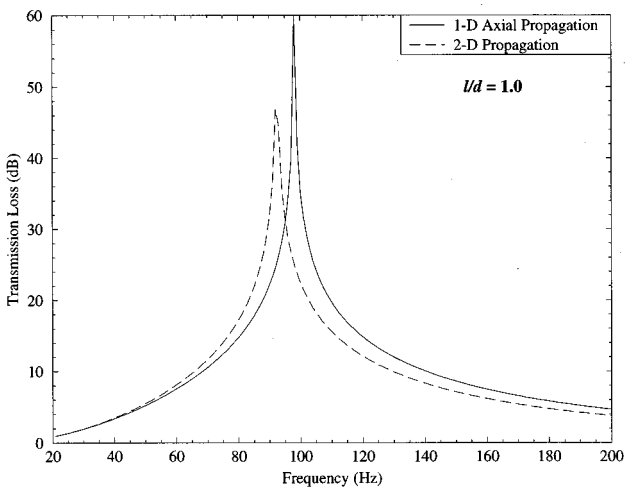


FIG. 5. Transmission loss comparison between one-dimensional axial and two-dimensional approach for Helmholtz resonator with $l/d=1.0$.

two-dimensional methods at the extremes of l/d where the volume is either very short with a large diameter leading to a one-dimensional radial wave or very long and narrow producing an axial planar wave. In between, however, for l/d ratios of approximately 0.1 to 3, hereafter to be referred as the mid-region, there is a considerable difference between the one- and two-dimensional methods with deviations reaching about 8%. The general trend in the resonance frequency at low and high l/d for the present concentric circular configurations is in qualitative agreement with the analysis of Chanaud (1994) on square-faced parallelepiped cavities, as expected.

The transmission loss of the foregoing two-dimensional resonator can be determined on an impedance tube setup as shown in Fig. 4. The complex interface between the impedance tube and the resonator neck where the two circular tubes come together is an obstacle for the development of an exact two- or three-dimensional relationship. Therefore, only one-dimensional waves are assumed to propagate in the impedance tube in order to isolate the multidimensional effects of the expansion. This approximation requires the presence of planar waves at the impedance tube–neck interface as shown in Fig. 4. Here the planar input wave E produces a planar wave A that travels up the resonator neck. At the expansion from the neck to the volume, the sudden area discontinuity excites the two-dimensional waves C and D in the volume as in Fig. 2. The boundary conditions at the neck–cavity interface require a two-dimensional wave B, while the boundary conditions at the neck–impedance tube interface require that B be one-dimensional. Both conditions are satisfied by assuming that B is two-dimensional at the neck–cavity volume interface, and that the radial modes decay sufficiently over the length of the neck before B reaches the impedance tube junction. This assumption is justified for frequencies well below the cutoff frequency for the neck. For the impedance-tube-mounted resonator, the equations for the reflection at the end of the resonator volume, Eqs. (29) and (30), remain unchanged. For the expansion from the resonator neck into the volume, Eqs. (23), (24), (27), and (28) are used with A_n $(n=1,2,...,\infty)$ set to zero, while retaining only the planar A_0 term. At the impedance tube interface, the pressure boundary condition gives

$$(A_0+B_0)=(E+F)=G \tag{31}$$

and the velocity boundary condition gives

$$(E-F)r_p^2=(A_0-B_0)r_c^2+Gr_p^2 \tag{32}$$

leading to a set of simultaneous equations if the input magnitude E is specified. The transmitted wave G can then be determined allowing for the calculation of the transmission loss across the resonator as

$$\mathrm{TL}=20\log_{10}\left|\frac{E}{G}\right|. \tag{33}$$

For a typical geometry within the mid-region of Fig. 3, for example $l/d=1.0$, Fig. 5 provides a transmission loss comparison between the axial one-dimensional method of Eq. (10) and the two-dimensional approach. The magnitude of deviation observed in this figure is large enough to lead to the misprediction of the primary transmission loss due to the narrow attenuation band of the Helmholtz resonator.

This analytical study illustrates the significance of the multidimensional effects in the vicinity of neck–volume interface. The remainder of the work concentrates then on the multidimensional physics in view of the three-dimensional

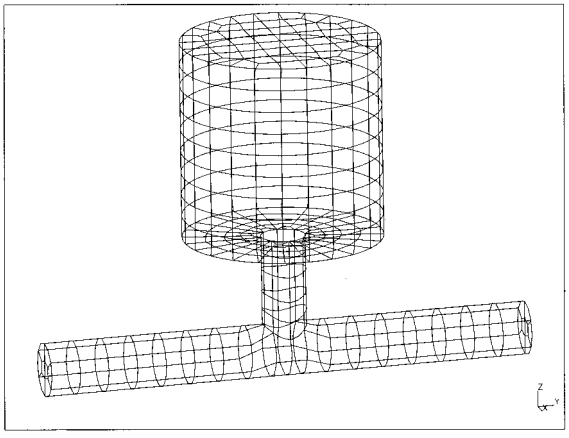


FIG. 6. Sample boundary element mesh for $l/d=1.0$.

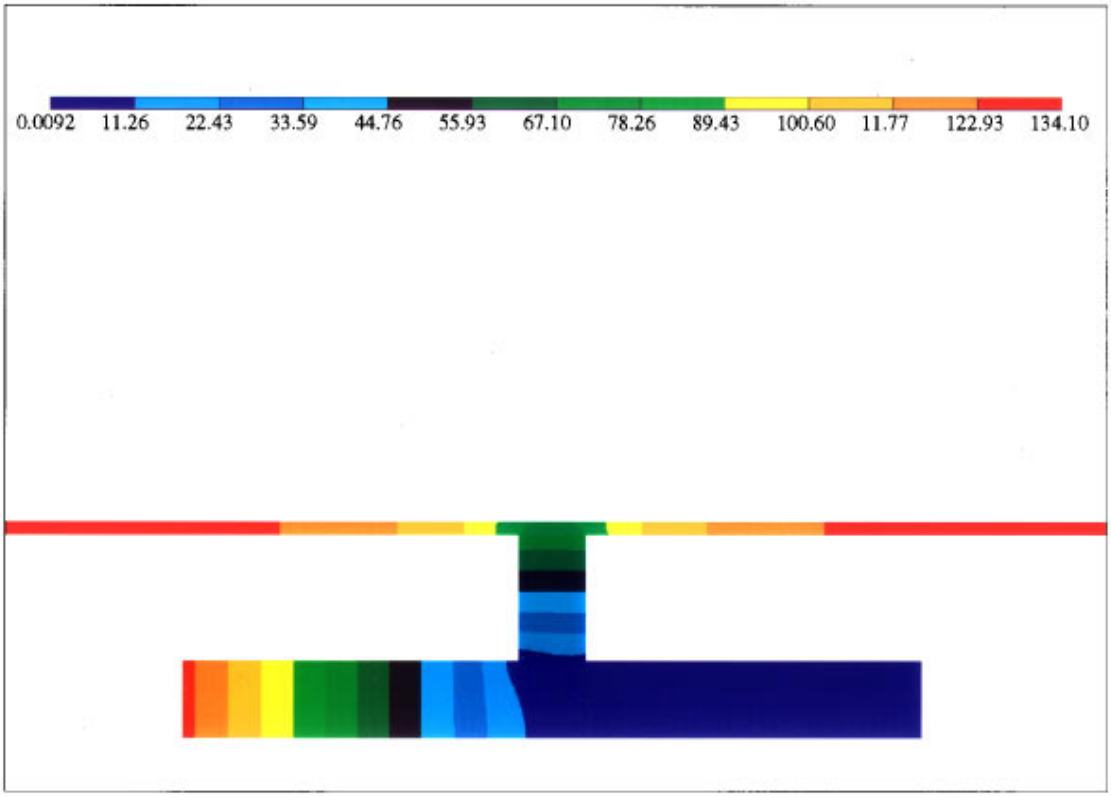


FIG. 7. Pressure magnitude (Pa) contours for Helmholtz resonator with $l/d=0.01$.

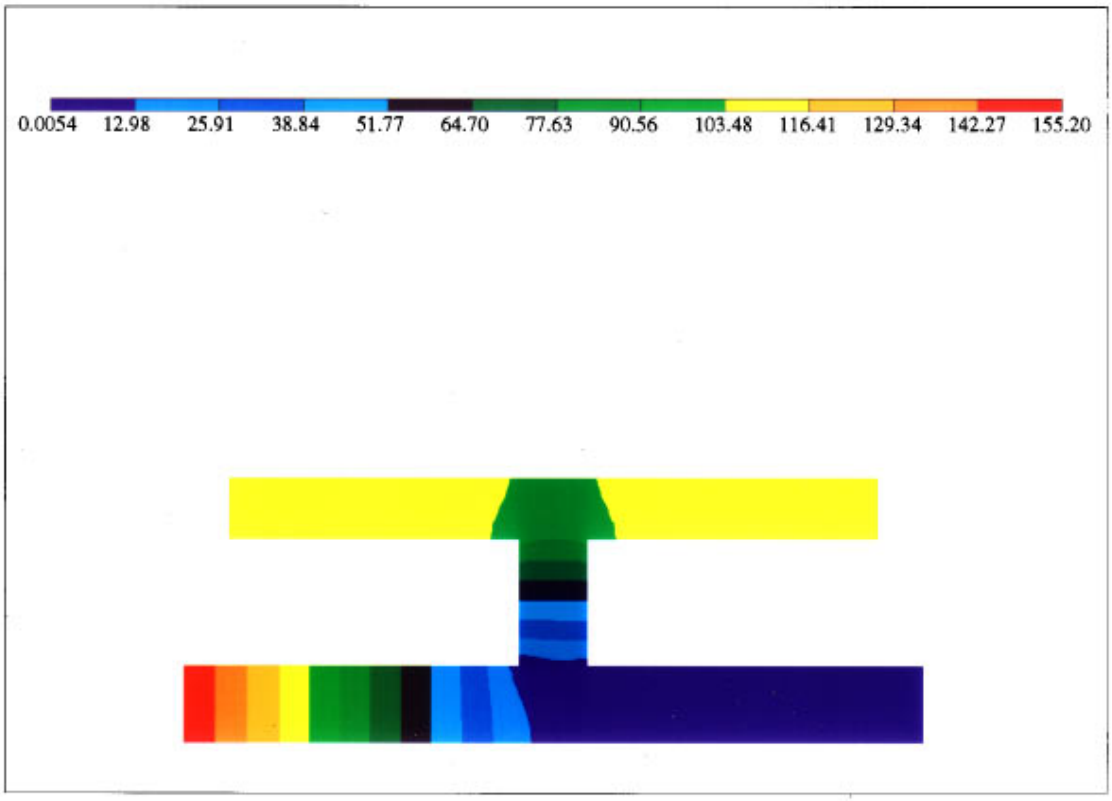


FIG. 8. Pressure magnitude (Pa) contours for Helmholtz resonator with $l/d=0.1$.

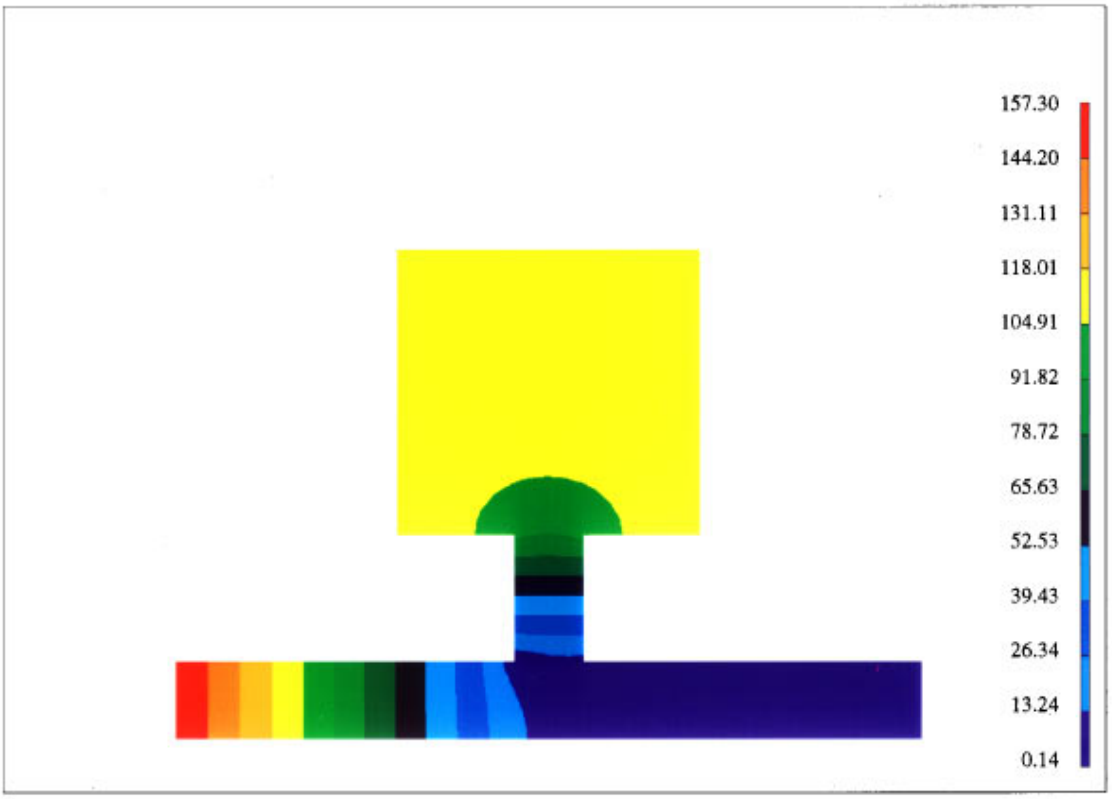


FIG. 9. Pressure magnitude (Pa) contours for Helmholtz resonator with $l/d=1.0$.

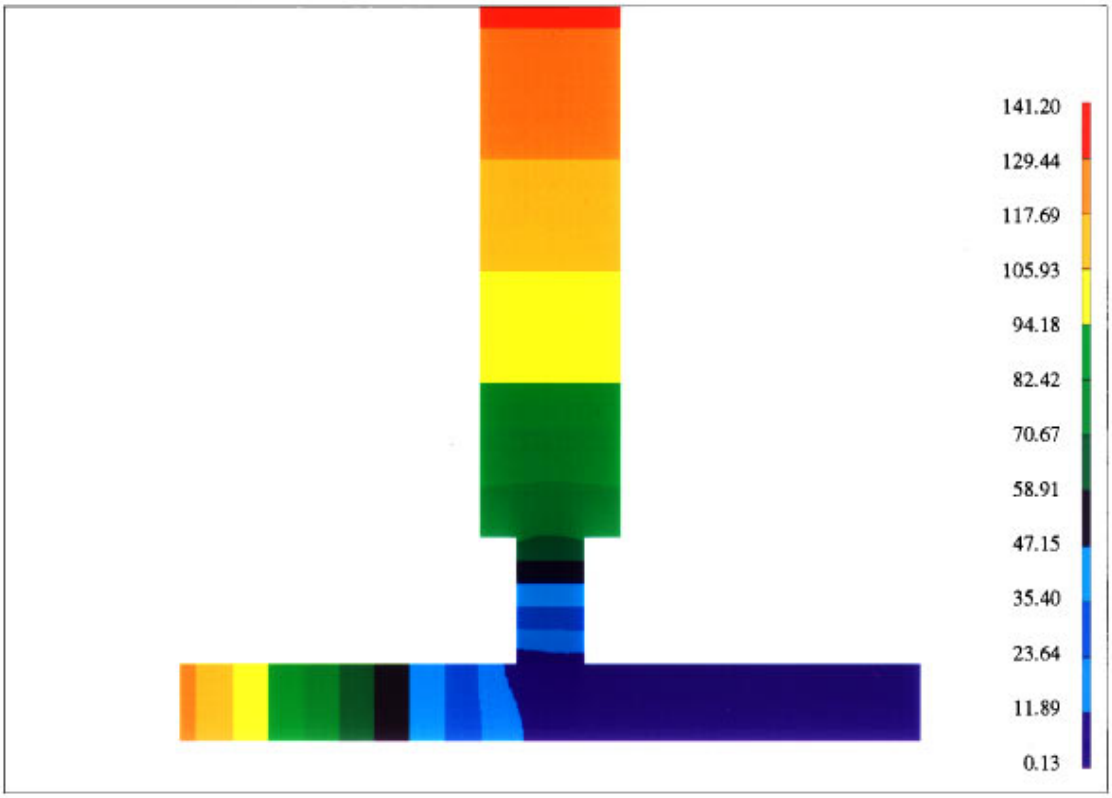


FIG. 10. Pressure magnitude (Pa) contours for Helmholtz resonator with $l/d=10$.

computations and experiments conducted in an extended impedance tube facility. The next section provides a brief description of the three-dimensional computational approach based on the direct boundary element method.

III. BOUNDARY ELEMENT METHOD

To analyze the multidimensional effects at the neck-to-volume area transition further, a three-dimensional direct boundary element method is used. A detailed account of this method, which is based on the linearized, inviscid wave equation [see Eq. (12)], can be found in numerous sources (Rayleigh, 1945, Art. 293; Seybert *et al.*, 1985; Soenarko and Seybert, 1991). The boundary element method was used to model the experimental apparatus: An oscillating velocity was input; an anechoic termination was implemented by setting the impedance of the termination equal to the characteristic impedance of the fluid, $\rho_0 c_0$; and the two-microphone technique was used to calculate the transmission loss. For the present investigation, the boundary element method was implemented using isoparametric quadrilateral and triangular elements. To ensure accuracy, a fine mesh spacing of less than 2.5 cm was maintained for all models. The largest mesh size of 2978 nodes occurred for a small cavity l/d of 0.01, which greatly increases the surface area to be discretized. As a result of this large surface area, no configuration below l/d of 0.01 was modeled with the boundary element method. A sample mesh for an l/d of 1.0 is shown in Fig. 6.

In order to investigate the differences between the one- and two-dimensional methods and determine the extent of nonplanar wave propagation particularly in the mid-region of l/d, pressure contours on the symmetry plane of the configurations were examined at their respective resonance frequencies using the boundary element method for l/d ratios of 0.01, 0.1, 1.0, and 10, as shown in Figs. 7 through 10. For all four configurations, some nonplanar wave bending is observed at the junction between the neck and impedance tube. Focusing on the interaction between the neck and the volume, at the low $l/d=0.01$, Fig. 7 ($f_r=77$ Hz) shows clearly one-dimensional radial propagation in the volume. Figure 8 ($f_r=90$ Hz) illustrates the contours at a lesser extreme of the mid-region in Fig. 3 with $l/d=0.1$. As the l/d ratio is increased to the $l/d=1.0$ configuration of Fig. 9 ($f_r=91$ Hz), multidimensional wave propagation becomes evident at the area transition between the neck and the volume. For the high $l/d=10$ case in Fig. 10 ($f_r=72$ Hz), some nonplanar bending is observed at the neck–volume interface, but the overall propagation is mostly planar. The extreme radial and axial configurations of $l/d=0.01$ and 10 exhibit a marked pressure variation over the length of the volume. The two mid-region configurations of $l/d=0.1$ and 1.0, however, show a rather small pressure variation in the volume. Thus the configurations in this region may be approximated by a lumped volume approach, provided the physics at the transitions is incorporated accurately. Examining the area transition for the $l/d=1.0$ and $l/d=10$ cases in Figs. 9 and 10 reveals that the multidimensional wave propagation is more pronounced for the larger area transition, as expected. In the one-dimensional axial model, it is assumed that the waves in the volume are planar immediately after the area transition and that the pressure and velocity suddenly change over this discontinuity when in reality there is a portion of the fluid in the volume that moves with the fluid in the neck. This effect has typically been accounted for in one-dimensional axial theory by adding a length correction to the neck which will be discussed in detail following a brief description of the experimental setup.

IV. EXPERIMENTS

The experimental apparatus consists of an extended impedance tube configuration, where the Helmholtz resonators are placed between a broad-frequency noise source and an anechoic termination. The two-microphone technique (Chung and Blaser, 1980; ASTM, 1990) is utilized to separate incident and reflected waves for calculation of the transmission loss across the element, with one pair of microphones placed before and another pair after the resonator. Although multidimensional waves are excited in the resonator volume, the selected impedance tube diameter ensures planar propagation at the microphones, with a cutoff frequency above 4 kHz for nonsymmetric modes. For further details of the experimental setup, refer to Selamet *et al.* (1994b). Eight Helmholtz resonators with circular concentric neck and cavity volumes, as described in Table I, were fabricated for the experimental study. The impedance tube diameter, the neck length and diameter, and the resonator volume (d_p, l_c, d_c, and V) are fixed for all resonators, while the length-to-diameter ratio, l/d, is varied. No resonator was built below $l/d=0.32$ configuration due to fabrication difficulties for the set of parameters employed in the study.

V. RESULTS AND DISCUSSION

Numerous corrections for both ends of the resonator neck proposed by a number of investigators lack universal applicability since the expressions are derived (1) usually to match one set of experimental data; and (2) to be used in the classical lumped approach, which is known to deviate from experiments as the l/d ratio of the volume increases, due to the neglect of wave motion effects in the neck and cavity volume (Selamet *et al.*, 1995a). Many of these end corrections also fail to incorporate the effect of expansion ratio from the resonator neck to the cavity, which are observed clearly in Fig. 3, and instead involve only the neck dimensions. An end correction that accounts for nonplanar wave propagation effects between the neck and volume was suggested by Ingard (1953) who modeled the neck as a piston oscillating into an expanded pipe of infinite length. His result for the end correction δ may be expressed in the dimensionless form as

$$\frac{\delta}{d_v}=2\sum_{n=1}^{\infty}\frac{1}{\alpha_{0n}^3}\left[\frac{J_1(\alpha_{0n}d_c/d_v)}{J_0(\alpha_{0n})}\right]^2, \tag{34}$$

where J_0 and J_1 are the Bessel functions of the first kind, $\alpha_{0n}=\gamma_{0n}r_v$, $J_1(\alpha_{0n})=0$, and $l_c'=l_c+\delta$. Equation (34) may also be approximated by a simple linear expression

$$\delta\approx 0.85\left(\frac{d_c}{2}\right)\left(1-1.25\,\frac{d_c}{d_v}\right) \tag{35}$$

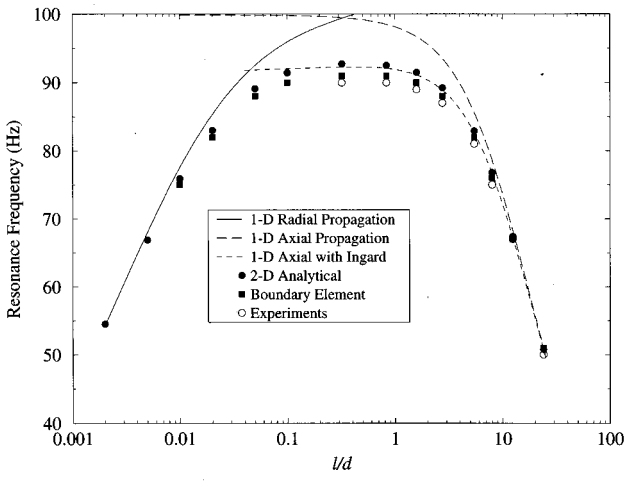


FIG. 11. Resonance frequency versus l/d ratio comparison among the axial and radial one-dimensional methods, the two-dimensional analytical approach, the boundary element method, and experiments.

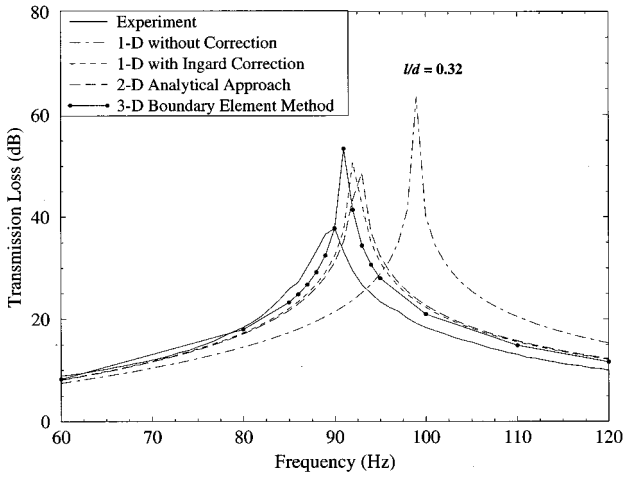


FIG. 12. Transmission loss comparison between experiment, corrected and uncorrected 1-D Eq. (10), 2-D analytical approach, and boundary element method for Helmholtz resonator $l/d=0.32$.

for $d_c/d_v<0.4$, which reduces to $\delta=0.85(d_c/2)$ as d_c/d_v becomes negligible. Recall the end corrections suggested by Rayleigh for a pipe opening with an infinite flange (King, 1936; Rayleigh, 1945, Art. 307; Miles, 1948) as

$$\frac{\pi}{4}\left(\frac{d_c}{2}\right)<\delta<\frac{8}{3\pi}\left(\frac{d_c}{2}\right), \tag{36}$$

or $0.78540<\delta/(d_c/2)<0.84883$. Thus Eq. (34) approaches Rayleigh's upper limit as $d_c/d_v\to 0$, as illustrated by Eq. (35). By assuming a quadratic velocity profile over the opening, Rayleigh improved this correction further as $\delta/(d_c/2)<0.82422$. Later, King (1936) determined the correction more accurately as $\delta/(d_c/2)\cong 0.82132$. At low l/d, the expansion is large and the correction from Eq. (34) is at a maximum, which reduces the resonance frequency of the classical approach and yields corrected one-dimensional predictions closer to the two-dimensional results. With increasing l/d ratio, however, the correction approaches zero and the resonance frequency increases, which contradicts the results depicted in Fig. 3. Although Eq. (34) involves nonplanar wave propagation at the neck–cavity interface, its neglect of one-dimensional axial propagation leads to a discrepancy as the l/d ratio is increased. Combination of Eq. (34) with Eq. (11) allows for one-dimensional axial wave propagation throughout the resonator and approximates the multidimensional physics at the area contraction. This corrected one-dimensional axial and the two-dimensional analytical approaches are compared in Fig. 11. For l/d ratios greater than about 0.1, the two methods agree well. At l/d ratios less than about 0.1, the two-dimensional model predicts a decrease in the resonance frequency as the volume begins to resonate radially (the one-dimensional radial propagation begins to dominate), whereas the one-dimensional axial predictions remain nearly constant.

Experimental results as well as the boundary element method predictions for the configurations listed in Table I are also included in Fig. 11. For the geometries investigated, the boundary element method provides a slight improvement over the other methods. This is attributed to the ability of the boundary element method to incorporate the complicated geometry and multidimensional wave propagation at the junction between the impedance tube and the neck which was left uncorrected for the one- and two-dimensional analysis. Thus the three-dimensional computations take the so-called ''radiation impedance'' effect into account inherently on both the cavity and the impedance tube sides of the neck.

For l/d ratios of 0.32 and 1.59 in the mid-region of Fig. 11, Figs. 12 and 13 show significant improvements in the attenuation predictions of the one-dimensional axial method when the correction factor is used. For a large l/d ratio of 23.92, which shows only a small deviation between the one-dimensional approach and experiments in Fig. 11, Fig. 14 illustrates the diminishing effect of the correction Eq. (34) with a decreased expansion ratio. These three figures also exhibit only minor differences between the approximate two-dimensional and the corrected one-dimensional approaches, as expected in view of the fact that all three use similar assumptions for l/d ratios greater than about 0.1. The three-

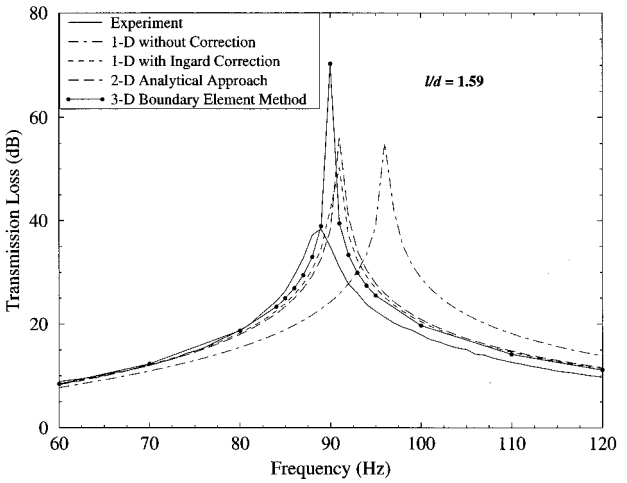


FIG. 13. Transmission loss comparison between experiment, corrected and uncorrected 1-D Eq. (10), 2-D analytical approach, and boundary element method for Helmholtz resonator $l/d=1.59$.

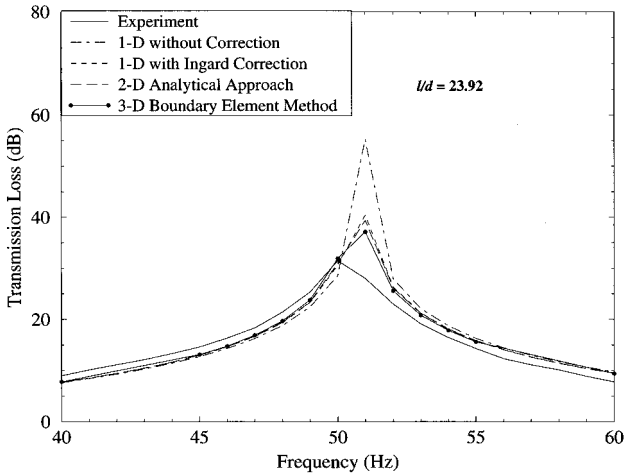


FIG. 14. Transmission loss comparison between experiment, corrected and uncorrected 1-D Eq. (10), 2-D analytical approach, and boundary element method for Helmholtz resonator $l/d=23.92$.

dimensional method in these figures gives slightly improved results for the reasons already indicated. In Figs. 11–14, the difference between the measured and computed (3-D BEM) resonance frequencies is a mere 1 Hz, which may possibly be attributed to rather minor structural vibrations and other losses in the experimental setup. The magnitude difference in the transmission loss between the analytical predictions and the experiments at resonance may be attributed to the neglect of viscothermal losses in the analytical treatments. The effect of viscous dissipation is expected to increase at either extreme of the l/d ratio, where the distance between the walls may become comparable to the boundary layer thickness of the oscillating air.

VI. CONCLUDING REMARKS

This study has shown the effect of both planar and nonplanar wave propagation on the resonance frequency and the wave attenuation of concentric Helmholtz resonators. The one-dimensional models including wave propagation in radial or axial directions and the experiments demonstrate that the resonance frequency changes as the cavity dimensions are varied. Full three-dimensional computations with the boundary element method illustrate that the deviations between the one-dimensional approaches and the experiments for cavities with l/d ratios of about 0.1 to 3 are mostly due to nonplanar wave propagation at the area discontinuities. A two-dimensional analytical approach was then introduced to examine the multidimensional wave propagation at the area discontinuity from the neck to the cavity volume for an axisymmetric resonator. The results for the transmission loss and the resonance frequency from the two-dimensional technique are found to agree well with those of the one-dimensional axial approach modified with Ingard's end correction for length-to-diameter ratios greater than 0.1. These results deviate slightly from the experiments, as expected, because of the difference in the neck–tube interface between the two configurations.

The effect of mean flow across the orifice is currently being investigated, which cannot only alter the resonance frequencies but, under the proper flow conditions, can turn the resonator into a noise generator rather than a silencer. Experimental work by Panton (1990) demonstrates that this effect is heavily dependent on the geometry of the orifice, as well as the nature of the boundary layer as it passes over the resonator. An experimental study combined with 3-D computations is also under progress for the low l/d range (less than 0.1) in combination with the effect of shear on the wave propagation in the same range.

TABLE I. Helmholtz resonator geometry ($d_p=4.859$ cm, $d_c=4.044$ cm, and $l_c=8.5$ cm).

Resonator	l_v (cm)	d_v (cm)	l_v/d_v or l/d	d_v/d_c
1	8.423	26.081	0.32	6.45
2	15.865	19.004	0.83	4.70
3	24.420	15.319	1.59	3.79
4	35.281	12.743	2.77	3.15
5	55.550	10.155	5.47	2.51
6	71.653	8.941	8.01	2.21
7	96.012	7.727	12.43	1.91
8	148.565	6.210	23.92	1.54

ACKNOWLEDGMENT

Support for this work by Ford Motor Company is gratefully acknowledged.

Alster, M. (**1972**). ''Improved calculation of resonant frequencies of Helmholtz resonators,'' J. Sound Vib. **24**(1), 63–85.

ASTM (**1990**). ASTM E 1050-1990, ''Standard test method for impedance and absorption of acoustical materials using a tube, two microphones and a digital frequency analysis system'' (American Society for Testing and Materials, Philadelphia, PA).

Chanaud, R. C. (**1994**). ''Effects of geometry on the resonance frequency of Helmholtz resonators,'' J. Sound Vib. **178**(3), 337–348.

Chuka, O. R. (**1973**). ''Methods for predicting the influence of intense sound, steady flow, and elevated temperature on the acoustic performance of automotive exhaust system components,'' Ph.D. thesis, University of Missouri-Rolla.

Chung, J. Y., and Blaser, D. A. (**1980**). ''Transfer function method of measuring in-duct acoustic properties: I Theory,'' J. Acoust. Soc. Am. **68**, 907–913.

Ingard, U. (**1953**). ''On the theory and design of acoustic resonators,'' J. Acoust. Soc. Am. **25**, 1037–1061.

King, L. V. (**1936**). ''On the electrical and acoustic conductivities of cylindrical tubes bounded by infinite flanges,'' Philos. Mag. **7**(21), 128–144.

Kinsler, L. E., Frey, A. R., Coppens, A. B., and Sanders, J. V. (**1982**). *Fundamentals of Acoustics* (Wiley, New York).

Miles, J. (**1944**). ''The reflection of sound due to a change in cross section of a circular tube,'' J. Acoust. Soc. Am. **16**, 14–19.

Miles, J. W. (**1946**). ''The analysis of plane discontinuities in cylindrical tubes, Parts I & II,'' J. Acoust. Soc. Am. **17**, 259–271, 272–284.

Miles, J. W. (**1948**). ''The coupling of a cylindrical tube to a half-infinite space,'' J. Acoust. Soc. Am. **20**, 652–664.

Miles, J. W. (**1971**). ''Resonant response of harbours: an equivalent-circuit analysis,'' J. Fluid Mech. **46**(2), 241–265.

Miles, J. W., and Lee, Y. K. (**1975**). ''Helmholtz resonance of harbours,'' J. Fluid Mech. **67**(3), 445–464.

Monkewitz, P. A., and Nguyen-Vo, N. M. (**1985**). ''The response of Helmholtz resonators to external excitation. Part 1. Single resonators,'' J. Fluid Mech. **151**, 477–497.

Munjal, M. L. (**1987**). *Acoustics of Ducts and Mufflers* (Wiley, New York).

Panton, R. L. (**1990**). ''Effect of orifice geometry on Helmholtz resonator excitation by grazing flow,'' AIAA J. **28**(1), 60–65.

Panton, R. L., and Miller, J. M. (**1975**). ''Resonant frequencies of cylindrical Helmholtz resonators,'' J. Acoust. Soc. Am. **57**, 1533–1535.

Radavich, P. M. (**1995**). ''The effects of multidimensional wave propagation on the acoustic performance of expansion chambers and Helmholtz resonators,'' Master's thesis, University of Michigan, Ann Arbor.

Rayleigh, J. W. S. (**1945**). *The Theory of Sound, Volume II* (Dover, New York), Art. 293.

Sahasrabudhe, A. D., Munjal, M. L., and Anantha Ramu, S. (**1995**). ''Analysis of inertance due to the higher order mode effects in a sudden area discontinuity,'' J. Sound Vib. **185**(3), 515–529.

Selamet, A., Dickey, N. S., and Novak, J. M. (**1993**). ''Computational simulation of Helmholtz resonators: lumped versus distributed volume,'' Noise Con '93, pp. 247–252.

Selamet, A., Dickey, N. S., Radavich, P. M., and Novak, J. M. (**1994a**). ''Theoretical, computational and experimental investigation of Helmholtz resonators: one-dimensional versus multi-dimensional approach,'' SAE 940612.

Selamet, A., Dickey, N. S., and Novak, J. M. (**1994b**). ''The Herschel–Quincke tube: a theoretical, computational, and experimental investigation,'' J. Acoust. Soc. Am. **96**, 3177–3185.

Selamet, A., Dickey, N. S., and Novak, J. M. (**1995a**). ''Theoretical, computational, and experimental investigation of Helmholtz resonators with fixed volume: lumped versus distributed analysis,'' J. Sound Vib. **187**(2), 358–367.

Selamet, A., and Radavich, P. M. (**1995b**). ''The effect of length on the acoustic attenuation performance of concentric expansion chambers: an analytical, computational, and experimental investigation,'' SAE 950544.

Selamet, A., and Radavich, P. M. (**1995c**). ''Helmholtz resonator: A multidimensional analytical, computational, and experimental study,'' SAE 951263.

Seybert, A. F., Soenarko, B., Rizzo, F. J., and Shippy, D. J. (**1985**). ''An advanced computational method for radiation and scattering of acoustic waves in three dimensions,'' J. Acoust. Soc. Am. **77**, 362–368.

Soenarko, B., and Seybert, A. F. (**1991**). ''Recent developments of the boundary element method to noise control problems in automotive engineering,'' *Proceedings of the 6th International Pacific Conference on Automotive Engineering, Seoul, South Korea*, pp. 1339–1346, SAE 912608.

Tang, P. K., and Sirignano, W. A. (**1973**). ''Theory of a generalized Helmholtz resonator,'' J. Sound Vib. **26**(2), 247–262.

The two-component representation of time-harmonic elastic body waves in the high- and intermediate-frequency regimes

Larissa Ju. Fradkin
School of Electrical, Electronic and Information Engineering, South Bank University, 103 Borough Road, London SE1 0AA, United Kingdom

Aleksei P. Kiselev
Institute of Mechanical Engineering, 61 Bol'shoi Prospekt, Vasilievskii Ostrov, St. Petersburg, 199178, Russia

(Received 27 December 1995; accepted for publication 8 August 1996)

Propagation and diffraction of time-harmonic elastic body waves through homogeneous and isotropic materials is revisited. The two-component representation of body waves is given which is applicable in the high- and intermediate-frequency regimes when the family of rays is (i) regular or (ii) possesses an irregularity due to the presence of a wedge, an acoustic point source located on the boundary, simple caustic, or focal line. Both physical and mathematical description of the phenomena under consideration is offered. The resulting expressions are asymptotic in character and easy to compute. They should prove useful in producing algorithms for describing acoustic fields in homogeneous and isotropic materials with isolated defects and in providing benchmark cases for testing numerical codes designed to solve elastic wave equations in more complicated situations. [S0001-4966(97)04212-4]

PACS numbers: 43.20.Bi, 43.20.Dk, 43.20.Gp, 43.35.Zc [JEG]

INTRODUCTION

In this paper we use modern ray theory to describe propagation and diffraction of time-harmonic waves in homogeneous and isotropic solids. The principles of modern ray theory have been first applied to elastic waves by Babič (1956) and then Karal and Keller (1959), Alekseev *et al.* (1961), Babich and Buldyrev (1991) and others. They allow us to generalize the concepts of geometrical acoustics, such as a plane wave, to provide a relatively simple description of diffraction effects. Unfortunately, the subject is still obscure to a nonspecialist, and in the majority of applications, diffraction effects are modeled numerically, by solving elastic wave equations or else by computing Kirchhoff's integrals. However, the resulting codes may be extremely costly to run and, unless they are tested against appropriate benchmarks, are not necessarily reliable. In view of this, it is our intention to compile an album of relevant benchmark cases and supplement each one of them with detailed physical and mathematical description. We believe that this should make the ray theory available to a wider audience. Our first selection presented below has been inspired by our interest in an asymptotic description of the near field of a circular transducer. The exposition should prove useful in mathematical modeling of nondestructive testing of homogeneous and isotropic media with isolated defects.

Throughout, we denote by $\mathbf{u}(x,t)$ the vector field of solid displacement which satisfies the wave equation

$$\partial_{tt}\mathbf{u}-a^2\nabla^2\mathbf{u}+b^2\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{u})=0, \tag{1}$$

with $\nabla^2\mathbf{u}\equiv\partial_{xx}\mathbf{u}+\partial_{yy}\mathbf{u}+\partial_{zz}\mathbf{u}$, t the time variable, x, y, and z the Cartesian coordinates of an arbitrary point $\mathbf{x}$, a and b characteristic wave speeds, and $a>b$. We work in *relatively high*, that is, *both* high and intermediate, frequency regimes, where frequency ω of a time-harmonic solution is such that the ratio of wavelength λ to a characteristic length scale l_o is of order unity or smaller. The characteristic length is problem dependent and sometimes may be difficult to find. The relevant mathematics lies outside the scope of this paper. We nondimensionalize all spatial variables by l_o and all temporal ones by combinations of l_o with relevant characteristic speeds. Thus, *all* quantities we refer to throughout this paper have been rendered *dimensionless*.

It is well known (e.g., Fung, 1965; Hudson, 1980) that any linear elastic body wave described by Eq. (1) can be decomposed into two modes, $\mathbf{u}^P$ and $\mathbf{u}^S$, one propagating with speed a and another, b. Following the accepted terminology we refer to these modes as P for primary and S for secondary, respectively. Mode conversion can only take place on a boundary or interface.

It is now well understood that, in the high-frequency regimes, solutions of Eq. (1) describing diffraction effects may be represented in the form of asymptotic series. The simplest such representation involving inverse powers of (dimensionless) frequency is known as *ray series* (Alekseev *et al.*, 1961; Babič, 1956; Karal and Keller, 1959; Červený *et al.*, 1977; Hudson, 1980). The zeroth-order term of this asymptotic series describes a regular family of *rays* along which wave propagation takes place. (The family of rays or ray field is called regular if only one ray passes through each point.) More generally, the zeroth-order term behaves in accordance with the principles of geometrical acoustics (referred to by some diffractionists as *crude* geometrical acoustics). These may be summarized as follows. In the high-frequency regime, in stationary media:

(i) waves and the energy they carry propagate with speed independent of frequency along the *rays* (lines of fast-

est propagation) which coincide with normals to the *wavefronts* (moving surfaces along which a waveform feature is being simultaneously received);

(ii) the amplitude of a wave attenuates along the ray according to the energy conservation law, so that the same amount of energy passes through each cross section of a ray tube (a tube containing all neighboring rays); and

(iii) in homogeneous and isotropic solids polarization of elastic *P* and *S* waves (that is, direction of their vector of displacement) is along and transverse to the ray, respectively.

In view of (iii) *P* and *S* waves are often referred to as longitudinal and transverse, or else compressional and shear, respectively. This is misleading, since usually polarization of the second term in the ray series is different to the first. For inhomogeneous and isotropic solids explicit expressions for the orthogonal components of the second terms have been obtained by Alekseev *et al.* (1961). Hudson (1980) and Kiselev (1983) presented these expressions in a simpler form. Tsvankin *et al.* (1983), Kiselev and Kashtan (1988), Babich and Kiselev (1989), Kiselev and Tsvankin (1989), and Kiselev and Roslov (1991) all argue that the above results provide excellent quantitative and qualitative explanation of the effect of ''anomalous'' [that is, contrary to principle (iii) above] polarization which had been observed both in seismological field data (Li *et al.*, 1987) and numerical experiments (Alekseev and Mikhailenko, 1982; Daley and Hron, 1987; Hron and Mikhailenko, 1981; Kiselev and Tsvankin, 1989). They show that anomalous polarization is a diffraction effect caused by curvature of wave fronts, amplitude variation along wave fronts, or inhomogeneities in the medium. Using the terminology introduced by Fock (1965), it can be seen as due to the resulting *transverse diffusion* of the displacement.

The ray fields obtained by using the ray representation may contain *irregularities* (''boundaries''), that is, point, line, or surface loci where different rays meet. By analogy with fluid dynamics the regions surrounding these irregularities are referred to as *boundary layers*. In the high-frequency limit the boundary layers become infinitesimally thin. In the theory of diffraction the concept of a boundary layer was first introduced by Keller (1958) and Buchal and Keller (1960) (see also Babič and Kirpičnikova, 1975). Ray asymptotics do not provide a good description of the displacement amplitude inside boundary layers. Occasionally, other types of solutions of Eq. (1) may be found and represented in the form of asymptotic series which involve *special functions* (and not just inverse powers) of wave numbers $k=\omega/a$ and $\varkappa=\omega/b$. Well-known examples of boundary layers are penumbra and the neighborhoods of critical rays, caustics, and focal lines.

Below we rederive known approximate expressions describing relatively high-frequency displacements in homogeneous and isotropic solids. Having restricted ourselves to homogeneous and isotropic media, we offer an extremely simple and unified approach to evaluating these expressions via elastic potentials. Some of our derivations are original. In a sense, all solutions are devised to reproduce features of a plane wave. What results is a two-term representation of displacement fields, where the two terms in question are always orthogonal to one another. The second terms are obtained by simply differentiating the first. As a rule, these representations are different to the full two-term expansions (obtaining which is often a daunting task), since the second terms lack the components parallel to the first. To emphasize this we call the resulting representation *two component* rather than two term. For some regular ray fields the two-component representation had been justified in previous research by comparison with the exact solution (e.g., Daley and Hron, 1987; Yanovskaya and Roslov, 1987; Kiselev and Kashtan, 1988; Kiselev and Tsvankin, 1989; Kiselev and Roslov, 1991). It had been also demonstrated that in regular ray fields the inclusion of the leading order orthogonal term may extend the region of validity of the one-term approximation to the intermediate-frequency regime. Since inside the boundary-layer regions diffraction effects are stronger than around regular rays, the corresponding orthogonal terms play an even bigger role.

The paper is organized as follows: in Sec. II we derive a two-component representation of the solid displacement in regions where ray fields are regular and therefore asymptotic series of ray type exist. General *P* and *S* waves, as well as examples of waves with plane, cylindrical and spherical fronts are considered. Particular attention is paid to the so-called head waves whose fronts are conical. In Sec. III boundary-layer regions are introduced surrounding various ray irregularities. The examples addressed are penumbra behind a wedge, neighborhoods of critical rays, simple caustics, and focal lines.

I. THE TWO-COMPONENT RAY REPRESENTATION

The ray representation generalizes the concept of a plane wave $\mathbf{U}\exp i(\mathbf{kx}-\omega\mathbf{t})$ by allowing the wavefront to deviate from plane, the wave amplitude to vary along the front, or the medium to contain inhomogeneities. The latter are not discussed in this paper. Let us describe the general scheme.

We assume that the ray representation is valid, that is, there is a time-harmonic solution of (1) in the form

$$\mathbf{u}^\alpha(x,t)=\mathbf{u}^\alpha(\mathbf{x})e^{-i\omega t}\equiv\mathbf{U}^\alpha(\mathbf{x},k)e^{-i\omega[t-\tau^\alpha(x)]}, \tag{2}$$

and the frequency ω is high and the amplitude $\mathbf{U}^\alpha$ varies slowly. More precisely, we assume that there exists a large parameter $k\gg 1$ and

$$\mathbf{U}^\alpha(\mathbf{x},k)=\sum_{n=0}^{\infty}\frac{\mathbf{U}_n^\alpha(\mathbf{x})}{(-ik)^{n+\nu}}=\frac{1}{(-ik)^\nu}\left[\mathbf{U}_0^\alpha+\frac{\mathbf{U}_1^\alpha}{-ik}+O(k^{-2})\right] \tag{3}$$

is an asymptotic series in inverse powers of k. Here $\alpha=P, S$, the underlying l_o is he characteristic scale of variation in $\mathbf{U}^\alpha$, the wave number $k=\omega/a$, ν is a real constant, and $\mathbf{U}_n^\alpha$ and $\tau^\alpha(x)$ are all independent of k (see Alekseev *et al.*, 1961; Červený *et al.*, 1977; Kravtsov and Orlov, 1990, Sec. 2.1.3). The assumption that k is large often allows us to simplify the wave equation: As will be seen below, in the presence of terms containing the unknown function, the terms containing its first derivatives divided by k may be omitted; or else in

the presence of terms containing the first derivatives, the terms containing the second derivatives divided by k may be dropped.

Let us substitute (2) and (3) into (1). Then by equating the coefficients of successive powers of k to zero we obtain the so-called eikonal equations for τ^α and transport equations for $\mathbf{U}_n^\alpha$. The eikonal equations are

$$(\nabla\tau^\alpha)^2=(c^\alpha)^{-2}, \tag{4}$$

where $c^P=a$ and $c^S=b$. This implies that for each α we have

$$\nabla\tau^\alpha\equiv\frac{d\tau^\alpha}{ds}\,\mathbf{s}=(c^\alpha)^{-1}\mathbf{s}, \tag{5}$$

with $\mathbf{s}(\mathbf{x})$ the unit vector tangential to a path of fastest variation in τ^α, and $d\tau^\alpha/ds$ the directional derivative. In general, solving the eikonal and transport equations is not a trivial task. A standard approach to this problem is described in Smirnov (1964) and Babič and Buldyrev (1991). The problems surrounding the choice of initial conditions lie outside the scope of this paper. The smooth level surfaces $\tau^\alpha(\mathbf{x})=$const, orthogonal to the paths of fastest variation in τ^α are known as wavefronts. The variation in τ^α along a level surface is obviously zero. For this reason, the paths of fastest variation provide extremals of the Fermat functional $\int_{M_0}^{M}(c^\alpha)^{-1}\,ds$ which is defined on paths connecting any two points M_0 and M. (Here s is the distance along a path.) The extremals are known as rays (Smirnov, 1964; Babič and Buldyrev, 1991; Kravtsov and Orlov, 1990), and $\mathbf{s}$ is sometimes referred to as a unit ray vector. As usual, each eikonal $\tau(\mathbf{x})$ may be interpreted as *the propagation time* from M_0 to M along the ray. It follows that in homogeneous and isotropic media rays are straight lines. We proceed by considering several important examples.

A. The two-component ray representation of a *P* wave

As is well known (e.g., Fung, 1965; Hudson, 1980), in a homogeneous and isotropic solid the primary displacement may be represented in the form

$$\mathbf{u}^P(x)=\nabla\phi(\mathbf{x}), \tag{6}$$

where ϕ is a scalar potential which satisfies the Helmholtz equation,

$$(\nabla^2+k^2)\,\phi=0, \tag{7}$$

and $k=\omega/a$. Similarly to (2), we seek a ray representation of ϕ,

$$\phi(\mathbf{x},k)=\Phi(\mathbf{x},k)e^{i\omega\tau^P(x)}. \tag{8}$$

This means that we assume that frequency ω is high and amplitude Φ varies slowly or that there exists a large dimensionless parameter $k\gg1$ (with the underlying l_0 the characteristic scale of variation in Φ) and

$$\Phi(\mathbf{x},k)=\sum_{n=0}^{\infty}\frac{\Phi_n(\mathbf{x})}{(-ik)^{n+\nu}}=\frac{1}{(-ik)^\nu}\left[\Phi_0+\frac{\Phi_1}{-ik}+O(k^{-2})\right], \tag{9}$$

is an asymptotic series. Here ν is a real constant.

Substituting (8) and (9) into the vector Helmholtz equation (7) and equating to zero, the coefficients of successive powers of k we obtain the eikonal equation for τ^P,

$$(\nabla\tau^P)^2=a^{-2}, \tag{10}$$

and the transport equation for Φ_0,

$$2(\nabla\Phi_0\cdot\mathbf{s})+\Phi_0\nabla\cdot\mathbf{s}=0. \tag{11}$$

Problems concerning transport equations for other Φ_n's lie outside the scope of this paper. Assuming for simplicity that $\nu=1$, substituting (8) and (9) into (6) and taking into account (2) and (3) we obtain

$$\mathbf{u}^P=\left[\mathbf{U}_0^P+\frac{\mathbf{U}_1^P}{-ik}+O(k^{-2})\right]e^{i\omega\tau^P}, \tag{12}$$

where the zeroth-order term is

$$\mathbf{U}_0^P=-\Phi_0\mathbf{s} \tag{13}$$

and the first-order term is

$$\mathbf{U}_1^P=\nabla\Phi_0-\Phi_1\mathbf{s}. \tag{14}$$

Obviously, we have

$$\mathbf{U}_0^P\|\mathbf{s} \tag{15}$$

and

$$\mathbf{U}_1^P=\mathbf{U}_\|^P+\mathbf{U}_\perp^P, \tag{16}$$

where $\mathbf{U}_\|^P(\mathbf{x})$ and $\mathbf{U}_\perp^P(\mathbf{x})$ are the components of $\mathbf{U}_1^P(\mathbf{x})$ which are tangential and transverse to the ray, respectively. The tangential component $\mathbf{U}_\|^P$ cannot be calculated so easily, and when it provides a small correction to the leading term is not of much interest. On the other hand, the transverse component $\mathbf{U}_\perp^P$ describes anomalous polarization and may be evaluated by using (12)–(15). Indeed, (6) implies

$$\mathbf{U}_\perp^P=\nabla_\perp\Phi_0=-\nabla_\perp(\mathbf{U}_0^P\cdot\mathbf{s}), \tag{17}$$

with $\nabla_\perp=\nabla-\mathbf{s}(\mathbf{s}\cdot\nabla)$ a component of the gradient which is transverse to the ray.

Thus, under the high-frequency/slowly varying amplitude assumption, the preferential direction (of fastest variation in primary displacement) is along the ray and the primary displacement $\mathbf{u}^P$ may be well approximated by its two-component representation

$$\tilde{\mathbf{u}}^P=\left[\mathbf{U}_0^P+\frac{\mathbf{U}_\perp^P}{-ik}\right]e^{i\omega\tau^P}. \tag{18}$$

This representation contains the leading terms of both polarizations and the second term is obtained by differentiating the first. Examples in which (18) favorably compares with the exact solution have been considered in Kiselev and Roslov (1991).

B. The two-component ray representation of an *S* wave

Analogously, the secondary displacement propagating in a homogeneous and isotropic solid may be represented as

$$\mathbf{u}^S(\mathbf{x}) = \nabla \times \boldsymbol{\psi}(\mathbf{x}), \tag{19}$$

where $\boldsymbol{\psi}$ is a vector potential which satisfies the vector Helmholtz equation,

$$(\nabla^2 + \varkappa^2)\boldsymbol{\psi} = 0, \tag{20}$$

with the wave number $\varkappa = \omega/b$. Similarly to (2), we seek a ray representation of $\boldsymbol{\psi}$,

$$\boldsymbol{\psi}(\mathbf{x},\omega) = \boldsymbol{\Psi}(\mathbf{x},\omega)e^{i\omega\tau^S(\mathbf{x})}. \tag{21}$$

This means that we assume that frequency ω is high and amplitude $\boldsymbol{\Psi}$ varies slowly, that is, there exists a large dimensionless parameter $\varkappa \gg 1$ (with the underlying l_0 being the characteristic scale of variation in $\boldsymbol{\Psi}$) and

$$\boldsymbol{\Psi}(\mathbf{x},\varkappa) = \sum_{n=0}^{\infty} \frac{\boldsymbol{\Psi}_n(\mathbf{x})}{(-i\varkappa)^{n+\nu}} = \frac{1}{(-i\varkappa)^{\nu}}\left[\boldsymbol{\Psi}_0 + \frac{\boldsymbol{\Psi}_1}{-i\varkappa} + O(\varkappa^{-2})\right], \tag{22}$$

is an asymptotic series. Here ν is a real constant. Therefore, the eikonal equation is

$$(\nabla\tau^S)^2 = b^{-2}, \tag{23}$$

and the transport equation for $\boldsymbol{\Psi}_0$ is

$$2(\mathbf{s}\cdot\nabla)\boldsymbol{\Psi}_0 + \boldsymbol{\Psi}_0(\nabla\cdot\mathbf{s}) = 0. \tag{24}$$

Assuming for simplicity that $\nu = 1$, substituting (21) and (22) into (19) and taking into account (2) and (3), we obtain

$$\mathbf{u}^S = \left[\mathbf{U}_0^S + \frac{\mathbf{U}_1^S}{-i\varkappa} + O(\varkappa^{-2})\right]e^{i\omega\tau^S}. \tag{25}$$

where the zeroth-order term is

$$\mathbf{U}_0^S = \boldsymbol{\Psi}_0 \times \mathbf{s} \tag{26}$$

and the first-order term is

$$\mathbf{U}_1^S = \nabla \times \boldsymbol{\Psi}_0 + \boldsymbol{\Psi}_1 \times \mathbf{s}. \tag{27}$$

We have

$$\mathbf{U}_0^S \perp \mathbf{s} \tag{28}$$

and

$$\mathbf{U}_1^S = \mathbf{U}_{\parallel}^S + \mathbf{U}_{\perp}^S, \tag{29}$$

where $\mathbf{U}_{\parallel}^S(\mathbf{x})$ and $\mathbf{U}_{\perp}^S(\mathbf{x})$ are the components of $\mathbf{U}_1^S(\mathbf{x})$ which are tangential and transverse to the ray, respectively. The transverse component $\mathbf{U}_{\perp}^P$ cannot be calculated so easily and when it provides a small correction to the leading term is not of much interest. On the other hand, the tangential component $\mathbf{U}_{\parallel}^P$ describes anomalous polarization and may be evaluated without difficulty. Indeed, (27) implies

$$\mathbf{U}_{\parallel}^S \equiv (\mathbf{U}_1^S\cdot\mathbf{s})\mathbf{s} = ([\nabla\times\boldsymbol{\Psi}_0]\cdot\mathbf{s})\mathbf{s} = (\nabla_{\perp}\cdot\mathbf{U}_0^S)\mathbf{s}, \tag{30}$$

where $\nabla_{\perp} \equiv [\nabla - \mathbf{s}(\mathbf{s}\cdot\nabla)]$ is the divergence in variables transverse to the ray and we have taken into account (26), relationships

$$([\nabla\times\mathbf{v}]\cdot\mathbf{w}) = (\nabla\cdot[\mathbf{v}\times\mathbf{w}]) + (\mathbf{v}\cdot[\nabla\times\mathbf{w}]), \tag{31}$$

$$\nabla\times\mathbf{s} = b\nabla\times\nabla\tau^S \equiv 0, \tag{32}$$

and (28). Above, $\mathbf{v}$ and $\mathbf{w}$ are arbitrary smooth vector functions.

Thus, under the high-frequency/slowly varying amplitude assumption, the preferential direction (of fastest variation in secondary displacement) is along the ray and the secondary displacement $\mathbf{u}^S$ may be well approximated by the two-component representation

$$\tilde{\mathbf{u}}^S = \left[\mathbf{U}_0^S + \frac{\mathbf{U}_{\parallel}^S}{-i\varkappa}\right]e^{i\omega\tau^S}. \tag{33}$$

It contains the leading terms of both polarizations and the second term is obtained by differentiating the first via (30). Examples in which this approximation favorably compares with the exact solution have been considered in Daley and Hron (1987), Yanovskaya and Roslov (1987), Kiselev and Kashtan (1988), Kiselev and Tsvankin (1989), and Kiselev and Roslov (1991).

C. Anomalous polarization of body waves with simple fronts

When rays may be introduced, the geometry of the displacement field may be conveniently described by using the so-called *ray coordinates* (Babič and Buldyrev, 1991; Červený *et al.*, 1977). When dealing with a regular 3D ray field, three such coordinates are needed: p and q to define a ray, and s to define position of a point on the ray. The Jacobian of transformation from the Cartesian to ray coordinates,

$$J(\mathbf{x}) \equiv \left|\frac{\partial(x,y,z)}{\partial(p,q,s)}\right|, \tag{34}$$

bears the name of *geometrical spreading*. This is due to the fact that a ray tube, that is, a collection of rays

$$p' \leqslant p \leqslant p'', \quad q' \leqslant q \leqslant q'' \tag{35}$$

has a cross section of infinitesimal area $J\,dp\,dq$. Thus, the larger J becomes the wider the rays spread. The zeroth-order potentials, solutions of the zeroth-order transport equations (11) and (24), attenuate along the ray according to

$$\Phi_0(\mathbf{x}) = \left[\frac{J(\mathbf{x}_o)}{J(\mathbf{x})}\right]^{1/2}\Phi_0(\mathbf{x}_o), \tag{36}$$

or

$$\boldsymbol{\Psi}_0(\mathbf{x}) = \left[\frac{J(\mathbf{x}_o)}{J(\mathbf{x})}\right]^{1/2}\boldsymbol{\Psi}_0(\mathbf{x}_o), \tag{37}$$

respectively, where $\mathbf{x}_o$ is an arbitrary fixed initial point (Babič and Buldyrev, 1991). The attenuation formulas (36) and (37) are in agreement with the energy principle (ii) described in the Introduction.

To obtain the 2-D version of the above we just drop y and q. Let us consider several important examples.

1. Waves with plane fronts

Let us introduce the Cartesian coordinates (x,y,z) and assume that a plane elastic P waves propagates in the z direction. Let its scalar potential be $\Phi(\mathbf{x})\exp(ikz)$.

For the plane wave, ray coordinates may be chosen to coincide with the Cartesian coordinates, so that x and y define a ray and z position on the ray. Then the *unit ray vector* is

$$\mathbf{s}=\mathbf{e}_z. \tag{38}$$

In this case we have

$$J/J_o=1, \tag{39}$$

where the subscript o denotes quantities evaluated on an initial wavefront $z=z_o$. Using (13) and attenuation formula (36) we obtain

$$\mathbf{U}_0^P=-\Phi_0(x,y)\mathbf{e}_z, \tag{40}$$

where $(-ik)^{-1}\Phi_0(x,y)$ is the scalar potential prescribed on $z=z_o$. For classical plane waves Φ_0 is constant. Then (17) gives

$$\mathbf{U}_\perp^P=\left(\frac{\partial\Phi_0}{\partial x}\,\mathbf{e}_x+\frac{\partial\Phi_0}{\partial y}\,\mathbf{e}_y\right). \tag{41}$$

The situation with S waves is similar. Let a vector potential be $\Psi(\mathbf{x})\exp(i\varkappa z)\mathbf{s}_\perp$, with $\mathbf{s}_\perp=s_1\mathbf{e}_x+s_2\mathbf{e}_y$ and s_1 and s_2 arbitrary constants. The unit ray vectors are again (38), and using (26) and attenuation formula (37) we have

$$\mathbf{U}_0^S=\Psi_0(x,y)\mathbf{s}_\perp', \tag{42}$$

with $(-i\varkappa)^{-1}\Psi_0(x,y)$ the amplitude of the vector potential prescribed on $z=z_o$ and $\mathbf{s}_\perp'=s_2\mathbf{e}_x-s_1\mathbf{e}_y$. Then (30) gives

$$\mathbf{U}_\parallel^S=\left(s_2\,\frac{\partial\Psi_0}{\partial x}-s_1\,\frac{\partial\Psi_0}{\partial y}\right)\mathbf{e}_z. \tag{43}$$

For both P- and S-plane waves the correction terms (41) and (43) are due to variations in the displacement amplitude along the wavefront. Direct application of (6) and (19) shows that for both P- and S-plane waves, the respective two-component representations (18) and (33) are exact.

2. Waves with cylindrical fronts

Let us introduce cylindrical coordinates (ρ,θ,z) and consider an elastic P wave with concentric cylindrical wavefronts and the axis running along the z axis. Let its scalar potential be $\Phi(\theta)\exp(ik\rho)$.

For the cylindrical wave ray coordinates may be chosen to coincide with the cylindrical coordinates, so that the distance along the z axis and angle θ define a ray and polar radius ρ position on the ray. The unit ray vectors are

$$\mathbf{s}=\mathbf{e}_\rho=(\cos\,\theta,\ \sin\,\theta,\ 0). \tag{44}$$

In this case we have

$$J/J_o=\rho/\rho_o, \tag{45}$$

where the subscript o denotes quantities evaluated on an initial wavefront $\rho=\rho_o$. By using (13) and the attenuation formula (36) we obtain

$$\mathbf{U}_0^P=-\frac{\Phi_0(\theta)}{\rho^{1/2}}\,\mathbf{e}_\rho, \tag{46}$$

where $(-ik)^{-1}\rho_o^{-1/2}\Phi_0(\theta)$ is the scalar potential prescribed on $\rho=\rho_o$. Then (17) gives

$$\mathbf{U}_\perp^P=\frac{1}{\rho^{3/2}}\,\frac{\partial\Phi_0}{\partial\theta}\,\mathbf{e}_\theta, \tag{47}$$

with $\mathbf{e}_\theta=\mathbf{s}_\perp=(-\sin\,\theta,\cos\,\theta,0)$. This means that the correction term decreases away from the source as $\rho^{-3/2}$.

The situation with S waves is similar. Let a vector potential be $\Psi(\mathbf{x})\exp(i\text{æ}\rho)\mathbf{e}_z$. The unit ray vectors are again (44) and using (26) and attenuation formula (37) we obtain

$$\mathbf{U}_0^S=\frac{\Psi(\theta)}{\rho^{1/2}}\,\mathbf{e}_\theta, \tag{48}$$

where $(-i\text{æ})^{-1}\rho_o^{-1/2}\Psi_0(\theta)\mathbf{e}_\theta$ is the amplitude of the vector potential prescribed on $\rho=\rho_o$. Then (30) gives

$$\mathbf{U}_\parallel^S=\frac{1}{\rho^{3/2}}\,\frac{\partial\Psi_0}{\partial\theta}\,\mathbf{e}_\rho. \tag{49}$$

Therefore, the $\rho^{-3/2}$ dependence is manifested again. Direct application of (6) and (19) shows that for the cylindrical P and S waves the first terms (46) and (48) are correct to the first order in $(k\rho)^{-1}$ and $(\varkappa\rho)^{-1}$, respectively. Therefore, the corresponding two-component representations (18) and (33) are approximate and not exact.

3. Waves with spherical fronts

Let us introduce the spherical coordinates (R,θ,ϕ) and consider an elastic P wave with fronts concentrical spheres centered at the origin of coordinates. Let their scalar potential be $\Phi(\mathbf{x})\exp(ikR)$.

For the spherical wave, the ray coordinates may be chosen to coincide with spherical coordinates, so that the angles θ and ϕ define a ray and spherical radius R position on the ray. Then the unit ray vectors are

$$\mathbf{s}=\mathbf{e}_R=(\sin\,\theta\,\cos\,\phi,\ \sin\,\theta\,\sin\,\phi,\ \cos\,\theta). \tag{50}$$

In this case we have

$$J/J_o=R^2/R_o^2, \tag{51}$$

with the subscript o denoting quantities evaluated on an initial wavefront $R=R_o$. When dealing with P waves we use (13) and attenuation formula (36) to obtain

$$\mathbf{U}_0^P=-\frac{\Phi_0(\theta,\phi)}{R}\,\mathbf{e}_R, \tag{52}$$

where $(-ik)^{-1}R_o^{-1}\Phi_0(\theta,\phi)$ is the scalar potential prescribed on $R=R_o$. Then (17) gives

$$\begin{aligned}\mathbf{U}_\perp^P&=-\left[\frac{1}{R}\,\frac{\partial(\mathbf{U}_0^P\cdot\mathbf{s})}{\partial\theta}\,\mathbf{e}_\theta+\frac{1}{R\,\sin\,\theta}\,\frac{\partial(\mathbf{U}_0^P\cdot\mathbf{s})}{\partial\phi}\,\mathbf{e}_\phi\right]\\&=\left[\frac{1}{R^2}\,\frac{\partial\Phi_0}{\partial\theta}\,\mathbf{e}_\theta+\frac{1}{R^2\,\sin\,\theta}\,\frac{\partial\Phi_0}{\partial\phi}\,\mathbf{e}_\phi\right].\end{aligned} \tag{53}$$

This means that the correction term decreases with distance faster than the leading term.

The situation with S waves is similar. Let a vector potential be $\Psi(\mathbf{x})\exp(i\varkappa R)\mathbf{s}_\perp$, with $\mathbf{s}_\perp = s_1\mathbf{e}_\theta + s_2\mathbf{e}_\phi$ and s_1 and s_2 arbitrary constants. The unit ray vectors $\mathbf{s}$ are again (46) and using (26) and attenuation formula (37) we obtain

$$\mathbf{U}_0^S = \frac{\Psi_0(\theta,\phi)}{R}\,\mathbf{s}'_\perp, \tag{54}$$

where $(-i\varkappa)^{-1}R_o^{-1}\Psi_0(\theta,\phi)$ is the amplitude of the vector potential prescribed on $R=R_o$ and $\mathbf{s}'_\perp = s_2\mathbf{e}_\theta - s_1\mathbf{e}_\phi$. Then (30) gives

$$\mathbf{U}_\parallel^S = \frac{1}{R^2\sin\theta}\left[s_2\frac{\partial(\Psi_0\sin\theta)}{\partial\theta} - s_1\frac{\partial\Psi_0}{\partial\phi}\right]\mathbf{e}_R. \tag{55}$$

As above, the correction term decreases with distance faster than the leading term. Kiselev and Roslov (1991) have found the accuracy of these formulas to be very high. Direct application of (6) and (19) shows that for the cylindrical P and S waves the first terms (52) and (54) are correct to the first order in $(k\rho)^{-1}$ and $(\varkappa\rho)^{-1}$, respectively. Therefore, the corresponding two-component representations (18) and (33) are approximate and not exact.

Another interesting case has been considered in Kiselev and Tsvankin (1989). The authors have studied reflection of a spherical S wave from a surface separating two different media when the incident wave is of P type. On the axis of symmetry $\mathbf{u}_0^S$ vanishes, and in the neighborhood of this axis both terms in the two-component representation play a part. It has been found that the accuracy of the representation is high even at distances of half wavelength.

4. Head waves with conical fronts

Many different types of P and S head waves are known in seismology (e.g., Cervený and Ravindra, 1971). We restrict ourselves to S head waves generated by a point source acting on half-plane (the 2-D Lamb problem) and introduce the Cartesian coordinate system $(\mathbf{e}_x,\mathbf{e}_z)$ and the associated polar system $(\mathbf{e}_\rho, \mathbf{e}_\theta)$. Let the origin of coordinates be at the source, the x axis lie on the boundary, and θ be the angle from the z axis measured anticlockwise (Fig. 1).

At each moment t the wave fronts of secondary and primary disturbances issuing from O are concentrical semicircles of radius bt and at, respectively. Let us consider two surface points $O_1(x_1,0)$ and $O_2(x_2,0)$, where $x_1,x_2>0$. At times $x_i/a \leqslant t < x_i/b$, $i=1,2$ each of these points experiences primary disturbance only. According to the generalized Huygens principle (applicable in the presence of boundaries), once reached by primary disturbance at time t_i each $O_i(x,0)$ becomes a secondary source of both primary and secondary waves. At a later time $t' = t_i + t'_i$ the secondary wave front issuing from O_i is a semicircle of radius bt'_i. The disturbance occurring in the shaded area in Fig. 1 is known as *the head wave*. To establish the form of its front we consider similar right angle triangles OBA, O_1B_1A, and O_2B_2A, with $OO_1 = at_1$, $OO_2 = at_2$, $O_1B_1 = bt'_1$, $O_2B_2 = bt'_2$, and $OB/OA = b/a$. It is easy to check that $t_1 + t'_1 = t_2 + t'_2$. Hence, at each moment t' the front of the head wave $\mathbf{u}^H$ is a segment of a straight line (e.g., Brekhovskikh, 1960; Achenbach, 1979). For obvious reasons, the ray forming the angle

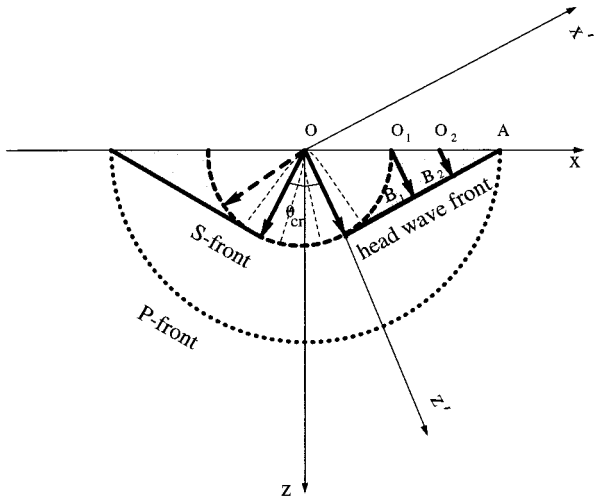


FIG. 1. Geometry of an acoustic vector field around a point source located on a boundary of a solid half-space. Bold dotted line: primary front; bold dashed line: secondary ray and wavefront; bold solid line: head wave rays and fronts. Thin dashed lines delineate boundary layers.

$\theta_{cr} = \sin^{-1} b/a$ with the z axis is called the critical ray. The picture to the left of the z axis is the mirror image of that to the right. For a 3-D point source the front of a head wave is the surface of a conical frustrum.

The generalized Huygens principle combined with the ray representation also allows us to establish how the amplitude of the head wave varies with distance. Indeed, at t_1 the zero displacement condition at O_1 is as follows:

$$[\mathbf{u}^P + \mathbf{u}^H]|_{O_1} = 0. \tag{56}$$

Let both primary and head waves be represented in the form of ray series (3) and substitute (2) and (3) into (56). Using (46) and (47) it follows that at t_1 the amplitude of the head wave at O_1 is $\sim x^{-3/2}$. This is due to attenuation which is experienced by the transverse component of the P wave as it travels along the boundary from O to O_1 (for a more traditional proof, see Friedrichs and Keller, 1955; Brekhovskikh, 1960; and Cervený and Ravindra, 1971). Note that as the head wave travels from O_1 to B_1 no further attenuation results. Indeed, the head wave has a conical front but at each instance only a portion of this front is generated by the old one. The rest is due to new sources appearing on the boundary. In order to be able to treat the head wave propagation as *front* propagation, we introduce an *effective* initial front which is orthogonal to the critical ray and passes through the origin. The subsequent fronts are parallel to it and as with ordinary plane waves there is no geometrical spreading [see (39)]. When the point source is 3D the resulting amplitude is $\sim x^{-2}$ [see (53)].

To simplify further calculations, let us rotate the original coordinate system around the origin. Let the new z axis be aligned with the critical ray (Fig. 1). Then, up to a constant factor the leading order approximation to the vector potential generated by a line source is

$$\boldsymbol{\psi}_1 = \frac{1}{i\varkappa}\frac{1}{(x')^{3/2}}\,e^{i\varkappa z'}\mathbf{e}_{y'}. \tag{57}$$

Here $(\mathbf{e}_{x'},\mathbf{e}_{y'},\mathbf{e}_{z'})$ is the 3-D Cartesian frame of reference defining the corresponding ray coordinates; x' and y' characterize the ray and z' position of the ray. Note that x' is the distance from B_1 to B measured along the front. The preferential direction (of fastest variation in secondary displacement) is along that of propagation of the head wave and according to (19), the two-component representation of the head wave is

$$\tilde{\mathbf{u}}^H=\left[\mathbf{U}_\perp^H+\frac{\mathbf{U}_\parallel^P}{-i\varkappa}\right]e^{i\varkappa z'}. \tag{58}$$

Here the first term is

$$\mathbf{U}_\perp^H=\frac{1}{(x')^{3/2}}\,\mathbf{e}_{x'}, \tag{59}$$

and using (30) we have

$$\mathbf{U}_\parallel^H=\frac{3}{2(x')^{5/2}}\,\mathbf{e}_{z'}. \tag{60}$$

Similar results may be obtained for other types of head waves.

II. THE TWO-COMPONENT BOUNDARY-LAYER REPRESENTATION

In this section we follow the general approach developed in Kiselev (1987) and expand acoustic potentials inside boundary layers into modes displaying features of a plane wave.

A. The two-component representation of the displacement inside the penumbra

Let the tip of a large 2-D wedge be the origin of the Cartesian coordinate system $(\mathbf{e}_x,\mathbf{e}_z)$ and the associated polar system $(\mathbf{e}_\rho,\mathbf{e}_\theta)$. Let a plane P-wave $\mathbf{u}=\exp(ikz)\mathbf{e}_z$ be incident on this wedge from the left, so that the z axis is aligned with the direction of wave propagation and only one of the faces of the wedge is insonified (see Fig. 2). Behind the wedge ($z>0$) in the well-insonified zone ($k\rho\gg1$), the incident and edge waves (the latter are waves diffracted from the tip of the wedge) may be distinguished, with the phase of the incident wave kz and the phase of the edge wave, $k\rho$. This means that two rays of different nature, one incident and another diffracted, run along the z axis. Surrounding this irregularity are parabolic surfaces of constant *phase difference*,

$$k(\rho-z)=\text{const}. \tag{61}$$

Therefore, the geometry of the displacement field around the wedge may be described by the parabolic coordinates

$$\begin{aligned} v&=\pm\sqrt{\sqrt{x^2+z^2}-z}=\sqrt{2\rho}\,\sin\,\theta/2,\\ w&=\sqrt{\sqrt{x^2+z^2}+z}=\sqrt{2\rho}\,\cos\,\theta/2, \end{aligned} \tag{62}$$

where the plus sign corresponds to the half-space above the z axis and minus to the other half-space. Note that we have

$$x=vw,$$

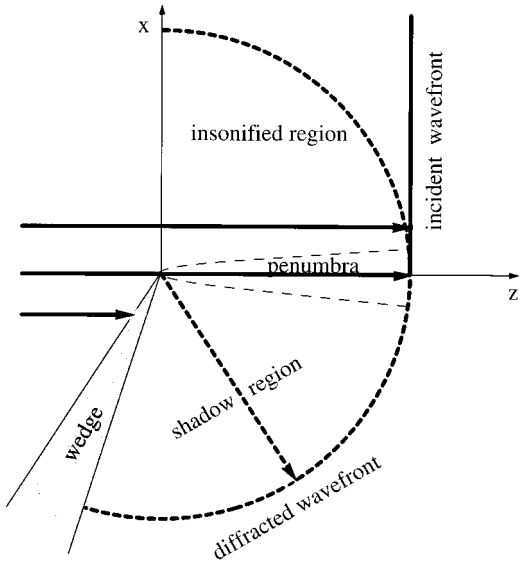


FIG. 2. Geometry of an acoustic scalar field around a wedge. Bold solid line: incident rays and wavefront; bold dashed line: diffracted ray and wavefront. Thin dashed line delineates a boundary layer.

$$z=\frac{w^2-v^2}{2}, \tag{63}$$

and hence the length of infinitesimal distance vector squared is

$$ds^2=dx^2+dz^2=(w^2+v^2)(dw^2+dv^2). \tag{64}$$

Introducing further dimensionless parabolic variables

$$\eta=\sqrt{k}v,\quad \xi=\sqrt{k}w, \tag{65}$$

the Helmholtz equation (7) becomes

$$[\partial_{\xi\xi}+\partial_{\eta\eta}]\phi+(\xi^2+\eta^2)\phi=0. \tag{66}$$

It is well known that this equation is separable and has the following modes:

$$\phi=AD_{-(1-i\lambda)/2}(\sqrt{2}e^{-i\pi/4}\xi)D_{-(1+i\lambda)/2}(-\sqrt{2}e^{-i\pi/4}\eta), \tag{67}$$

where λ is a separation constant and $D_p(q)$'s are the parabolic cylinder functions (e.g., Morse and Feshbach, 1953; Gradsteyn and Ryzhik, 1963, 9.255.2). The parabolic cylinder functions have the following integral representation:

$$D_p(q)=\frac{e^{-q^2/4}}{\Gamma(-p)}\int_0^\infty e^{-qv-v^2/2}v^{-p-1}\,dv,\quad \text{for Re } p<0, \tag{68}$$

and

$$D_p(q)=\frac{1}{\sqrt{\pi}}\,2^{p+1/2}e^{-i(\pi/2)p}e^{q^2/4}\int_{-\infty}^\infty v^p e^{-2v^2+2ivq}\,dv,\quad \text{for Re } p>-1 \tag{69}$$

(Gradsteyn and Ryzhik, 1963, 9.241.2 and 1, respectively). Note that $\Gamma(-p)$ is the gamma function and in (69), $\arg v^p=p\pi i$ when $v<0$. Choosing $A=1/(k\sqrt{\pi i})$, $\lambda=-i$ and using Gradsteyn and Ryzhik (1963, 9.254.1) (67) becomes

$$\phi=\frac{1}{ik}\sqrt{\frac{1}{i\pi}}\,e^{ikz}\int_{-\infty}^{\eta}e^{is^2}\,ds=\frac{1}{ik}F(\eta)e^{ikz}, \tag{70}$$

with the function $F(\eta)$ known as a Fresnel integral (Morse and Feshbach, 1953; Borovikov and Kinber, 1994). Instead of using the general form (67), we could proceed by seeking solution of the Fresnel diffraction problem (66) in the form of $\Phi(\eta)\exp(ikz)$ and assuming $\Phi(\eta)$ to be a slowly varying amplitude. This would lead to the equation for $\Phi(\eta)$, one of whose solutions is $F(\eta)$. We prefer to use slightly more cumbersome formulas (67)–(69) to be able to treat penumbra and neighborhoods of critical rays (see the next section) in the same manner.

Whatever the approach, the above choice of the solution is made to assure that as $k\rho\to\infty$ the asymptotic behavior of ϕ is as follows:

$$\phi=\frac{1}{ik}\frac{1}{2\sqrt{\pi k\rho}\,\sin\,\theta/2}\,e^{ik\rho+i\pi/4}[1+O((k\rho)^{-1})],$$

$$\text{if}\ \ \sin\,\theta/2<0,$$

$$=\frac{1}{ik}\left\{\frac{1}{2\sqrt{\pi k\rho}\,\sin\,\theta/2}\,e^{ik\rho+i\pi/4}[1+O((k\rho)^{-1})]+e^{ikz}\right\},$$

$$\text{if}\ \ \sin\,\theta/2>0 \tag{71}$$

(Gradsteyn and Ryzhik, 1963, Section 8.25). This behavior is entirely consistent with the physical picture in the presence of a wedge (see Morse and Feshbach, 1953; Borovikov and Kinber, 1994): In the geometrical shadow region (where $\sin\,\theta/2<0$) there is only a diffracted cylindrical wave. Above the wedge (where $\sin\,\theta/2>0$) two waves are present, the diffracted cylindrical, $\rho^{-1/2}\exp(ik\rho)$, and plane, $\exp(ikz)$. The latter travels unimpeded to the right. Interestingly, solution (70) is exact, and not the first term of an asymptotic series. This can be seen as due to the fact that the preferential directions in this case are along parabolas and the Helmholtz equation separates in parabolic coordinates exactly.

To obtain the two-component representation of the primary wave inside the penumbra of a wedge ($|x|\ll z$) we apply (6) to (70) and, similarly to (18), we have

$$\tilde{\mathbf{u}}^P=\left[\mathbf{U}_0^P+\frac{\mathbf{U}_\perp^P}{(-ik)^{1/2}}\right]e^{ikz}, \tag{72}$$

with

$$\mathbf{U}_0^P=F(\pm x\sqrt{k/2z})\mathbf{e}_z \tag{73}$$

and

$$\mathbf{U}_\perp^P=-\frac{1}{(-ik)^{1/2}}\,\frac{d}{dx}\,F(\pm x\sqrt{k/2z})\mathbf{e}_x=\mp\sqrt{1/2\pi z}\;\mathbf{e}_x\,. \tag{74}$$

Here the argument of the Fresnel function and $\mathbf{U}_\perp^P$ are both approximated to the leading order in x/z, the top sign corresponds to the insonified, and the bottom to the geometrical shadow region. Term (73) depends on k but remains bounded with k. The underlying l_o may be chosen to be the characteristic scale of variation in F. Then, inside the penumbra, the potential (70) may be seen as factored into a fast-oscillating exponent and slowly varying amplitude. It follows that the locus of irregularity, $x=0$, coincides with the direction of fastest variation in total primary displacement.

Similarly, for S waves inside the penumbra of a wedge we have

$$\boldsymbol{\psi}_0=\frac{1}{i\varkappa}\,F(\eta)e^{i\varkappa z}\mathbf{e}_y\,, \tag{75}$$

and using (19) their two-component representation is

$$\tilde{\mathbf{u}}^S=\left[\mathbf{U}_0^S+\frac{\mathbf{U}_\parallel^S}{(-i\varkappa)^{1/2}}\right]e^{i\varkappa z}, \tag{76}$$

with

$$\mathbf{U}_0^S=-F(\pm x\sqrt{\varkappa/2z})\mathbf{e}_x \tag{77}$$

and

$$\mathbf{U}_\parallel^S=\pm\sqrt{1/2\pi z}\;\mathbf{e}_z\,. \tag{78}$$

B. The two-component representation of the displacement in the vicinity of a critical ray

We return to the 2-D Lamb problem. According to Fig. 1 there is a direction along which two rays of different nature, one carrying the head wave and the other the cylindrical secondary wave run in parallel. The resulting irregularity is characterized by the so-called critical angle $\theta_o=\sin^{-1}b/a$. Let us introduce the same coordinate systems $(\mathbf{e}_{x'},\mathbf{e}_{z'})$ and $(\mathbf{e}_{\rho'},\mathbf{e}_{\theta'})$ as in Fig. 1, with the origin at the source and the z' axis aligned with the critical ray. Then the head wave is $\mathbf{u}=(x')^{-3/2}\exp(i\varkappa z')\;\mathbf{e}_{y'}$ [see (57)]. Inside the solid ($z'>0$) and away from the critical ray ($\varkappa\rho'\gg1$), the head and secondary waves may be distinguished, with the phase of the head wave $\varkappa z'$ and the phase of the secondary wave, $\varkappa\rho'$. Thus, inside this region, similarly to the penumbra case, the level surfaces of constant *phase difference* are parabolas. Introducing new variables

$$v'=\pm\sqrt{\sqrt{x'^2+z'^2}-z'}=\sqrt{2\rho'}\,\sin\,\theta'/2,$$

$$w'=\sqrt{\sqrt{x'^2+z'^2}+z'}=\sqrt{2\rho'}\,\cos\,\theta'/2, \tag{79}$$

$$\xi'=\sqrt{\varkappa}w',\quad \eta'=\sqrt{\varkappa}v'$$

we choose the general modes similar to (67),

$$\boldsymbol{\psi}=AD_{-(1-i\lambda)/2}(\sqrt{2}e^{-i\pi/4}\xi')D_{-(1+i\lambda)/2}(-\sqrt{2}e^{-i\pi/4}\eta')\mathbf{e}_{y'}\,. \tag{80}$$

Further specifying $A=4\sqrt{i\varkappa}$ and $\lambda=2i$ we obtain

$$\boldsymbol{\psi}=4(i\varkappa)^{1/2}D_{-3/2}(\sqrt{2}e^{-i\pi/4}\xi')D_{1/2}(-\sqrt{2}e^{-i\pi/4}\eta')\mathbf{e}_{y'}\,. \tag{81}$$

The above choice of the coefficient, order and form of the argument of the parabolic cylinder function is made to assure that as $\varkappa\rho'\to\infty$ the asymptotic behavior of $\boldsymbol{\psi}$ is as follows:

$$\boldsymbol{\psi}=\frac{4i\ \tan\ \theta'/2}{(\rho'\ \sin\ \theta')^{1/2}}\,e^{i\varkappa\rho'}[1+O((\varkappa\rho')^{-1})]\mathbf{e}_{y'},$$

$$\text{if}\quad \sin\ \theta'/2<0,$$

$$=\left[\frac{4i\ \tan\ \theta'/2}{(\rho'\ \sin\ \theta')^{1/2}}\,e^{i\varkappa\rho'}+\frac{1}{i\varkappa(\rho'\ \sin\ \theta')^{3/2}}\,e^{i\varkappa z'}\right]$$

$$\times[1+O(\varkappa\rho')^{-1})]\mathbf{e}_{y'},\quad \text{if}\quad \sin\ \theta'/2>0 \tag{82}$$

[Gradsteyn and Ryzhik, 1963, 9.246(1) and (2), respectively; also we used the fact that $\Gamma(1/2)=-2\sqrt{\tau}-ibid.$, 8.338(3)]. This behavior is entirely consistent with the physical picture in the high-frequency regime (Brekhovskikh, 1960; Červený and Ravindra, 1971; Achenbach, 1979): To the left (where $\sin\theta'/2<0$) there is only the secondary cylindrical wave. To the right (where $\sin\theta'/2>0$) two waves are present, the secondary cylindrical, $|x'|^{-1/2}\exp(i\varkappa\rho')$, and the head wave, $|x'|^{-3/2}\exp(i\varkappa z')$, which is traveling down. As in the penumbra case, solution (81) is exact and not the first term of an asymptotic series.

In the vicinity of a critical ray ($|x'|\ll z'$) (81) takes the form

$$\boldsymbol{\psi}=4i(-i\varkappa)^{1/2}D_{-3/2}(2e^{-i\pi/4}\sqrt{\varkappa z'})$$

$$\times D_{1/2}(-e^{-i\pi/4}x'\sqrt{\varkappa/z'})\mathbf{e}_{y'}. \tag{83}$$

Here the arguments of the parabolic cylinder functions approximate the exact argument to the leading order in x'/z'. To obtain the corresponding asymptotic two-component representation of the acoustic wave we apply (19) to (83) and, similarly to (33), we have

$$\tilde{\mathbf{u}}=\left[\mathbf{u}_0+\frac{\mathbf{u}_{\|}}{(-i\varkappa)^{1/2}}\right], \tag{84}$$

with

$$\mathbf{u}_0=-4i(-i\varkappa)^{3/2}D_{-3/2}(2e^{-i\pi/4}\sqrt{\varkappa z'})$$

$$\times D_{1/2}(-e^{-i\pi/4}x'\sqrt{\varkappa/z'})\mathbf{e}_{x'}$$

$$=-\sqrt{2}iz'^{-3/2}D_{1/2}(-e^{-i\pi/4}x'\sqrt{\varkappa/z'})$$

$$\times[1+O((\varkappa z')^{-1})]e^{i\varkappa z'}\mathbf{e}_{x'} \tag{85}$$

and

$$\mathbf{u}_{\|}=4(i\varkappa)^{3/2}(iz')^{-1/2}D_{-3/2}(2e^{-i\pi/4}\sqrt{\varkappa z'})$$

$$\times D_{3/2}(-e^{-i\pi/4}x'\sqrt{\varkappa/z'})\mathbf{e}_{z'}$$

$$=-\sqrt{2}iz'^{-5/4}D_{3/2}(-e^{-i\pi/4}x'\sqrt{\varkappa/z'})$$

$$\times[1+O((\varkappa z')^{-1})]e^{i\varkappa z'}\mathbf{e}_{z'}. \tag{86}$$

The asymptotics in (85) and (86) are based on Gradsteyn and Ryzhik [1963; 9.246(1)] and apply when $\varkappa z'$ is large but $\sqrt{\varkappa z'}x'/z'$ is not. Both (85) and (86) depend on $\varkappa$ but remain bounded with $\varkappa$. The vector potential $\boldsymbol{\psi}$ in (83) contains an oscillating factor $\exp(i\varkappa z')$ [see (68), (69), and (79)]. The underlying l_o may be chosen to be the characteristic scale of variation in the amplitude of $\boldsymbol{\psi}$. It follows that for high frequencies the direction of fastest variation in total secondary displacement is along the critical ray.

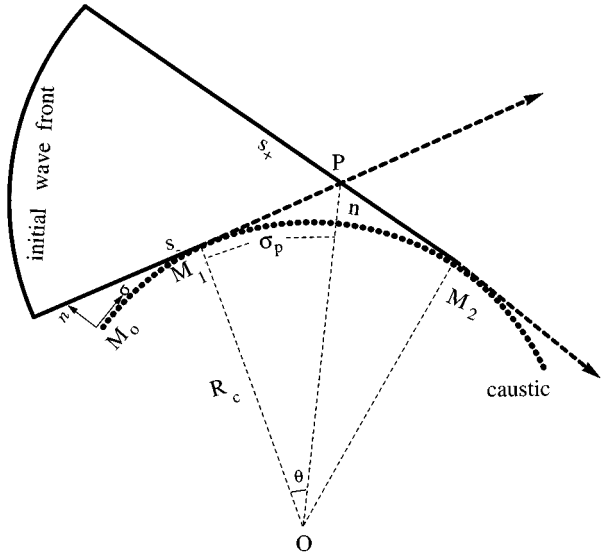


FIG. 3. Geometry of an acoustic scalar field around a simple caustic. Bold solid line: incoming rays; bold dashed line: outgoing rays.

Expression (81) was previously obtained by Brekhovskikh (1960) using the Fourier decomposition (see also Červený and Ravindra, 1971). Our derivation is simpler and follows that of Tsepelev (1975).

C. The two-component representation of the displacement in the vicinity of a simple caustic

Let a P wave $\mathbf{u}=\exp(iks)\mathbf{s}$ propagate along the rays which issue from a 2-D initial wave front and form a caustic; that is, let the family of rays have an envelope (Fig. 3). As above, $\mathbf{s}$ is a unit ray vector. Let us restrict ourselves to a *simple* portion of this caustic, that is, assume that at any tangency point M the second derivatives of the caustic curve and the second derivatives of the tangent ray do not coincide. This assumption excludes turning points and asymptotes and means that locally the caustic behaves as a circle (e.g., Babič and Buldyrev, 1991; Borovikov and Kinber, 1994).

The above definitions imply that at every tangency point M the gradient $\nabla\tau=(d\tau/ds)\mathbf{s}$ is directed along the caustic. Therefore, *on* the caustic the direction of the fastest phase variation is *along* the caustic and we have

$$a\Delta\tau=a\int_{M_o}^{M_1}\frac{d\tau}{ds}\,ds=\sigma_{o1}, \tag{87}$$

where $\Delta\tau$ is the phase difference between waves reaching points M_o and M_1, and σ_{o1} is the length of the arc joining M_o and M_1. Formula (87) allows us to describe each point P in this vicinity using the local caustic coordinate frame $[\mathbf{e}_\sigma(\mathbf{x}),\mathbf{e}_n(\mathbf{x})]$ with its origin at M_1 and unit vectors $\mathbf{e}_\sigma(\mathbf{x})$ and $\mathbf{e}_n(\mathbf{x})$ tangential and normal to the caustic, respectively. The distance σ is measured along the caustic from an arbitrarily but fixed point M_o.

Since locally any arc of a simple caustic behaves as a circle, the caustic coordinate system is simply related to the polar coordinate frame $(\mathbf{e}_\rho,\mathbf{e}_\theta)$ whose origin is at the center of curvature O and whose underlying Cartesian coordinate system $(\mathbf{e}_x,\mathbf{e}_z)$ has its z axis aligned with OM_1. Let positive n

correspond to the outward normal direction, that is, the insonified side of the caustic. Then $n=\rho-R_c$ and the infinitesimal arc length along the curve is $d\sigma=R_c\,d\theta$, with R_c the radius of curvature at M_1. If $K=R_c^{-1}$ denotes the curvature, then the length of infinitesimal vector squared is

$$ds^2=d\rho^2+\rho^2\,d\theta^2=dn^2+(1+nK)^2\,d\sigma^2. \tag{88}$$

Points P which lie inside the insonified side may also be described by ray coordinates. However, it has to be noted that in this neighborhood any point P is the crossing point of *two* rays of different nature, M_1P which has passed the caustic already and PM_2 which is going to touch it at M_2. We specify ray coordinates of points lying on the outgoing ray M_1P as (σ_{o1},s_-). Thus, M_1P is characterized by its tangency point and position on M_1P by the phase $s_-=a\tau_-$. This is the phase gained traveling between the initial wave front and P. Similarly, ray coordinates of points lying on the incoming ray PM_2 are (σ_{o2},s_+).

Considering right-angled triangles M_iPO ($i=1$ or 2) we have

$$\frac{R_c}{R_c+n}=\cos\,\theta=1-\frac{\theta^2}{2}+O(\theta^4), \tag{89}$$

and

$$M_iP=(R_c+n)\sin\,\theta=(R_c+n)\left[\theta-\frac{\theta^3}{6}+O(\theta^5)\right]. \tag{90}$$

Relationship (89) gives

$$\theta=\sqrt{2Kn}+O(n). \tag{91}$$

Substituting (91) into (90) and taking into account that $\sigma_P=R_c\theta$, (90) implies

$$M_iP=\sigma_P+\tfrac{2}{3}\sqrt{2K}n^{3/2}+O(n^2). \tag{92}$$

Note that if we did not include the second-order term in (89) we would obtain no relationship between θ and n. Also, neglecting the third-order term in (90) would lead to inconsistency, since $n\theta$ and $R_c\theta^3$ in (90) are of the same order.

Now, the wave leaving P along M_1P has the total phase

$$s_-=s_o+\sigma_{o1}+M_1P=\sigma+\tfrac{2}{3}2Kn^{3/2}+O(n^2), \tag{93}$$

where the constant s_o is the phase along the ray tangent to the caustic at M_o. We remind the reader that $\sigma=s_o+\sigma_{o1}+\sigma_P$. Similarly, taking into account (87) the wave reaching P along M_2P has the total phase

$$s_+=s_o+\sigma_{o1}+2\sigma_P-M_2P=\sigma-\tfrac{2}{3}\sqrt{2K}n^{3/2}+O(n^2). \tag{94}$$

Inverting (93) and (94) in the vicinity of a simple caustic to the leading order in n, the caustic coordinates are related to the ray coordinates in the following manner:

$$\sigma=\frac{s_++s_-}{2},\qquad n=\left[-\frac{3}{4\sqrt{2K}}(s_+-s_-)\right]^{2/3}. \tag{95}$$

Having established the above relationships between coordinate systems, we can estimate Jacobians $J_\pm$ of the incoming and outgoing rays, respectively. The resulting estimates are

$$J_\pm=\frac{\partial(x,z)}{\partial(\theta,\rho)}\cdot\frac{\partial(\theta,\rho)}{\partial(\sigma,n)}\cdot\frac{\partial(\sigma,n)}{\partial(s_\pm,s_\pm)}\sim n^{1/2} \tag{96}$$

(cf. Landau and Lifschitz, 1979; Babič and Buldyrev, 1991; Borovikov and Kinber, 1994).

To continue, (88) implies that in the vicinity of a simple caustic the Helmholtz equation (7) may be rewritten in caustic coordinates to give

$$(\nabla^2+k^2)\phi=\frac{1}{1+nK}\left\{\frac{\partial}{\partial\sigma}\left(\frac{1}{1+nK}\frac{\partial}{\partial\sigma}\right)+\frac{\partial}{\partial n}\left[(1+nK)\frac{\partial}{\partial n}\right]\right\}\phi+k^2\phi=0. \tag{97}$$

Assuming that these coordinates allow asymptotic separation of variables, we seek solution of (97) in the form

$$\phi(\mathbf{x},k)=\Phi(\eta,k)e^{ik\sigma}, \tag{98}$$

where we have

$$\eta=-(2k^2K)^{1/3}n. \tag{99}$$

Then substituting (98) into (97), in the vicinity of a caustic ($nK\ll1$) Eq. (97) reduces to

$$\left[\frac{d^2}{d\eta^2}-\left(\frac{K}{k}\right)^{2/3}\frac{d}{d\eta}-\eta\right]\Phi(\eta,k)=0. \tag{100}$$

Let us assume further that wave number k is large and in the vicinity of a simple caustic amplitude Φ varies slowly. This is equivalent to assuming that there exists a large parameter $k\gg1$ (with the underlying l_o the characteristic scale of variation in Φ in this direction) and $\Phi(\eta,k)$ can be represented as an asymptotic series in inverse powers of k. Thus, the potential ϕ in (98) can be seen as factored into a fast-oscillating exponent which depends on the *mean phase* and amplitude which depends on η, that is, the distance to the caustic only. We can now equate the coefficients of successive powers of k in (100) to zero to see that in the vicinity of a simple caustic ($nK\ll1$) the first term $\Phi_0(\eta)$ in this asymptotic series satisfies the Airy equation,

$$\left[\frac{d^2}{d\eta^2}-\eta\right]\Phi_0=0. \tag{101}$$

One solution of (101) is an Airy function,

$$\mathrm{Ai}(\eta)\equiv\frac{1}{\sqrt{\pi}}\int_0^\infty\cos(\beta\eta+\beta^3/3)\,d\beta. \tag{102}$$

This function possesses the following asymptotics

$$\begin{aligned}\mathrm{Ai}(\eta)&=[\tfrac{1}{2}|\eta|^{-1/4}[e^{i(-(2/3)|\eta|^{3/2}+\pi/4)}+e^{i((2/3)|\eta|^{3/2}-\pi/4)}]\\&\quad\times[1+O(|\eta|^{-3})]\qquad\text{as}\quad\eta\to-\infty,\\&=\tfrac{1}{2}\eta^{-1/4}e^{-(2/3)\eta^{3/2}+i\pi/4}[1+O(\eta^{-3/2})]\quad\text{as}\quad\eta\to\infty\end{aligned} \tag{103}$$

[Babič and Buldyrev, 1991, (A.1.26) and (A.1.25), respectively]. Therefore, as $|\eta|\to\infty$ the zeroth approximation to the scalar potential,

$$\phi_0=\frac{2}{ik}\,\mathrm{Ai}(\eta)e^{i(k\sigma-\pi/4)} \tag{104}$$

behaves as follows:

$$\phi_0=\frac{1}{ik}\,|\eta|^{-1/4}[e^{iks_+}+e^{iks_--i\pi/2}+O(|\eta|^3)],\quad \text{if}\quad \eta<0,$$

$$=\frac{1}{ik}\,\eta^{-1/4}e^{-(2/3)\eta^{3/2}+ik\sigma}[1+O(\eta^{-3/2})],\qquad \text{if}\quad \eta>0. \tag{105}$$

Here we used (95) and the fact that it implies

$$\frac{s_+-s_-}{2}=-\frac{2}{3ik}\,\eta^{3/2} \tag{106}$$

to the leading order in η. The asymptotic behavior (105) is entirely consistent with the physical picture (e.g., Babič and Kirpičnikova, 1975; Pierce, 1981; Babič and Buldyrev, 1991): In the outward region of a simple caustic (where $n>0$ and hence $\eta<0$) there exist two time-harmonic waves, one approaching and the other moving away from it. In the inward region (where $n<0$ and hence $\eta>0$) only one (exponentially) fast decreasing wave is present. Note that the outgoing wave has the $-\pi/2$ phase shift and that $\eta^{-1/4}$ is the geometrical spreading associated with both rays [see (36), (96), and (99)]. The presence of the shift reflects the fact that on passing the caustic, due to diffraction, rays change their nature.

To obtain the two-component representation of the primary wave in the vicinity of a simple caustic we apply (6) to (104) and, similarly to (18), we have

$$\tilde{\mathbf{u}}^P=\left[\mathbf{U}_0^P+\frac{\mathbf{U}_\perp^P}{-ik^{1/3}}\right]e^{ik\sigma}, \tag{107}$$

where

$$\mathbf{U}_0^P=\mathrm{Ai}(-(2k^2K)^{1/3}n)\mathbf{e}_\sigma \tag{108}$$

and the second term is

$$\mathbf{U}_\perp^P=-k^{-2/3}\,\frac{d\ \mathrm{Ai}(-(2k^2K)^{1/3}n)}{dn}\,\mathbf{e}_n$$

$$=-(2K)^{1/3}\,\mathrm{Ai}'(-(2k^2K)^{1/3}n)\mathbf{e}_n. \tag{109}$$

Here $'$ denotes differentiation with respect to the argument of the Airy function. Both the Airy function in (108) and its derivative in (109) depend on k but remain bounded with k. In general, both have zeros, but on the caustic itself ($n=0$) the second to first term ratio is

$$\frac{|\mathbf{U}_\perp^P|}{|\mathbf{U}_0^P|}=(2K)^{1/3}\,\frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}\approx 0.73(2K)^{1/3}. \tag{110}$$

The full asymptotic series for $\mathbf{u}$ can be found, e.g., in Babič and Buldyrev (1991).

Similarly, for an S wave in the vicinity of a simple caustic we can choose

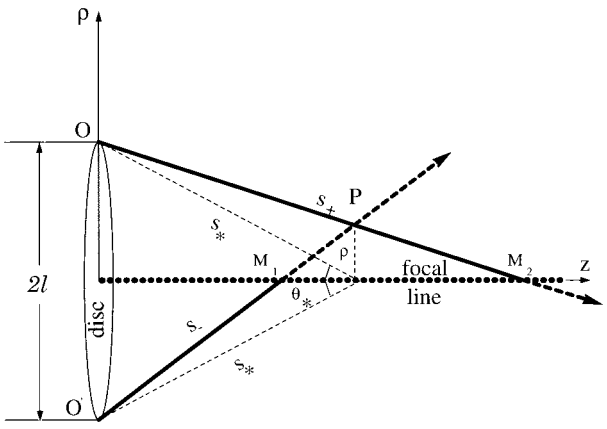


FIG. 4. Geometry of an acoustic scalar field around a focal line. Bold solid line: incoming rays; bold dashed line: outgoing rays.

$$\boldsymbol{\psi}_0=\frac{2}{i\varkappa}\,\mathrm{Ai}(-(2\varkappa^2K)^{1/3}n)e^{i(\varkappa\sigma-\pi/4)}\mathbf{e}_y, \tag{111}$$

where $(\mathbf{e}_\sigma,\mathbf{e}_y,\mathbf{e}_n)$ is a local 3-D coordinate system. Then using (19) its two-component representation is

$$\tilde{\mathbf{u}}^S=\left[\mathbf{U}_0^S+\frac{\mathbf{U}_\parallel^S}{-i\varkappa^{1/3}}\right]e^{i\varkappa x}, \tag{112}$$

with

$$\mathbf{U}_0^S=-\mathrm{Ai}(-(2\varkappa^2K)^{1/3}n)\mathbf{e}_n \tag{113}$$

and

$$\mathbf{U}_\parallel^S=(2K)^{1/3}\,\mathrm{Ai}'(-(2\varkappa^2K)^{1/3}n)\mathbf{e}_\sigma. \tag{114}$$

It follows that on the caustic the second to first term ratio is

$$\frac{|\mathbf{U}_\parallel^S|}{|\mathbf{U}_0^S|}=(2K)^{1/3}\,\frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}\approx 0.73(2K)^{1/3}. \tag{115}$$

D. The two-component representation of the displacement in the vicinity of a focal line

Let a P wave $\mathbf{u}=\exp(iks)\mathbf{s}$ be incident from the left on an infinitely thin lens-shaped disk with the diameter OO' of length $2l$ (see Fig. 4). Let the center of the disk be the origin of the Cartesian coordinate system $(\mathbf{e}_x,\mathbf{e}_y,\mathbf{e}_z)$ and cylindrical polar system $(\mathbf{e}_\rho,\mathbf{e}_\theta,\mathbf{e}_z)$, and let the plane of the disk be normal to the z axis. Consider a plane which passes through the axis of the disk and an arbitrary point $P(\rho,\theta,z)$ lying in the vicinity of the focal line ($\rho\ll l$). Focal lines are degenerate caustics and as for simple caustics, each point P is the crossing point of *two* rays of different nature, $O'P$ which has already crossed the focal line at $M_1(0,\theta,z_1)$ and OP which is going to cross it at $M_2(0,\theta,z_2)$. We specify ray coordinates of points lying on the outgoing ray $O'P$ as (θ,z_1,s_-). Thus, $O'P$ is characterized by the θ coordinate of its origin O and the z coordinate of its intersection with the focal line. The position of P on the ray is characterized by the phase s_-. This is the phase gained traveling between the initial wave front and P. Similarly, ray coordinates of points lying on the incoming ray OP are (θ,z_2,s_+). It can be established without difficulty that we have

$$z_{1,2}=z\left[1\pm\frac{\rho}{l}+O\left(\frac{\rho^2}{l^2}\right)\right] \tag{116}$$

and

$$s_{\pm}=\sqrt{z^2+(l\pm\rho)^2}=\sqrt{z^2+l^2}\left[1\pm\frac{\rho l}{z^2+l^2}+O\left(\frac{\rho^2}{l^2}\right)\right]. \tag{117}$$

Therefore, the Jacobians $J_{\pm}$ of the incoming and outgoing rays both satisfy

$$J_{\pm}\sim\rho. \tag{118}$$

Let us now introduce variables

$$s_*\equiv\sqrt{z^2+l^2}=\frac{s_++s_-}{2},$$
$$\eta\equiv k\rho\ \sin\ \theta_*(z)=k\ \frac{s_+-s_-}{2}. \tag{119}$$

where $\sin\theta_*(z)\equiv l/\sqrt{z^2+l^2}$ and equalities hold to the leading order in ρ/l. Comparing (106) with (119) we seek solution of the Helmholtz equation (7) in a form similar to (98), that is,

$$\phi(\mathbf{x},k)=\Phi(\eta,k)e^{iks_*}=\Phi(k\rho\ \sin\ \theta_*(z),k)e^{iks_*(z)}. \tag{120}$$

Let us now assume that the wave number k is large and the problem is axisymmetric. Substituting (120) into (7) in the vicinity of a focal line ($\rho\ll l$) we see that the Helmholtz equation (7) reduces to

$$\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\ \rho\ \frac{\partial}{\partial\rho}+k^2\ \sin^2\ \theta_*(z)\right]\Phi+O(k)=0. \tag{121}$$

Let us assume further that in the vicinity of a focal line, the amplitude $\Phi(\eta, k)$ varies slowly; that is, there exists a relatively large parameter $k\gg1$ (with the underlying l_o the characteristic scale of variation in Φ) and $\Phi(\eta, k)$ can be represented as an asymptotic series in inverse powers of k. Then the potential ϕ in (120) can be seen as factored into a fast-oscillating exponent which depends on the *mean phase* and the amplitude which depends on η, that is, the distances to the focal line and to the disk only. We can now equate the coefficients of successive powers of k in (131) to zero to see that in the vicinity of a focal line ($\rho\ll l$) the first term $\Phi_0(\eta)$ in this asymptotic series satisfies the zeroth order Bessel equation.

$$\left[\frac{1}{\eta}\frac{\partial}{\partial\eta}\ \eta\ \frac{\partial}{\partial\eta}+1\right]\Phi_0=0. \tag{122}$$

Returning to the cylindrical coordinates, one solution of (122) is

$$\Phi_0=\frac{1}{ik}\ \sqrt{2\pi}\ e^{-i\pi/4}J_o(\eta), \tag{123}$$

where

$$J_o(\eta)=\frac{2}{\pi}\int_0^{\pi/2}\cos(\eta\ \sin\ \theta)d\theta \tag{124}$$

is a Bessel function of zeroth order [Gradsteyn and Ryzhik, 1963, 8.411(2)]. This is the only solution which is regular at $\eta=0$. As $|\eta|\to\infty$ the corresponding zeroth approximation to the scalar potential behaves as follows:

$$\phi_0=\frac{1}{ik}\left\{\frac{1}{\eta^{1/2}}[e^{iks_+}+e^{iks_--i\pi/2}]+O(\eta^{-2})\right\} \tag{125}$$

(Gradsteyn and Ryzhik 1963, 8.451). This behavior is entirely consistent with the physical picture: At any point on any side of the focal line two rays intersect, one approaching the line and the other moving away from it. Note that the outgoing wave has the $-\pi/2$ phase shift and that $\eta^{-1/2}$ is the geometrical spreading associated with both rays [see (118) and (119)].

Above we used the treatment of the acoustic potential in the vicinity of a focal line ($\rho\ll l$) as first presented by Kiseley (1980). To obtain the two-component representation of the primary displacement inside this region, we apply (6) to (123) and write

$$\tilde{\mathbf{u}}^P=[\mathbf{U}_0^P+\mathbf{U}_\perp^P]e^{iks_*(z)}, \tag{126}$$

where

$$\mathbf{U}_0^P=\sqrt{2\pi}\ \cos\ \theta_*(z)J_o(k\rho\ \sin\ \theta_*(z)e^{-i\pi/4}\mathbf{e}_z, \tag{127}$$

and

$$\mathbf{U}_\perp^P=\frac{1}{ik}\ \sqrt{2\pi}\ \frac{dJ_o(k\rho\ \sin\ \theta_*(z))}{d\rho}\ e^{-i\pi/4}\mathbf{e}_\rho$$
$$=\sqrt{2\pi}\ \sin\ \theta_*(z)J_1(k\rho\ \sin\ \theta_*(z)e^{i\pi/4}\mathbf{e}_\rho, \tag{128}$$

where the Bessel function of the first order $J_1(v) = -J_o'(v)$, and $'$ denotes differentiation with respect to the argument of the Bessel function. Note that the focal line does not coincide with the direction of fastest variation in primary displacement. Note too that outside zeros of J_o and J_1 both terms are of the same order. On the other hand, on the focal line the second to first term ratio is

$$\frac{|U_\perp^P|}{|U_0^P|}=\frac{J_o'(0)}{J_o(0)}=0. \tag{129}$$

Similarly, for S waves in the vicinity of a focal line we have

$$\boldsymbol{\psi}_0=\frac{1}{i\varkappa}\ \sqrt{2\pi}J_o(\varkappa\rho\ \sin\ \theta_*(z))e^{i(\varkappa s_*(z)-\pi/4)}\mathbf{e}_\theta. \tag{130}$$

Then, using (19), their two-component representation is

$$\tilde{\mathbf{u}}^S=[\mathbf{U}_0^S+\mathbf{U}_\parallel^S]e^{i\varkappa s_*(z)}, \tag{131}$$

with

$$\mathbf{U}_0^S=\sqrt{2\pi}\ \cos\ \theta_*(z)J_0(\varkappa\rho\ \sin\ \theta_*(z))e^{3\pi i/4}\mathbf{e}_\rho \tag{132}$$

and

$$\mathbf{U}_\parallel^S=\sqrt{2\pi}\ \sin\ \theta_*(z)J_1(\varkappa\rho\ \sin\ \theta_*(z))e^{5\pi i/4}\mathbf{e}_z. \tag{133}$$

Again, on the focal line the second to first term ratio is

$$\frac{|U_\parallel^S|}{|U_0^S|}=0. \tag{134}$$

III. CONCLUSIONS

We have shown that when studying propagation and diffraction of high-frequency elastic body waves in homogeneous and isotropic materials it is often possible to find modes which display features of a plane wave. This is done by

(1) introducing a local coordinate system (v_1, v_2, v_3), with a unit vector $\mathbf{e}_3$ oriented along a preferential direction determined by geometry of the ray field; and
(2) separating variables (usually approximately), so that the acoustic *potentials* are represented as products of two factors dependent, respectively, on v_3 only and the other two coordinates only. As a rule, the former factor turns out to be a fast-oscillating exponent.

It then follows via (6) and (19) that the *displacements* may be represented as a sum of two orthogonal terms, the first one of which is aligned with the preferential direction $\mathbf{e}_3$ or is perpendicular to it, depending on whether the waves are primary or secondary. When such direction coincides with the direction of fastest variation in the displacement, the magnitude of the second term proves to be smaller than the first and their ratio decreases with frequency. The focal line is one phenomenon in our album for which there is no direction of fastest variation in the total displacement. Remarkably, the corresponding acoustic potentials may still be factored and the form of the two factors may be guessed at by treating focal lines as degenerate caustics. Notably, although the second terms in the two-component representation are zero on the focal line, outside such zeros both components are of the same order.

According to various authors, the two-component representation often describes diffraction effects in the intermediate- as well as the high-frequency regimes, and it has been shown that such representations may be found for inhomogeneous and weakly anisotropic solids as well. In view of this, and since the second terms in these representations are simple to evaluate, it is our belief that they may be used to a great effect in mathematical modeling of nondestructive testing as well as seismology.

ACKNOWLEDGMENTS

Part of this work was carried out by the second author at the South Bank University, London as an U.K. Royal Society Ex-Quota Fellow. We are grateful to Dr. J. A. Hudson for a careful reading of the manuscript and insightful comments.

Achenbach J. D. (**1979**). *Wave Propagation in Elastic Solids* (North-Holland, New York).

Alekseev, A. S., Babič, V. M., and Gel'chinsky B. Ya. (**1961**). ''Ray method of calculation of intensity of wave fronts,'' in *Voprosy Dynamicheskoi Teorii Rasprostraneniya Seismicheskikh Voln* (Problems in the Dynamic Theory of Propagation of Seismic Waves), edited by G. I. Petrashenl (Leningrad University, Leningrad), Vol. 5, pp. 3–24.

Alekseev, A. S., and Mikhailenko, B. G. (**1982**). '' 'Nongeometrical phenomena' in the theory of propagation of seismic waves,'' Dokl. Akad. Nauk SSSR **267**(5), 1079–1803.

Babič, V. M. (**1956**). ''Ray method of calculation of intensity of wave fronts,'' Dokl. Akad. Nauk SSSR **110**(3), 355–367.

Babič, V. M., and Buldyrev, V. S. (**1991**). *Short-Wavelength Diffraction Theory: Asymptotic Methods* (Springer-Verlag, Berlin).

Babič, V. M., and Kirpičnikova, N. Ya. (**1975**). *The Boundary-Layer Method in Diffraction Problems* (Springer-Verlag, Berlin).

Babich, V. M., and Kiselev, A. P. (**1989**). ''Non-geometrical waves–are there any? An asymptotic description of some ''non-geometrical'' phenomena in seismic wave propagation,'' Geophys. J. Int. **99**(2), 415–420.

Borovikov, V. A., and Kinber, B. Ye. (**1994**). *Geometrical Theory of Diffraction* (IEE Electromagnetic series Stewenage), Vol. 37.

Brekhovskikh, L. M. (**1960**). *Waves in Layered Media* (Academic, New York).

Buchal, R. N., and Keller, J. B. (**1960**). ''Boundary layer problems in diffraction theory,'' Comm. Pure Appl. Math. **13**(1), 85–114.

Červený, V., and Ravindra R. (**1971**). *Theory of Seismic Head Waves* (University of Toronto, Toronto).

Červený, V., Molotkov, I. A., and Pšenčík, I. (**1977**). *Ray Method in Seismology* (Karlov University, Prague).

Daley, P. F., and Hron, F. (**1987**). ''Reflection of spherical *P*-wave on a free surface (near-vertical incidence),'' Bull. Seismol. Soc. Am. **77**, 1057–1070.

Fung, Y. C. (**1965**). *Foundations of Solid Mechanics* (Prentice-Hall, Englewood Cliffs, NJ).

Fock, V. A. (**1965**). *Electromagnetic Diffraction and Propagation Problems* (Pergamon, Oxford).

Friedrichs, K. O., and Keller, J. B. (**1955**). ''Geometrical acoustics. II. Diffraction, reflection, and refraction of a weak spherical or cylindrical shock at a plane interface,'' J. Appl. Phys. **26**(8), 961–66.

Gradshteyn, I. S., and Ryzhik, I. M. (**1963**). *Table of Integrals, Series and Products* (Academic, New York).

Hron, F., and Mikhailenko, B. G. (**1981**). ''Numerical modeling of nongeometrical effects by Alekseev–Mikhailenko method,'' Bull. Seismol. Soc. Am. **71**(4), 1011–1029.

Hudson, J. A. (**1980**). *The Excitation and Propagation of Elastic Waves* (Cambridge U.P., Cambridge).

Karal, F. C., and Keller, J. B. (**1959**). ''Elastic Wave Propagation in Homogeneous and Inhomogeneous Media,'' J. Acoust. Soc. Am. 31, 694–705.

Keller, J. B. (**1958**). ''Geometrical Theory of Diffraction,'' Proc. Symp. Appl. Math. **8**, 27–52.

Kiselev, A. P. (**1980**). ''Focal Boundary Layer in Axisymmetric Diffraction Problems,'' Izv. VUZ Radiofiz. **13**(3), 333–352 [English transl.: Radiophys. Quantum Electron. **23**, 241–252 (1980)].

Kiselev, A. P. (**1983**). ''Extrinsic Components of Elastic Waves,'' Izv. Akad. Nauk SSSR, Fiz. Zemli **9**, 51–61 [English transl.: Izv. Acad. Sci. SSSR, Phys. Solid Earth **19**(9), 707–710 (1983)].

Kiselev, A. P. (**1987**). ''Depolarization in Diffraction,'' Zh. Tekh. Fiz. **56**, 1184–1185 [English transl.: Sov. Phys. Tech. Phys. **32**(2), 695–696 (1987)].

Kiselev, A. P., and Tsvankin, I. D. (**1989**). ''A method of Comparison of Exact and Asymptotic Wave Field Computations,'' Geophys. J. **96**(2), 253–258.

Kiselev, A. P., and Roslov, Yu. V. (**1991**). ''Use of Additional Components for Numerical Modeling of Polarization Anomalies of Elastic Body Waves,'' Geol. Geofiz. **4**, 121–131 [English transl.: Sov. Geol. Geophys. **32**(4), 105–114 (1991)].

Kiselev, Yu. V., and Kashtan, B. M. (**1988**). ''Etalon Problems for the Most Simple Piecewisely Homogeneous Media with Parallel Plane Interfaces,'' in *Voprosy Dynamicheskoi Teorii Rasprostraneniya Seismicheskikh Voln* (Problems in the Dynamic Theory of Propagation of Seismic Waves), edited by G. I. Petrashen (Leningrad University, Leningrad), Vol. 27, pp. 98–116.

Kravtsov, Yu. A., and Orlov, Yu. I. (**1990**). *Geometrical Optics in Inhomogeneous Media* (Springer-Verlag, Berlin).

Landau, L. D., and Lifschitz, E. M. (**1975**). *The Classical Theory of Fields. Course of Theoretical Physics* (Pergamon, Oxford), Vol. 2.

Li, Y. G., Leary, P. C., and Aki, K. (**1987**). ''Observation and modeling of fault zone fracture anisotropy. II. *P*-wave polarization anomalies,'' Geophys. J. **91**(2), 485–492.

Morse, P. M., and Feshbach, H. (**1953**). *Methods of Theoretical Physics* (McGraw-Hill, New York).

Pierce, A. P. (**1981**). *Acoustics—an Introduction to its Physical Principles and Applications* (McGraw Hill, New York).

Smirnov, V. I. (**1964**). *Course of Higher Mathematics* (Pergamon Oxford), Vol. 4.

Tsepelev, N. V. (**1975**). ''On some solutions of the Helmholmz equation,'' Zap. Nauchn. Semin. Leningr. otd. Mat. Inst. Akad. Nauk SSSR **51**, 497–500 [English transl.: J. Soviet Math. **11**(3), 197–200 (1979)].

Tsvankin, I. D., Kalinin, A. V., and Pivovarov, B. L. (**1983**). ''Refraction of a spherical wave at an interface closely adjacent to a surface,'' Izv. Akad. Nauk. SSSR Fiz. Zemli **10**, 32–45.

Tsvankin, I. D., and Kalinin, A. V. (**1984**). ''The wave effects at the generation of converted waves,'' Izv. Akad. Nauk SSSR Fiz. Zemli **2**, 34–40.

Yanovskaya, T. V., and Roslov, Yu. V. (**1987**). ''A Contribution of first order ray approximation to the wavefield reflected from a free surface of inhomogeneous halfspace,'' Vestn. Leningr. Univ. Ser. Fiz. Khim. **2**, 66–72.

Ray shooting, eigenray search, and complex resonances for submerged structures

I-Tai Lu and Y. Q. Yao
Department of Electrical Engineering/Weber Research Institute, Polytechnic University, Route 110, Farmingdale, New York 11735

(Received 7 February 1996; revised 5 July 1996; accepted 8 July 1996)

The general theory of a ray-mode approach for time-harmonic responses and a wavefront-resonant approach for transient responses of submerged structures [I-T. Lu, J. Acoust. Soc. Am. **99**, 114–132 (1996)] is applied to a prototype structure consisting of a cylindrical pipe, hemispherical endcaps, a bulkhead, and a rib. The acoustic scattering returns are synthesized systematically in terms of angular spectra of surface modes of the structural elements where bookkeeping of various spectral objects such as ray, collective ray, modes, ray mode, resonance, etc. can be maintained. Ray shooting and eigenray searching procedures for the guided modes of structural elements are derived. An approximate solution for the resonances of the global submerged structure is also obtained. Numerical examples of typical ray trajectories of these guided modes, travel times of wavefronts, and complex frequencies of resonances are presented. [S0001-4966(97)02501-0]

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

INTRODUCTION

As discussed in Ref. 1, a submerged structure can be considered as a collection of plate and shell elements when the bulk acoustic wavelengths are much smaller than the lateral dimensions of these elements. Each plate or shell element supports some guided modes. These modes propagate independently in an element but couple to one another at boundaries where this element joins other elements. A convenient analyzing strategy is to divide the structure into subregions where the waves in each subregion are treated separately and the coupling mechanism at each junction or joint[2,3] is also analyzed separately. The total response can then be obtained by combining these propagation results in each subregion and coupling factors for each junction systematically by employing a ray-type approach.[4–7] This approach can usually reduce the complexity of the problem greatly. A survey of this approach and a detailed list of recent publications have been given in Ref. 1.

The submerged structures are usually very complex in practical applications. Since it is practically impossible to compute the exact high-frequency responses of a complex structure, one always resorts to some simplified models. In many situations (especially in a noisy and time-varying environment), it is then much more useful to identify and compute some good physical observables than compute the total response of the simplified model by brute force. In other words, a characterization procedure should focus on a few robust, key observables with good physical interpretations. By good physical interpretation we mean that the observables should be linked to some set of spectral objects possessed of two important characteristics: (a) They are good propagators, in the sense that they retain their unique characters and remain identifiable as they propagate, and that they are insensitive to small perturbations of the model; and (b) they are efficient, in the sense that relatively few of them are required to characterize the entire scattered or radiated wave field. Examples of physical observables are travel times of rays in early times and complex frequencies of resonances in late times. Here, we examine the theory in Ref. 1 for evaluating wavefront travel times and resonant frequencies in detail by numerical examples.

Consider a submerged structure consisting of a truncated cylindrical pipe, two hemispherical endcaps, a bulkhead, and a ring (see Fig. 1). Its building elements are flat plates, spherical shells, and a cylindrical shell. The wave motion on each individual element (when it is without truncation) can be solved by the technique of separation of variables. But the structure in Fig. 1 is not separable. Since the structure has rotational symmetry about the z axis, the angular spectral decomposition with respect to the ϕ coordinate is applicable to all structural elements, and the overall response can be synthesized in terms of the ϕ spectra. Also, a local guided mode representation (where the modes in the endcaps, ring, pipe, and bulkhead are defined along the r coordinate, z coordinate, ρ coordinate, and z coordinate, respectively) satisfies the full elastic equations and boundary conditions at the surfaces of the plate or shell elements. Using a combination of guided modes and angular spectra, the problem is reduced to a ''one-dimensional'' problem in the remaining coordinate, which is ρ for the bulkhead and the ring, θ for the endcaps, and z for the pipe. Since the original structure is not separable, the local separability has been assumed here when the guided modes on shell/plate elements have been employed. For convenience, we will denote the common ϕ coordinate of all structural elements by ''w,'' the normal coordinate to the surface of each plate element by ''u,'' and the remaining (third) coordinate of each plate or shell element by ''v'' in later formulations and discussions.

The solution in the remaining coordinate ''v'' is not truly a one-dimensional problem because sources and receivers do not usually lie on the hull surface. However, we can project sources and receivers onto the surface of the hull for

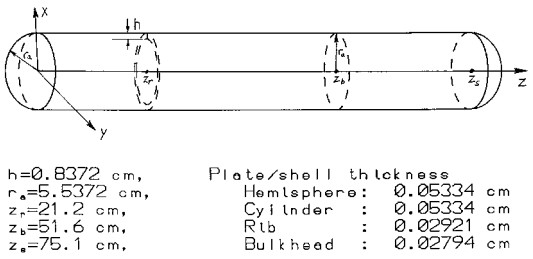


FIG. 1. A structure consists of a finite cylindrical pipe with a ring, a bulkhead, and two hemispherical endcaps. The parameters used for computation are also shown.

all possible excitation and observing mechanisms. Consider as an example a leaky mode of a fluid loaded shell element. With a given angular spectrum, the mode can radiate into the surrounded fluid (and vice versa) through a specific angle determined by the phase matching mechanism. Thus, any point on the surface of the shell element can be extended into the surrounding fluid following the radiation trajectory of the corresponding leaky mode. By reciprocity, sources and receivers in the fluid can be projected onto the hull. Then, the reduced problem in ''v'' coordinate becomes one dimensional. The essential feature of this formulation is that the system equations become algebraic. The solution to this reduced one-dimensional problem can be synthesized by traveling waves to include all possible coupling phenomena, by composite reflection and transmission matrices, by resonances,[8–12] or by a more flexible arrangement which retains some traveling fields, and account for the remaining ones collectively through use of modified collective reflection and transmission matrices and/or through use of some resonant modes.[13–15] These ideas have been developed in our previous work[1] and will not be repeated here. This paper will focus on the evaluation of travel times of wavefronts and complex frequencies of resonances. Simple, efficient, and robust procedures can be obtained in the reduced ''one dimensional'' domain. Section I describes the spectral approach for canonical configurations. Using these results, the original nonseparable structure is reduced to a one-dimensional problem in ''v'' coordinate. In Sec. II, ray shooting and eigenray searching procedures and global resonant conditions are formulated. Numerical examples of typical ray trajectories of these guided modes, travel times of wavefronts, and complex frequencies of resonances are presented in Sec. III. Concluding remarks are made in Sec. IV.

I. SPECTRAL APPROACH FOR SEPARABLE CONFIGURATIONS

Based on the local separability assumption, the classical spectral approach for separable configurations is employed as the basic building blocks of our approach. Here, the general considerations in Ref. 1 are now applied to two separable geometrical configurations, which can serve as prototype models for acoustic scattering from submerged structures; specifically, the spherical and cylindrical shells model endcaps and the hull, respectively, of the structure in Fig. 1. Without loss of generality, the discussion of Green's function formulation for thin plates is omitted because the bulkhead and the ring formed by plate elements do not contact with the surrounding fluid and, hence, only source-free solutions are required for the scattering problem. In order to highlight the fact that the procedure can be applied to all separable configurations, some of the results are listed in a tabular format. **The relevant formulas in Ref. 1 will not be duplicated here but will be referred to as (1.x) throughout this paper**.

The coordinate systems, three-dimensional wave equations, and boundary conditions for the cylindrical and spherical configurations are first shown in Table I. Note that ''a'' is the shell radius and ''Z'' is the shell impedance. The two one-dimensional problems in the (''u'', ''v'') coordinates and their solutions are summarized in Tables II and III, respectively. In the radius coordinate ''u'' (see Table II), the domain extends from the shell surface (radius a) to infinity. In high frequencies, the turning point is assumed to be less than the radius and hence the simple WKB solution in (1.26) is valid everywhere in the domain of interest: $a \leqslant u < \infty$. The reflection coefficient Γ_{u1} contains the resonant poles which

TABLE I. The coordinate system, wave equations, and boundary conditions for cylindrical and spherical shells.

Configuration / Coordinate		Cylindrical shell	Spherical shell
		$u=\rho,\ v=z,\ w=\phi$	$u=r,\ v=\theta,\ w=\phi$
3-D wave equations		$\left[\frac{1}{u}\frac{\partial}{\partial u}u\frac{\partial}{\partial u}+\frac{1}{u^2}\frac{\partial^2}{\partial v^2}+\frac{\partial^2}{\partial w^2}+k^2\right]G$ $=-\frac{1}{u}\delta(u-u')\delta(v-v')\delta(w-w')$	$\frac{1}{u^2}\frac{\partial}{\partial u}u^2\frac{\partial}{\partial u}G+\frac{1}{u^2\sin v}\frac{\partial}{\partial v}\sin v\frac{\partial}{\partial v}G$ $+\frac{1}{u^2\sin^2 v}\frac{\partial^2}{\partial w^2}G+k^2G$ $=-\frac{1}{u^2\sin v}\delta(u-u')\delta(v-v')\delta(w-w')$
Boundary conditions	u-coordinate	$G=Z\frac{\partial G}{\partial u}$ at $u_1=a$;	Radiation condition at $u_2=\infty$
	v-coordinate	Radiation condition at $v_{1,2}=\pm\infty$	Finiteness condition at $v_{1,2}=0,\pi$
	w-coordinate	Periodicity condition: $\frac{\partial G}{\partial w}=0$ at $\lvert w-w'\rvert=\pi$	

TABLE II. One-dimensional characteristic problem in the radius domain (u).

Configuration	Cylindrical shell ($u=\rho$)	Spherical shell ($u=r$)
Equation	$\left[\frac{d}{du}u\frac{d}{du}+(k^2-\lambda_v)u-\frac{\lambda_w}{u}\right]g_u=-\delta(u-u')$	$\left[\frac{d}{du}u^2\frac{d}{du}+k^2u^2-\lambda_v\right]g_u=-\delta(u-u')$
Boundary condition	$g_u=Z\frac{\partial g}{\partial u}$ at $u_1=a$; Radiation condition at $u_2=\infty$	
p and γ_u^2 in (1.23)	$p=u,\ \gamma_u^2=(k^2-\lambda_v)-(\lambda_w-0.25)/u^2$	$p=u^2,\ \gamma_u^2=k^2-\lambda_v/u^2$
Γ_{ui} in (1.25)	$\Gamma_{u2}=0;\quad \Gamma_{u1}=\begin{cases}-i\\ (i\gamma_u Z+1)/(i\gamma_u Z-1)\end{cases}$	if $\sqrt{\frac{\lambda_w-0.25}{k^2-\lambda_v}}\gtrless u_1$
u_{ti} in (1.25)	$u_{t1}=\max\left[\sqrt{\frac{\lambda_w-0.25}{k^2-\lambda_v}},u_1\right];\quad u_{t2}=\infty$	$u_{t1}=\max\left[\frac{\sqrt{\lambda_v}}{k},u_1\right];\quad u_{t2}=\infty$
Representation in (1.12) with $J=1$	$g_u^{(0)}=C_u\theta^{i\lvert\int_u^{u'}\gamma_u(\tau)d\tau\rvert},\quad R_u^{(1)}=C_u\Gamma_{u1}\theta^{i\{\int_{u_{t1}}^{u}+\int_{u_{t1}}^{u'}\}\gamma_u(\tau)d\tau}$	

correspond to guided waves on the structure (shell) surfaces.

In the lateral domain ''v'' (see Table III), the solution for the cylindrical configuration is exact. Regarding the spherical shell, there are two turning points and the WKB solution in (1.26) is valid between the two turning points but is not valid near the turning points or in the shadow regions. However, a uniform asymptotic approach using Airy function in these regions is straightforward[16] and is omitted here. Also, note that the scattering problem of a spherical shell insonified by a point source is intrinsically two dimensional because of the rotational symmetry about the axis passing through source and center of the structure. Here, we purposely formulate the problem in a three-dimensional format so that we can use these results in later sections for truncated spherical shells which can model endcaps. Because we are interested in high frequencies and far zones, the wave motion in this coordinate is better represented in terms of traveling waves, as in (1.9). For the cylindrical shell, (1.9) is reduced to the original form in (1.4) or in (1.26) because both reflection coefficients (Γ_{v1} and Γ_{v2}) are zero (see the first row of Table I in Ref. 1). For the spherical shell, $g_v^{(j)}$ can be represented in terms of the power series expansion in (1.9) and can further be decomposed into four ray species (denoted by the subscript α) as shown in the last row of Table III.

The one-dimensional problems in the azimuthal coordinate (''w'' or ϕ) for both cylindrical and spherical configurations are governed by the equation

$$\left(\frac{d^2}{dw^2}+\lambda_w\right)g_w=-\delta(w-w'), \tag{1}$$

and the periodicity condition

$$\frac{dg_w}{dw}=0 \quad \text{at} \quad |w-w'|=\pi. \tag{2}$$

TABLE III. One-dimensional characteristic problem in the lateral domain (v).

Configuration	Cylindrical shell ($v=z$)	Spherical shell ($v=\theta$)
Equation	$\left[\frac{d^2}{dv^2}+\lambda_v\right]g_v=-\delta(v-v')$	$\left[\frac{d}{dv}\sin v\frac{d}{dv}-\frac{\lambda_w}{\sin v}+\lambda_v\sin v\right]g_v=-\delta(v-v')$
Boundary condition	Radiation condition at $v_{1,2}=\pm\infty$	Finiteness of g_v at $v_{1,2}=0,\pi$
p and γ_v^2 in (1.23)	$p=1,\ \gamma_v^2=\lambda_v$	$p=\sin v,\ \gamma_v^2=\lambda_v+0.25-(\lambda_w-1/4)/\sin^2 v$
Γ_{vi} in (1.25)	$\Gamma_{v1}=\Gamma_{v2}=0$	$\Gamma_{v1}=\Gamma_{v2}=-i$
v_{ti} in (1.25)	$v_{t1}=-\infty,\ v_{t2}=\infty$	$v_{t1}=\theta_t,\quad v_{t2}=\pi-\theta_t;\quad \theta_t\equiv\sin^{-1}\sqrt{\frac{\lambda_w-0.25}{\lambda_v+0.25}}$
Representation in (1.9) $g_v=\sum_{m=0}^{\infty}g_m^{(m)}$ $g_v^{(m)}=\sum_{\alpha=1}^{4}g_{v\alpha}^{(m)}$ C_v and Λ_v are in (1.26)	$g_v=-\frac{\theta^{i\sqrt{\lambda_v}\lvert v-v'\rvert}}{2i\sqrt{\lambda_v}}$	$g_{v1}^{(m)}=C_v(-i)^{2m}\exp i\left[\left\lvert\int_v^{v'}\gamma_v\,d\tau\right\rvert+2m\Lambda_v\right]$ $g_{v2}^{(m)}=C_v(-i)^{2m+1}\exp i\left[\left\{\int_{v_{t1}}^{v}+\int_{v_{t1}}^{v'}\right\}\gamma_v\,d\tau+2m\Lambda_v\right]$ $g_{v3}^{(m)}=C_v(-i)^{2m+1}\exp i\left[\left\{\int_{v}^{v_{t2}}+\int_{v}^{v_{t2}}\right\}\gamma_v\,d\tau+2m\Lambda_v\right]$ $g_{v4}^{(m)}=C_v(-i)^{2(m+1)}\exp i\left[-\left\lvert\int_v^{v'}\gamma_v\,d\tau\right\rvert+2(m+1)\Lambda_v\right]$

TABLE IV. The three-dimensional Green's function G.

Configuration	Cylindrical shell	Spherical shell
Representation	$G=\sum_{n=0}^{\infty}\sum_{\beta=1}^{2}[G_{\beta}^{(n)}+R_{\beta}^{(n)}]$	$G=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\sum_{\alpha=1}^{4}\sum_{\beta=1}^{2}[G_{\alpha\beta}^{(mn)}+R_{\alpha\beta}^{(mn)}]$
Magnitudes in (1.27) $\begin{bmatrix}A^{(\mathscr{E})}\\ B^{(\mathscr{E})}/D\end{bmatrix}$	$\begin{bmatrix}A_{\beta}^{(n)}\\ B_{\beta}^{(n)}/D\end{bmatrix}=\frac{C_u}{16\pi^2\sqrt{\lambda_v\lambda_w}}\begin{bmatrix}1\\ \Gamma_{u1}\end{bmatrix}$	$\begin{bmatrix}A_{\alpha\beta}^{(mn)}\\ B_{\alpha\beta}^{(mn)}/D\end{bmatrix}=\frac{-i(-1)^m C_u C_v\delta_{\alpha}}{8\pi^2\sqrt{\lambda_w}}\begin{bmatrix}1\\ \Gamma_{u1}\end{bmatrix},\quad \delta_{\alpha}=\begin{cases}1, & \alpha=1\\ -i, & \alpha=2,3\\ -1, & \alpha=4\end{cases}$
Phases in (1.27) $\begin{bmatrix}\varphi^{(\mathscr{E})}\\ \psi^{(\mathscr{E})}\end{bmatrix}$ with $\zeta_{\beta}=\begin{cases}2\pi-\lvert w-w'\rvert, & \beta=1\\ \lvert w-w'\rvert, & \beta=2\end{cases}$	$\begin{bmatrix}\varphi_{\beta}^{(n)}\\ \psi_{\beta}^{(n)}\end{bmatrix}=\begin{bmatrix}\left\lvert\int_u^{u'}\gamma_u\,d\tau\right\rvert\\ \left\{\int_a^{u}+\int_a^{u'}\right\}\gamma_u\,d\tau\end{bmatrix}+\sqrt{\lambda_v}\lvert v-v'\rvert+\sqrt{\lambda_w}(2n\pi+\zeta_{\beta})$	$\begin{bmatrix}\varphi_{\alpha\beta}^{(mn)}\\ \psi_{\alpha\beta}^{(mn)}\end{bmatrix}=\begin{bmatrix}\left\lvert\int_u^{u'}\gamma_u\,d\tau\right\rvert\\ \left\{\int_a^{u}+\int_a^{u'}\right\}\gamma_u\,d\tau\end{bmatrix}+[2m\Lambda_v+\xi_{\alpha}]+[\sqrt{\lambda_w}(2n\pi+\zeta_{\beta})]$ $\xi_{\alpha}=\begin{cases}\left\lvert\int_v^{v'}\gamma_v\,d\tau\right\rvert, & \alpha=1\\ \left\{\int_{v_{t1}}^{v}+\int_{v_{t3}}^{v'}\right\}\gamma_v\,d\tau, & \alpha=2\\ \left\{\int_{v}^{v_{t2}}+\int_{v'}^{v_{t2}}\right\}\gamma_v\,d\tau, & \alpha=3\\ 2\Lambda_v-\left\lvert\int_v^{v'}\gamma_v\,d\tau\right\rvert, & \alpha=4\end{cases}$

The equation in (1) has already been in the standard form as shown in (1.22), and its exact solution is

$$g_w=-\frac{\cos\sqrt{\lambda_w}(\pi-|w-w'|)}{2\sqrt{\lambda_w}\,\sin(\pi\sqrt{\lambda_w})}. \tag{3}$$

If the wave motion in this coordinate is represented in terms of traveling waves as in (1.9), the Green's function can be decomposed into two ray species denoted by the subscript β:

$$g_w=\sum_{n=0}^{x}g_w^{(n)},\quad g_w^{(n)}=\sum_{\beta=1}^{2}g_{w\beta}^{(n)},$$

$$g_{w1}^{(n)}=\frac{i}{2\sqrt{\lambda_w}}\,e^{i\sqrt{\lambda_w}[(2n+1)\pi-|w-w'|]}, \tag{4}$$

$$g_{w2}^{(n)}=\frac{i}{2\sqrt{\lambda_w}}\,e^{i\sqrt{\lambda_w}[2n\pi+|w-w'|]}.$$

Using the results from Ref. 1, the amplitudes and phases in (1.27) for the ray and the remainder integrals can be obtained and are summarized in Table IV. Here, the symbol $(\mathscr{E})$ representing all possible subscripts such as ''jmn'' in (1.27) can then be clearly defined here. First, the superscript (j) for the ''u''-coordinate traveling wave representation in (1.17) is deleted because only one $g_u^{(i)}$ term is required (also see Table II). Second, the superscript (m) for the ''v''-coordinate traveling wave representation in the cylindrical shell (but not in the spherical shell) is also deleted because no ray expansion is needed. Third, the index n in the ''w'' coordinate remains intact for both cylindrical and spherical shells as shown in Eq. (4). Finally, in addition to the superscripts (n) or (mn), we need the subscripts α to describe different ray species in the ''v'' coordinate for the spherical shell and the subscript β to describe different ray species in the ''w'' coordinates for both configurations (see Table IV).

TABLE V. The mode wave numbers.

Modes	Spherical shell (ξ)	Cylindrical shell (k_z)	Rib (k_ρ)	Bulkhead (k_ρ)
Compressional wave	6.387 040 6 $+j$0.050 068 3	1.158 659 9 $+j$0.000 580 7	1.159 711 7 $+j$0.0	1.159 710 8 $+j$0.0
Torsional wave	11.035 309 6 $+j$0.0	1.974 405 3 $+j$0.0	1.974 405 3 $+j$0.0	1.974 405 3 $+j$0.0
Bending wave	51.608 295 1 $+j$0.0	9.141 861 6 $+j$0.0	11.850 343 5 $+j$0.0	12.111 225 3 $+j$0.0
First leaky bending wave	2.618 998 7 $+j$48.419 589	0.475 081 3 $+j$8.695 920 3	0.0 $+j$11.603 856 6	0.0 $+j$11.870 156 9

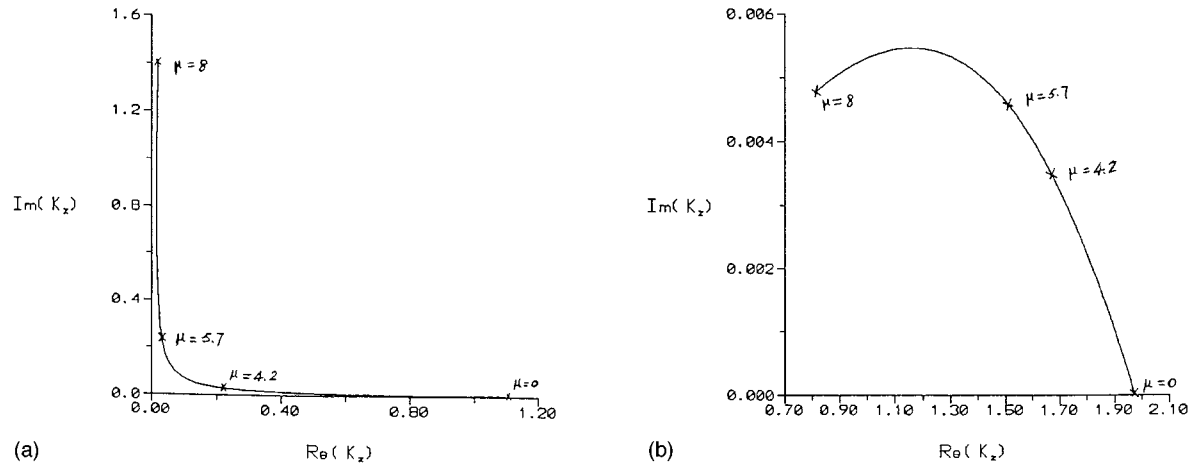


FIG. 2. The loci of the modal wave number k_z of a fluid loaded cylindrical shell as the azimuthal wave number μ changing from 0 to 8: (a) Compressional wave, (b) torsional wave.

Note that the resonant denominator D in (1.27) is the denominator of Γ_{u1} (i.e., $i\gamma_u Z-1$).

II. FORMULATIONS

Figure 1 shows the structure and relevant parameters used for numerical computation. The structure consists of a finite cylindrical pipe with a ring, a bulkhead, and two hemispherical endcaps. In this section, we will show some typical ray trajectories, wavefront travel times, and complex resonant frequencies.

A ray trajectory can be specified by its phase term. A typical phase term [see (1.28)] is

$$\varphi=\varphi_w+\varphi_v+\varphi_u, \quad \varphi_w\equiv\mu\Delta\phi,$$

$$\varphi_v\equiv\int_{\text{pipe}}\gamma_v\,dv+\int_{\text{endcaps}}\gamma_v\,dv+\int_{\text{bulkhead}}\gamma_v\,dv+\int_{\text{rib}}\gamma_v\,dv, \tag{5}$$

$$\varphi_u\equiv\int_{\text{source}}\gamma_u\,du+\int_{\text{receiver}}\gamma_u\,du,$$

where φ_I, $I=u,v,w$, denotes the phase changes in the ''I'' coordinate. Here, γ_u and γ_v are the wave number in the u and v coordinates, respectively, and μ is the wave number in the common azimuthal coordinate ''ϕ'' or ''w.'' The total increment of the azimuthal angle is denoted by $\Delta\phi$. The lower limits of the ''v'' integrals in Eqs. (5) denote the projection of all possible ray segments onto the ''v'' coordinate of the related structure elements. The lower limits (''source'' and ''receiver'') of the ''u'' integrals denote the ''u'' projection of ray segments in water for excitation and reradiation, respectively. The wave number γ_u and γ_v in the integrands are chosen according to their corresponding modes and structure elements.

The eigenray search is furnished by solving the following stationary phase condition:

$$\frac{d\varphi}{d\mu}=0, \quad \text{at } \mu=\mu_s, \quad \gamma_v=\gamma_{vs}, \quad \gamma_u=\gamma_{us}. \tag{6}$$

Note that the mode wave number γ_v of a corresponding structure mode is a function of the azimuthal spectral variable μ. Thus, if μ is fixed, the corresponding γ_v's are fixed and the corresponding radius wave number γ_u is also determined. In other words, the ray phase in Eqs. (5) is a function of the spectral variable μ which can be considered as the ray parameter. The stationary phase condition in Eq. (6) determines the unique ray parameter μ_s and hence the eigenray if the double derivative of the phase φ with respect to μ at μ_s is

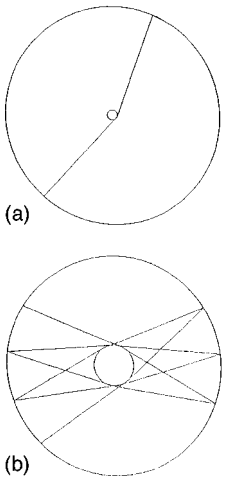


FIG. 3. Two typical ray trajectories on the bulkhead: (a) μ=0.3; (b) μ=1.2.

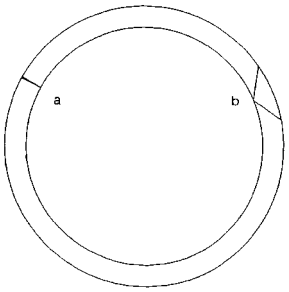


FIG. 4. Two typical ray trajectories on the ring: (a) μ=0.2; (b) μ=4.7.

not equal to zero. (If the double derivative of the phase φ is also zero but the triple derivative of the phase φ is not zero, the receiver is located at a caustics.) In the eigenray searching procedure, any root searching routine can be employed to find the ray parameter μ by solving Eq. (6). The travel time of a ray arrival (wavefront) can be obtained by differentiating the phase term in Eqs. (5) with respect to the angular frequency ω when the ray parameter μ is chosen to satisfy the stationary phase condition in Eq. (6).

A ray tracing procedure can also be obtained based on the stationary phase condition in Eq. (6). In the ray tracing procedure, the ray parameter μ is specified, and the ray trajectory is yielded by solving for the increments in ''u'' and ''v'' coordinates with a small increment in ''ϕ'' coordinate. Consider a small increment $\Delta\phi$ of the azimuthal angle for a given ray parameter μ. The increments in the u and v coordinates (i.e., Δu and Δv) can then be related to $\Delta\phi$ and μ by Eq. (6):

$$\Delta\phi = -\frac{d\gamma_v}{d\mu}\,\Delta v - \frac{d\gamma_u}{d\mu}\,\Delta u. \tag{7}$$

There are two unknowns (Δu and Δv) in Eq. (7), and we need an additional equation to determine the two unknowns. If the ray trajectory lies on the surface of a structure element, $\Delta u = 0$ and Δv are uniquely determined by Eq. (7). If the ray is about to take off from the structure element into the surrounding fluid, or about to couple into the structure element from the fluid, the phase matching condition

$$\frac{\gamma_v}{\Delta v} = \frac{\gamma_u}{\Delta u} \tag{8}$$

provides the second equation to solve the two unknowns. Although Eqs. (7) and (8) may be employed to trace rays further away from the structure, we prefer to use a ray tracing procedure based on the rectangular coordinate system in far zones. This is due to the fact that the wave numbers γ_u and γ_v are obtained based on WKB approximations in a cylindrical or spherical coordinate system.

The complex frequencies of the structure in Fig. 1 can be obtained by solving (1.61) or (1.83). Here, we will show the results of a further simplified structure of Fig. 1: a cylindrical pipe with two hemispherical endcaps. The junctions between the pipe and endcaps are assumed to be perfectly welded. Hence, the various coupling matrices in (1.54) and (1.55) becomes

$$\begin{aligned} &[\mathbf{t}_p^{U,D}] = [I], \quad p = 1,2,3,4,5, \\ &[\mathbf{r}_p^{U,D}] = [0], \quad p = 2,3,4, \\ &[\mathbf{r}_0^{U}] = [\mathbf{r}_5^{D}] = -i[I], \end{aligned} \tag{9}$$

where $[I]$ is the identity matrix and $[0]$ is the zero matrix. The ''$-i$'' in Eqs. (9) is the reflection coefficient at caustics. Because there is no mode coupling, the resonant condition is simplified as

$$\int_{\Theta}^{\pi-\Theta} \gamma_{v,\text{sphere}}\, dv + \int_0^L \gamma_{v,\text{pipe}}\, dv = \left(n + \frac{1}{2}\right)\pi, \tag{10}$$

where Θ and $\pi - \Theta$ are the turning points in the left and right spherical endcaps, respectively. The two γ_v's in Eq. (10) have to be related to the same mode, e.g., the compressional mode, etc.

III. NUMERICAL EXAMPLES

Consider the structure in Fig. 1. In our numerical examples, the bulk compressional wave and shear wave speeds of the structure material are 5719 and 3001 m/s, respectively. The sound speed in the surrounding fluid is 1488 m/s. The frequency used for computation is 94.3 kHz. Thus, the corresponding wave numbers of bulk compressional wave (k_p), bulk shear wave (k_s), and pressure wave in water (k_w) are 1.0545, 1.9918, and 3.9811 cm^{-1}, respectively. The spectral eigenvalues in the ''v'' coordinate of the fundamental guided modes on the structure elements are listed in Table V when $\mu = 0$. Here, $\xi = (\lambda_v + 0.25)^{1/2}$. The parameters ξ, k_ν and k_ρ are the wave numbers on the ''v'' coordinate for endcaps, pipe, and bulkhead/ring, respectively. Figure 2(a) and (b) show, respectively, in the complex k_z plane the loci of the compressional and torsional wave numbers of a fluid loaded cylindrical shell as the azimuthal wavenumber μ changing from 0 to 8. Note that the scale of the imaginary part is exaggerated in Fig. 2(b).

We will first show ray trajectories [using ray shooting Eqs. (7) and (8)] in each individual structure element. Figure 3 shows two typical ray trajectories on the bulkhead. Figure 4 shows two typical ray trajectories on the ring. The outer circles in Figs. 3 and 4 represent the joints where the bulkhead and the ring join the cylindrical pipe, respectively. The inner circles in Fig. 3 denote the traces of turning points for the corresponding azimuthal wave numbers (μ's). But the inner circle in Fig. 4 denotes the edge of the ring. For convenience, let the source and receiver be located at the junction where the bulkhead or the ring joins the cylindrical pipe (i.e., the radius coordinates of both source and receiver are 5.5372 cm). The associated turning point is drawn as a circle with radius R in Fig. 3 but not in Fig. 4. The total increment of the azimuthal angle is denoted by $\Delta\phi$. Figure 3(a) shows a ray with small μ (0.3). The ray emanated from the source propagates inwardly almost along the radius direction. It is reflected by the caustics (the corresponding turning point is denoted by $R = 0.2587$ cm) and propagates outwardly until reaching the receiver. The azimuthal angle proceeded by this ray is 198°. Figure 3(b) shows the ray with the azimuthal wave number equal to 1.2. The associated turning point is a circle with radius equal to $R = 1.035$ cm. The azimuthal angle proceeded by this ray is 1001.5°. Here, we let the ray bounce back and forth in between the joint and the caustics three complete cycles.

In Fig. 4, the first ray denoted by ''a'' has a small azimuthal wave number: $\mu = 0.2$. Hence, the ray propagates inwardly almost along the radius direction and reflected at the edge (truncation) of the ring. The azimuthal angle proceeded by this ray is 1.28°. Note that the associate turning point is 0.17 cm, which is smaller then the radius of the truncation of the ring, i.e., 4.7 cm. Therefore, the ray is reflected by the truncation before reaching the turning point. The second ray denoted by ''b'' has a large azimuthal wave number: $\mu = 4.7$.

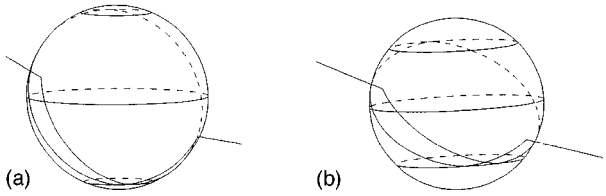


FIG. 5. Two typical ray trajectories on a spherical shell: (a) $(r',\theta',\phi')=(8.11\text{ cm},\ 70°,\ 0°)$, $(r,\theta,\phi)=(8.141\text{ cm},\ 110°,\ 501.2°)$, $\Theta=23.04°$, $\mu=2.5$; (b) $(r',\theta',\phi')=(10.16\text{ cm},\ 70°,\ 0)$, $(r,\theta,\phi)=(10.64\text{ cm},\ 110°,\ 491.8°)$, $\Theta=47.37°$, $\mu=4.7$.

The ray propagates inwardly and is reflected at the truncation of the ring. Note that the associate turning point is 4.05 cm which is still smaller than the radius of the truncation of the ring, i.e., 4.7 cm. The ray cannot reach the turning point. The azimuthal angle proceeded by this ray is 25.05°.

Figure 5 shows two typical ray trajectories on a fluid loaded spherical shell. The source and receiver coordinates are denoted by (r',θ',ϕ') and (r,θ,ϕ), respectively. With each ray, the two associated turning points are specified by Θ and π-Θ, which are also shown in Fig. 5 by two circles on the top and the bottom portions of the sphere, respectively. The circles at the equator are not part of the ray trajectories but only shown here for providing a three-dimensional view of the sphere. For a smaller azimuthal wave number μ, the elevation wave number (along the unit vector $\mathbf{a}_\vartheta$) becomes larger and the angle Θ of the turning point becomes smaller. That means a ray with a smaller azimuthal wave number μ is more likely to move up and down along the elevation direction than a ray with a larger azimuthal wave number μ, and vice versa. This phenomenon can be seen clearly by comparing Fig. 5(a) with Fig. 5(b). Figure 6 shows two typical ray trajectories on an infinite cylindrical shell. The source and receiver coordinates are denoted by (ρ',z',ϕ') and (ρ,z,ϕ), respectively. For a smaller azimuthal wave number μ, the longitudinal wave number (along z coordinate) becomes larger. That means a ray with a smaller azimuthal wave number μ is more likely to propagate along the axis of the cylinder than a ray with a larger azimuthal wave number μ, and vice versa [comparing Fig. 6(a) with Fig. 6(b)].

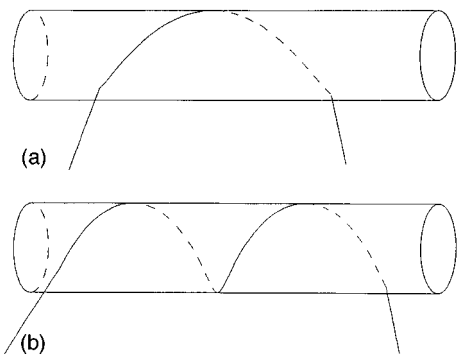


FIG. 6. Two ray trajectories on an infinite cylindrical pipe: (a) $(\rho',z',\phi')=(11.64\text{ cm},\ 55.5\text{ cm},\ 0°)$, $(\rho,z,\phi)=(11.64\text{ cm},\ 13.50\text{ cm},\ 320.6°)$, $\mu=3.7$; (b) $(\rho',z',\phi')=(10.45\text{ cm},\ 65.5\text{ cm},\ 0°)$, $(\rho,z,\phi)=(12.94\text{ cm},\ 6.5\text{ cm},\ 674.9°)$, $\mu=4.7$.

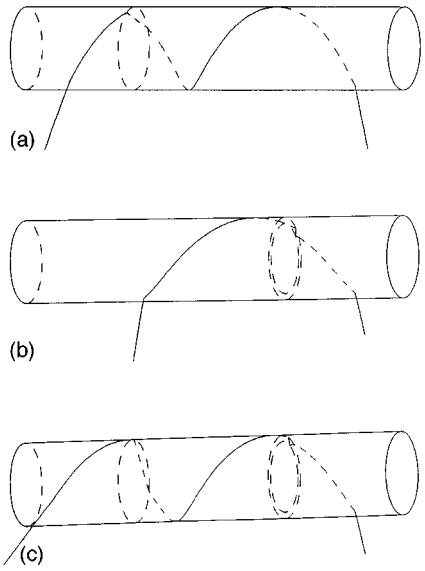


FIG. 7. Three ray trajectories on an infinite cylindrical pipe with an ring and/or an bulkhead: (a) $(x',y',z')=(-3.126\text{ cm},\ 12.28\text{ cm},\ 65\text{ cm})$, $(x,y,z)=(3.102\text{ cm},\ 12.28\text{ cm},\ 9.5\text{ cm})$, $\mu=4.55$; (b) $(x',y',z')=(-3.792\text{ cm},\ 11.0\text{ cm},\ 65.5\text{ cm})$, $(x,y,z)=(2.268\text{ cm},\ 11.42\text{ cm},\ 23.5\text{ cm})$, $\mu=3.7$; (a) $(x',y',z')=(-3.454\text{ cm},\ 11.66\text{ cm},\ 56.6\text{ cm})$, $(x,y,z)=(9.005\text{ cm},\ 13.95\text{ cm},\ 5.5\text{ cm})$, $\mu=4.2$.

Having shown the ray trajectories in each individual structure elements, one can employ the principle of local separability and the coupling rules of waves at joints to derive ray trajectories of nonseparable structures. Figure 7 shows three typical ray trajectories on an infinite cylindrical pipe with an ring and/or an bulkhead. The ray emanated from the source in fluid excites the compressional wave of the cylinder. When the wave hits the joint of bulkhead or ring, the energy will be reflected, transmitted, and coupled into the internal structure elements. The energy propagating inside the internal structure elements will eventually coupled back to the fluid loaded shell and then radiate into the surrounding fluid. Figure 7(a), (b), and (c) show rays coupled into the bulkhead only, the ring only, and both the bulkhead and the ring, respectively. Figure 8 shows one typical ray trajectory on the complete structure in Fig. 1. The ray excites the compressional wave at the left end cap, then propagates through the cylindrical shell, bulkhead, and ring, and finally takes off at the joint of the right end cap. In Figs. 7 and 8, the source and receiver coordinate are denoted in a global rectangular coordinate system by (x',y',z') and (x,y,z), respectively. In all cases, the ray parameter is specified by μ.

Figure 9 shows a source–receiver arrangement for com-

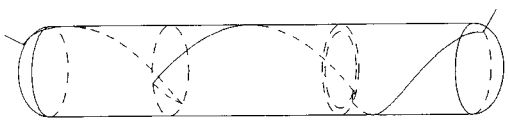


FIG. 8. A typical ray trajectory on the structure in Fig. 1: $(x',y',z')=(2.124\text{ cm},\ -7.288\text{ cm},\ 80.41\text{ cm})$, $(x,y,z)=(7.808\text{ cm},\ -3.840\text{ cm},\ -6.322\text{ cm})$, $\mu=3.75$.

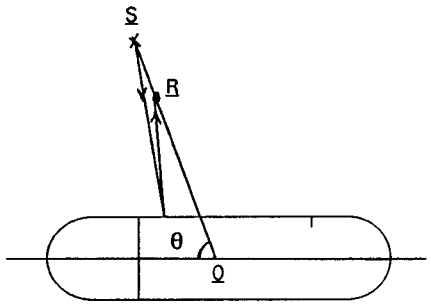


FIG. 9. A source–receiver arrangement for computing the travel times of wavefronts. The distance between the source and the center of the structure is 180 in., and the distance between the source and the center of the structure is 123 in.

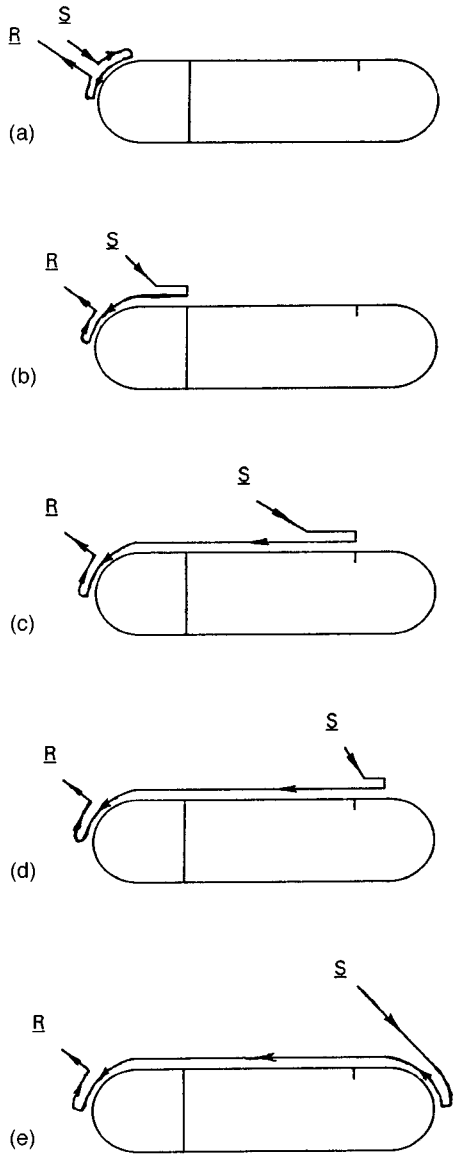


FIG. 10. The five possible excitation and re-radiation arrangements of the first ray species with $0°<\theta<90°$. The corresponding ray travel times are listed in Table VI. These rays are chosen to have gone through two reflections. One reflection occurs at the caustics of the left endcap. The other reflection occurs at the joints between the left endcap and the pipe, between the pipe and the bulkhead, between the pipe and the ring, and between the pipe and the right endcap, respectively, for cases *a*, *b*, *c*, and *d*, respectively. But for, case *e*, the second reflection occurs at the caustics of the right endcap.

TABLE VI. Travel times corresponding to the rays shown in Fig. 10.

Species I θ	Ray	
	t (μs)	Fig.
18°	58.2	
24°	78.8	
30°	104.7	
36°	135.5	
42°	170.9	10(a)
48°	210.4	
54°	253.3	
60°	299.3	
66°	347.6	
68°	364.1	
70°	464.9	10(b)
72°	596.9	10(c)
76°	696.5	10(d)
78°	732.3	
80°	732.6	10(e)
84°	733.1	
90°	733.2	

puting the travel times of wavefronts which are derived by eigenray search using Eq. (6). The distance between the source and the center of the structure is 180 in. and the distance between the receiver and the center of the structure is 123 in. When the acoustic wave couples to the structure wave, the incident wave can propagate either clockwise or counterclockwise along the hull. Similarly, when the wave takes off from the structure and couples energy into the surrounding fluid, the re-radiate wave can also propagate either clockwise or counterclockwise along the hull. Therefore, we can define four ray species. Let the rays of the first species have both incident and re-radiated waves propagating clockwise. The excitation can occur at the right endcap, the cylindrical pipe, or the left endcap. Similarly, the re-radiation can occur at the right endcap, the cylindrical pipe, or the left endcap. Figure 10 shows all possible combinations of excitations and re-radiations of the first species when the angle θ

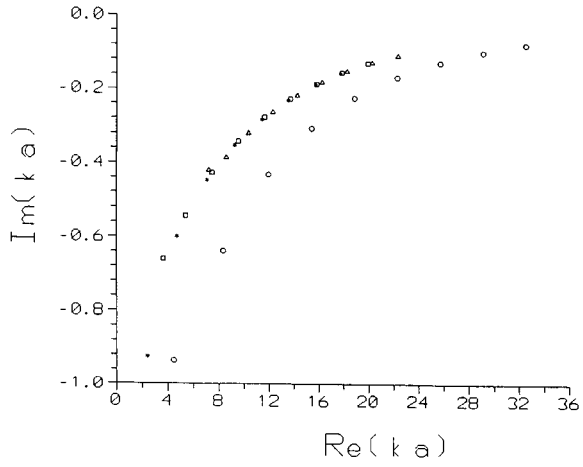


FIG. 11. The Lamb's poles for $L/2a=1$. The poles associated with $\mu=0$, 1, and 2 are denoted by asterisk, square, and triangle, respectively. The case for a spherical shell ($L=0$) is denoted by circle. The index ''*n*'' in Eq. (10) is listed in Table VII.

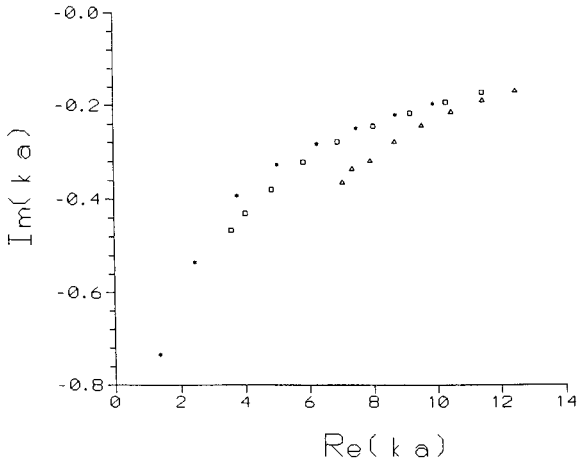


FIG. 12. The Lamb's poles for $L/2a=3$. The poles associated with $\mu=0$, 1, and 2 are denoted by asterisk, square, and triangle, respectively. The index ''n'' in Eq. (10) is listed in Table VII.

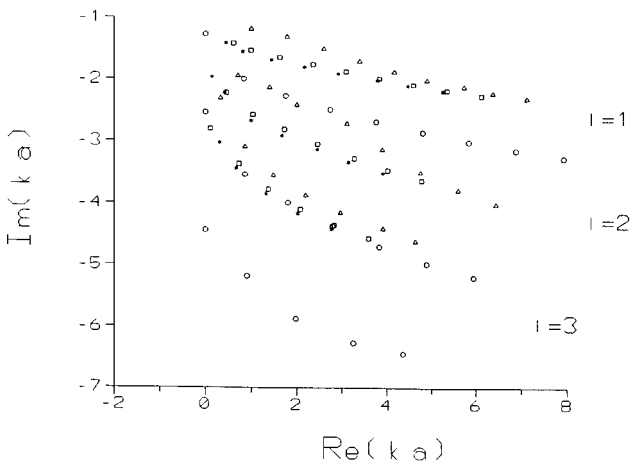


FIG. 14. The Franz poles for $L/2a=1$. The poles associated with $\mu=0$, 1, and 2 are denoted by asterisk, square, and triangle, respectively. The case for a spherical shell ($L=0$) is denoted by circle. The index ''n'' in Eq. (10) is listed in Table VII.

in Fig. 9 varies from 0° to 90°. Without loss of generality, we will consider motion of the compressional waves on the hull only. The reverberations of waves between joints are also not considered here. The travel times corresponding to the rays in Fig. 10 are shown in Table VI. For convenience, we show only the results with $\mu=0$. The travel times of other rays can be computed accordingly.

The complex frequencies of a cylindrical pipe with two hemispherical endcaps are computed using Eq. (10). Following the terminologies in Ref. 9 for a spherical shell, the Lamb's poles for $L/2a=1$, 3, and 6.673 are shown in Figs. 11, 12, and 13, respectively. Also, the Franz poles for $L/2a=1$, 3, and 6.673 are shown in Figs. 14, 15, and 16, respectively. Here, ''L'' is the length of the cylindrical pipe and ''a'' is the radius of the exterior surface of the cylinder (or the endcaps). The radius of the interior surface is $0.9904a$. The Lame's parameters for the material are $\lambda=1.2432\times10^{12}$ dyn/cm^2, $\mu=0.7929\times10^{12}$ dyn/cm^2, and the density $\rho=8.96$ g/cm^3. In Figs. 11–16, the poles associated with $\mu=0$, 1, and 2 are denoted by an asterisk, square, and triangle, respectively. The case of a spherical shell ($L=0$) is also generated and denoted by circle in Figs. 11 and 14. Note that the poles corresponding to different azimuthal spectral variable μ degenerate for the spherical shell because of the spherical symmetry. In Figs. 11–16, the index ''n'' corresponds to the standing wave pattern specified by the ''n'' in (10) and is listed in Table VII. The index ''l'' in Figs. 11–16 corresponds to various guided modes in the structure elements. From Figs. 11–16, we observe that larger ''L'' implies smaller resonant frequencies. Also, larger ''L'' causes larger separations among modes with different azimuthal order μ.

IV. CONCLUSION

In the hybrid ray-mode (wavefront-resonance) approach,[1] the acoustic scattering returns are synthesized systematically in terms of spectra of surface modes of the structural elements where bookkeeping of various spectral objects can be maintained. This formulation provides a uni-

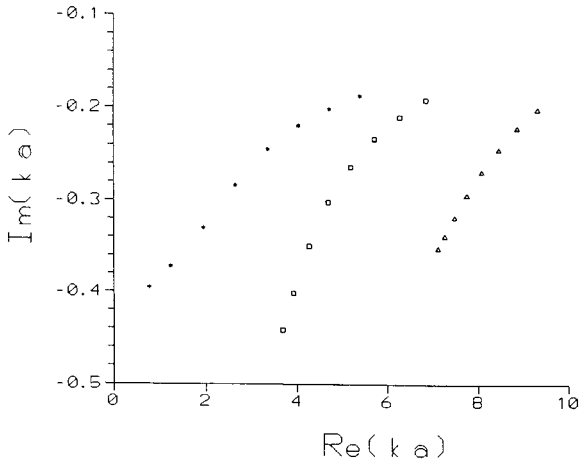


FIG. 13. The Lamb's poles for $L/2a=6.673$. The poles associated with $\mu=0$, 1, and 2 are denoted by asterisk, square, and triangle, respectively. The index ''n'' in Eq. (10) is listed in Table VII.

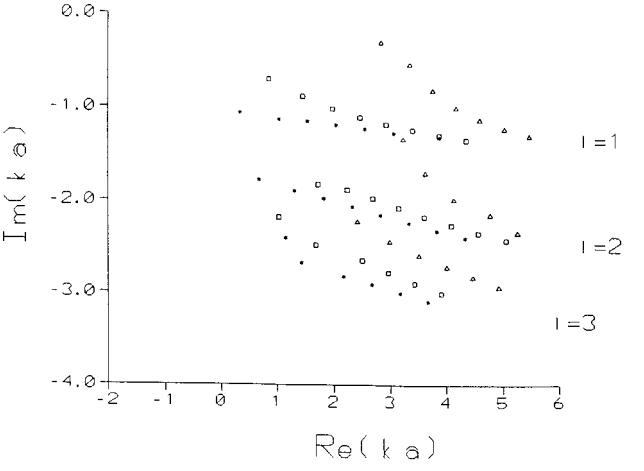


FIG. 15. The Franz poles for $L/2a=3$. The poles associated with $\mu=0$, 1, and 2 are denoted by asterisk, square, and triangle, respectively. The index ''n'' in Eq. (10) is listed in Table VII.

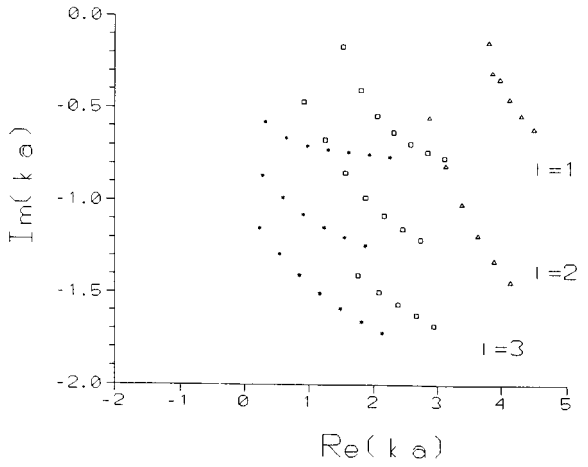


FIG. 16. The Franz poles for $L/2a=6.673$. The poles associated with $\mu=0$, 1, and 2 are denoted by asterisk, square, and triangle, respectively. The index ''n'' in Eq. (10) is listed in Table VII.

fied and systematic approach for deriving mode, ray, spectral integral, ray-mode, collective ray, eigenray–eigenmode, wavefront resonance, etc. in both time-harmonic and time-dependant domains. To understand the physical phenomena, we can single out any specific physical process or treat a particular class of processes collectively. In this paper, the general theory is employed to analyze a canonical structure consisting of a cylindrical pipe, two hemispherical endcaps, a bulkhead, and a ring. Using a combination of guided modes and angular spectra, the problem is reduced to a ''one-dimensional'' problem. The essential feature of this formulation is that the system equations become algebraic. It implies the local separability. Therefore, simple procedures can be derived for computing two good physical observables: global resonant frequencies and local ray travel times. Numerical examples show that the approach is very efficient and robust.

We have shown numerical results of ray trajectories of guided modes, travel times of wavefronts, and complex frequencies of resonances of the canonical structure in Fig. 1. Generally speaking, the travel time of the wavefront along a ray can be simply determined by elementary physics if the trajectory is defined. However, to find the ray trajectory is nontrivial and is usually based on a trial-and-error procedure in conventional approaches. This kind of conventional approach can become very inefficient and does not always guarantee convergence. In our approach, the ray trajectory of a eigenray is determined systematically by solving the stationary phase condition in Eq. (6). Moreover, the difficulty of ray proliferation is overcome because systematic bookkeeping and alternative representations are available. The complex resonant frequencies for an elongated cylindrical shell with hemispherical endcaps are, to the best of our knowledge, derived for the first time by a ray approach. A longer cylindrical pipe implies smaller resonant frequencies and also causes larger separations among resonances with different azimuthal order μ.

As mentioned in Ref. 1, the accuracy of our results depends on the validity of the following three subapproaches of the problem: (a) modal solutions for each plate or shell component; (b) the ray method employed; and (c) coupling at junctions. In our numerical examples, condition (a) is satisfied because each component of the structure in Fig. 1 is separable if it is without truncations. Condition (c) is also satisfied because of similar reasons, and a detailed work has been published in Ref. 2. Regarding condition (b), it is well known that the ray approach is valid in high-frequency regimes. To quantify these parameter regimes, one has to compare ray results with benchmark results using canonical examples such as spherical shells where rigorous approaches exist. From our experiences, the field magnitude result is accurate within 5% error when $ka>5$. However, the ray travel time result is accurate within 5% error when $ka>3$. A lot of times, accurate travel time results can be obtained even with $ka<3$. Although the results in Figs. 11–16 are in the low-frequency range of this type of structure, the results is quite accurate except for very small ka. For the spherical shell (see Figs. 11 and 14), the problem is separable and our approach is rigorous. For the elongated cylindrical shell (see Figs. 11, 12, 15, and 16), the resonant frequencies obtained by the ray-type approach are accurate within 5% error when $ka>3$ using a conservative assessment. Note that, in a low-frequency range, computation of resonant frequencies for this canonical structure can be made using pure numerical methods, e.g., the Helmholtz integral equation based boundary element–finite element method wherein the fluid-structure system matrix is set up and the complex eigenvalues can be accurately extracted. The present method also sets up its Matrix–Green's function and extracts the eigenfre-

TABLE VII. The index n for Figs. 11–16.

		Fig. 11 ($L=0$)	Fig. 11 ($L=2a$)	Fig. 12	Fig. 13	Fig. 14 ($L=0$)	Fig. 14 ($L=2a$)	Fig. 15	Fig. 16
	$\mu=0$		$n=2$–9	$n=2$–9	$n=1$–9		$n=2$–9	$n=3$–9	$n=3$–9
$l=1$	$\mu=1$	$n=1$–9	$n=2$–9	$n=1$–9	$n=1$–9	$n=1$–9	$n=1$–9	$n=2$–9	$n=3$–9
	$\mu=2$		$n=1$–9	$n=1$–9	$n=1$–9		$n=1$–9	$n=3$–9	$n=4$–9
	$\mu=0$						$n=3$–9	$n=5$–11	$n=4$–9
$l=2$	$\mu=1$					$n=4$–9	$n=3$–9	$n=4$–11	$n=4$–9
	$\mu=2$						$n=2$–9	$n=5$–9	$n=6$–11
	$\mu=0$						$n=5$–9	$n=7$–11	$n=5$–11
$l=3$	$\mu=1$					$n=5$–9	$n=4$–9	$n=6$–11	$n=5$–9
	$\mu=2$						$n=3$–9	$n=6$–11	

quencies therefrom. The two approaches sound similar but in fact are very different. First, the dimension of the system matrix obtained from numerical approaches is very large. It increases with the cube of frequency. But the dimension of the system matrix obtained from our approach is very small and is independent of frequency. Second, the numerical approaches have to work on a three-dimensional structure but our approach work on a reduced one-dimensional structure. Finally, numerical approaches cannot identify or separate an individual wave mechanism but our approach can. It is also interesting to see that a high-frequency approach can still yield quite accurate estimations of resonant frequencies in low frequencies.

[1] I-T. Lu, ''A ray-mode (wavefront-resonance) approach for analyzing acoustic radiation and scattering by submerged structures'' J. Acoust. Soc. Am. **99**, 114–132 (1996).

[2] I-T. Lu, H. B. Bertoni, and H. Y. Chen, ''Coupling of plate modes at joints,'' J. Acoust. Soc. Am. **92**, 510–526 (1992).

[3] I-T. Lu, H. Bertoni, and Y. Q. Yao, ''Diffracted effects due to finiteness of guided structure modes,'' J. Acoust. Soc. Am. **98**, 2841–2844 (1995).

[4] P. L. Marston, ''GTD for backscattering from elastic spheres and cylinders in water and the coupling of surface elastic waves with the acoustic field,'' J. Acoust. Soc. Am. **83**, 25–37 (1988).

[5] A. D. Pierce, ''Wave propagation on thin-walled elastic cylindrical shells,'' in *Elastic Wave Propagation*, edited by M. F. McCarthy and M. A. Hayes (Elsevier, New York, 1989), pp. 205–210.

[6] L. B. Felsen and I. T. Lu, ''Ray treatment of wave propagation on thin-walled curved elastic plates with truncations,'' J. Acoust. Soc. Am. **86**, 360–374 (1989).

[7] A. N. Norris and D. A. Rebinsky, ''Membrane and flexural waves on thin shells,'' ASME J. Vib. Acoust. (in press).

[8] L. Flax, G. C. Gaunaurd, and H. Uberall, ''Theory of resonance scattering,'' in *Physical Acoustics*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1981), Vol. 10, pp. 1–60.

[9] G. C. Gaunard and M. F. Werby, ''Lamb and creeping waves around submerged spherical shells resonantly excited by sound scattering,'' J. Acoust. Soc. Am. **82**, 2021–2033 (1987).

[10] G. C. Gaunaurd and M. F. Werby, ''Acoustic resonance scattering by submerged elastic shells,'' Appl. Mech. Rev. **43**, 171–208 (1990).

[11] J. Yahner, and C. B. Burroughs, ''Frequencies of resonance of axisymmetric modes of open prolate spheroidal shells,'' J. Acoust. Soc. Am. **94**, 213–220 (1993).

[12] G. C. Gaunaurd, H. Huang, and W. Wertman, ''Acoustic scattering by elastic spherical shells that have multiple massive components attached by compliant mounts,'' J. Acoust. Soc. Am. **94**, 2924–2935 (1993).

[13] I. T. Lu, L. B. Felsen, and A. H. Kamel, ''Eigenreverberations, eigenmodes, and hybrid combinations; A new approach to propagation in layered multiwave media,'' Wave Motion **6**, 435–467 (1984).

[14] I. T. Lu and L. B. Felsen ''Matrix Green's functions for array-type sources and receivers in multiwave layered media,'' Geophys. J. R. Astron. Soc. **84**, 31–48 (1986).

[15] I. T. Lu, ''Forward modeling and data inversion for beam propagation in a stratified medium. I: Theory,'' J. Acoust. Soc. Am. **91**, 1228–1238 (1992).

[16] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves* (Prentice-Hall, Englewood Cliffs, NJ, 1973).

Spatial resolution of diffraction tomography

Thomas A. Dickens and Graham A. Winbow
Exxon Production Research Company, P.O. Box 2189, Houston, Texas 77252-2189

(Received 20 June 1994; accepted for publication 20 August 1996)

Diffraction tomography is an imaging technique applicable to crosshole seismic data and aimed at achieving optimal spatial resolution away from the borehole. In principle the method can form acoustic images equivalent to extending acoustic well logs away from the wellbore and into the formation with a spatial resolution less than one wavelength of the radiation employed to gather the crosshole data. This paper reports on the capability of diffraction tomography to produce high-resolution reconstructions of simple targets from limited-view-angle data. The goal is to quantify the resolution and velocity-reconstruction capability of diffraction tomography with realistic source–receiver geometries. Simple targets (disks and low-contrast sequences of layers) are used for this study. The scattering from these targets can be calculated without approximation, making them ideal test cases for the algorithm. The resolution capability of diffraction tomography is determined to be on the order of one wavelength for several experimental geometries. It is shown that the image-formation characteristics of diffraction tomography, in terms of its ability to determine object boundaries and velocities, are closely related to the experimental geometry. Reflection and vertical seismic profiling (VSP) experiments tend to reproduce boundaries well, while crosshole experiments give the best overall reconstruction of both target boundaries and velocity. The quantitative accuracy of the velocity reconstruction depends upon the match between the spatial-frequency content of the object and the spatial-frequency response of the algorithm. For some targets, the velocity cannot be correctly reproduced from limited-view-angle data. [S0001-4966(97)04812-1]

PACS numbers: 43.20.Fn, 43.40.Ph [ANN]

INTRODUCTION

Theoretically, diffraction tomography is a geophysical imaging technique with high potential to achieve maximum spatial resolution with minimum image distortion for cross-well seismic data. The method may also be applicable to acoustic imaging in medicine (Gelius *et al.*, 1991). In sum, it is a scheme which at least theoretically would effectively extend acoustic well logs away from the wellbores and into the interior of reservoirs. We anticipate that practical success in achieving this goal would result in more efficient exploitation of existing hydrocarbon reservoirs. It remains to develop diffraction tomography into a form in which its full imaging potential may be applied to real reservoirs.

In its present form (Devaney, 1984; Wu and Toksöz, 1987), diffraction tomography is a scheme for imaging a target of arbitrary shape and structure immersed in a uniform medium. If the velocity contrast and size of the target are not too large [Eq. (3) below], then an image of the target can be directly reconstructed that is exact within the limitations of (1) source and receiver geometry and (2) the wavelength of the illuminating radiation.

In this form, diffraction tomography provides a theoretical ''laboratory'' in which we can explore all aspects of image formation for synthetic data, for arbitrary source and receiver geometries such as crosshole transmission, vertical seismic profiling (VSP), and surface seismic. Several authors (e.g., Devaney, 1984; Wu and Toksöz, 1987; Lo *et al.*, 1988; Gelius *et al.*, 1991) have studied the ability of diffraction tomography to image point targets and disks. Others have examined more complicated targets (Pratt and Worthington, 1988). We have attempted to provide a more quantitative characterization of the performance of diffraction tomography by studying its resolution and its ability to reconstruct target velocity from limited-view-angle data.

There is no obvious aspect of diffraction tomography that restricts its application to particular structures. We expect that, at least for synthetic data, diffraction tomography will produce wavelength-scale resolution for arbitrary dip angles and complex structures, provided source and receiver geometry is appropriate.

Diffraction tomography constructs an image for each separate frequency. Data taken at many frequencies, formed for example from time-domain data, can be used to provide an enhanced image, and in this way the final image corresponds to inverting full time-domain data.

As a processing scheme, diffraction tomography treats equally all types of events: reflections, refractions, and diffractions. This is in contrast to other imaging methods such as

(1) reflection imaging (Lazaratos *et al.*, 1991, 1992; Stewart and Marchisio, 1991), which in its present form uses only data from reflections coming from beneath the source and receiver,
(2) traveltime tomography (Bregman *et al.*, 1989), which uses only the first arrival time for each recorded wavetrain, and
(3) crosswell logging (Krohn, 1990), which uses only the information from guided waves transmitted through layered formations between boreholes.

Of course, the classical filtered-backpropagation diffraction tomography algorithm (Devaney, 1984; Wu and Toksöz,

1987) we study in this paper is limited to the case of weakly scattering inhomogeneities in a uniform background medium, with evenly spaced sources and receivers. The aim of this paper is to study generic features of diffraction tomography, such as its spatial resolution, sensitivity to experiment geometry, and range of validity as the target scattering strength is varied. The principles learned from this study should apply to more complex diffraction tomography algorithms that can be used in situations in which, for example, the background velocity is not constant.

In this paper we discuss the task of imaging a disk-shaped target and a system of layers. Each target is assumed to be immersed in a uniform background medium. For these idealized problems we give a detailed analysis of the roles in image formation of

(1) reflection and transmission events,
(2) source and receiver geometry, and
(3) the achievable horizontal and vertical spatial resolution.

The scattered wave field for each object we study is calculated exactly, without making use of a weak-scattering assumption. This allows us to study the effects of target size and velocity contrast on the imaging ability of diffraction tomography, which is based upon a weak-scattering (Born) approximation.

In the present work we can completely separate the effect of the background properties of the host medium from other imaging distortions. This separation is, of course, more difficult to achieve for nonuniform background velocities. This issue is explored by Devaney and Zhang (1991); Dickens (1992, 1994); Gelius (1995); Harris and Wang (1996); and Reiter and Rodi (1996).

I. THE COMPLEMENTARY NATURE OF TRAVELTIME AND DIFFRACTION TOMOGRAPHY

As stated in the Introduction, in its present form, diffraction tomography is an algorithm for imaging weakly scattering, isolated structures embedded in a medium of constant velocity. [By weakly scattering, we mean a small object with relatively low velocity contrast with respect to the background, as in Eq. (3) below.] However, these conditions will rarely be met in real reservoir imaging situations. The algorithm will have to be extended to the case of weakly scattering structures embedded in an arbitrary background medium. The two-dimensional velocity structure of the background medium will be required as an input to the algorithm, and a natural way to obtain these velocities is with traveltime tomography [see, for example, Bregman *et al.* (1989)].

The application of traveltime tomography does not depend upon any assumptions about the strength of the scattering of a target, nor does it require careful selection of an input velocity model. It therefore provides an excellent means of obtaining the velocity structure of the imaged area with little *a priori* information. However, the spatial resolution of traveltime tomography appears to be controlled by the size of the Fresnel zone of the incident radiation (Williamson, 1991; Williamson and Worthington, 1993). This size exceeds a wavelength for typical applications.

We will see below that diffraction tomography can provide resolution on the order of one wavelength or less. Given the need of diffraction tomography to be given good input velocity information, it is natural to envision beginning with traveltime tomography, and using its velocity reconstruction as the input for diffraction tomography, which would then improve upon the resolution of the traveltime tomography image. The two methods complement each other well; traveltime tomography yields velocities from virtually any starting model, with relatively lesser spatial resolution, while diffraction tomography, for an acoustic medium with weakly scattering inhomogeneities, provides better resolution, while requiring good input velocity information. In essence, traveltime tomography yields the low-spatial-frequency (or wave number) part of the object spectrum, and diffraction tomography yields the high-frequency part. [See Pratt and Goulty (1991) for an example of the use of traveltime tomography to provide the initial velocity estimate for a wave-equation based imaging algorithm.]

II. BASIC THEORY OF DIFFRACTION TOMOGRAPHY

Diffraction tomography is based on the ability to reconstruct an object from its Fourier spectrum, and the observation that, in the Born approximation, the object's spectrum can be simply related to measurements of the acoustic field scattered from it. That is, diffraction tomography uses the simple Fourier relationship

$$O(\mathbf{r}) = \frac{1}{(2\pi)^2} \int \int d\mathbf{K} \widetilde{O}(\mathbf{K}) e^{i\mathbf{K}\cdot\mathbf{r}}, \tag{1}$$

where

$$O(\mathbf{r}) = 1 - \frac{C^2}{C^2(\mathbf{r})}, \tag{2}$$

with C the background velocity and $C(\mathbf{r})$ the velocity at $\mathbf{r}$. Here, $O(\mathbf{r})$ is called the ''object function,'' and represents fluctuations in velocity about the constant background value, which become the sources of scattered waves. $\widetilde{O}(\mathbf{K})$ can be related to the scattered field in the Born approximation, and Eq. (1) then gives a reconstruction of the object function. The Born approximation holds when (Kak and Slaney, 1988)

$$\frac{\delta C}{C} \frac{D}{\lambda} \ll 1, \tag{3}$$

where $\delta C/C$ is the relative velocity contrast and D is a measure of the size of the scattering region.

If $\widetilde{O}(\mathbf{K})$ were known for all $\mathbf{K}$, the object could be perfectly reconstructed. However, for a finite bandwidth and for source–receiver geometries that are physically realizable, it is possible to determine $\widetilde{O}(\mathbf{K})$ in only a comparatively small region of wave-number space.

Consider first a crosshole experiment with geometry shown in Fig. 1(a). (We will consider multi-frequency experiments later.) If $k_0 = 2\pi/\lambda$ is the wave number of the acoustic field, then the wave vector $\mathbf{K}$ has x and z components given by

$$K_z = -k_0(\cos\theta + \cos\phi), \tag{4}$$

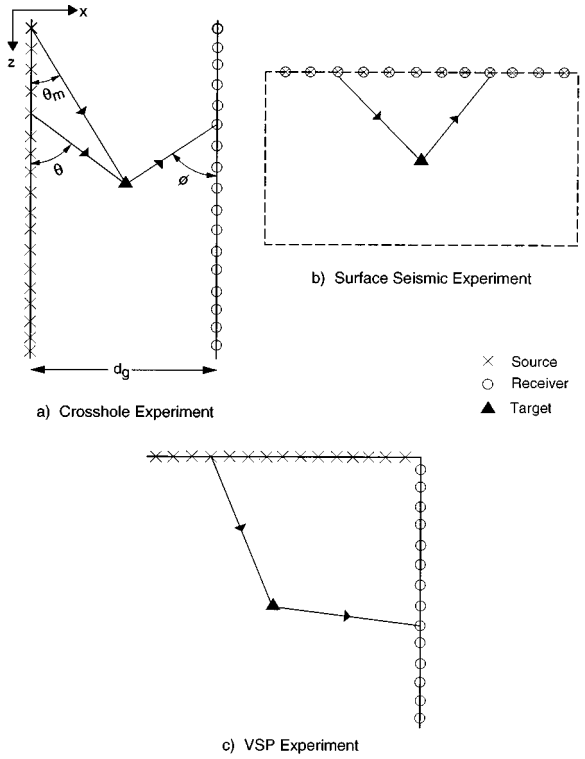


FIG. 1. Experimental geometries considered in this paper.

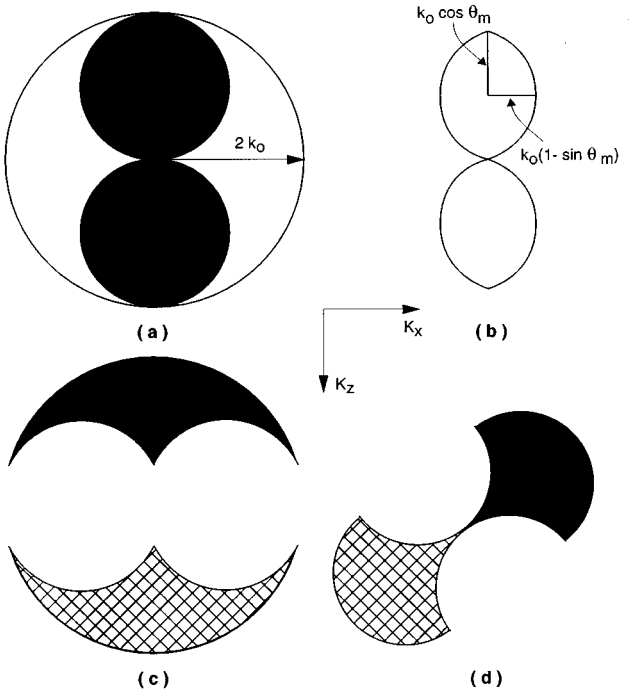


FIG. 2. Wave-number-space coverage diagrams for (a) crosshole experiment with infinite borehole length, (b) crosshole experiment with finite borehole length, (c) surface seismic experiment, and (d) VSP experiment. Crosshatched areas show additional coverage obtained using conjugate symmetry of the object function.

$$K_x = k_0(\sin\,\phi - \sin\,\theta), \tag{5}$$

where θ and ϕ are defined in Fig. 1(a). Using these equations and assuming that the boreholes are infinite in length ($0° \leqslant (\theta, \phi) \leqslant 180°$), it can be shown that the coverage would be a ''figure eight'' as shown by the blackened area in Fig. 2(a) (Devaney, 1984). (If one were able to completely surround the object being imaged, obtaining both reflection and transmission data, the outer circle of radius $2k_0$ would be completely covered.) In practice, the boreholes are finite in extent, which has the effect of reducing the coverage to a lenticular region (Lo *et al.*, 1990), as is shown in Fig. 2(b). This reduces the resolution that can be achieved and also substantially affects the accuracy of the velocity reconstruction, as will be discussed further in a later section.

Figure 1(b) and (c) show the geometries for surface seismic and offset vertical seismic profiling (VSP). Fig. 2(c) and (d) show the corresponding coverages obtained (solid diagrams) in these cases. These coverages are again for finite source and receiver line lengths. Note the poor coverage in the surface seismic case, especially in the K_x direction. If the object is lossless [$O(\mathbf{r})$ is real], then the conjugate symmetry of the wave-number-space data ($\widetilde{O}(-\mathbf{K}) = \widetilde{O}^*(\mathbf{K})$) automatically extends the coverage to the crosshatched regions shown in Fig. 2(c) and (d) (Devaney, 1984).

In this paper, the filtered backpropagation reconstruction algorithm for line sources described in Wu and Toksöz (1987) was implemented using the fast Fourier transform, which yields an algorithm of asymptotic order $N^3 \log N$, where N is the number of sources and receivers. (A formulation in terms of the FFT is apparently not possible for the VSP reconstruction; the algorithm is of order N^4.) Reconstruction proceeds quickly on a Sun SPARCstation 1+ workstation, requiring only approximately 15 s per frequency when $N = 64$.

III. SPATIAL RESOLUTION

We next examine the spatial resolution attainable by diffraction tomography. The natural way to study this problem is to examine the image of a point target (one much smaller than a wavelength). Resolution is *defined* here as the full width at the half-amplitude point of the point-target response function.

The point target scattering data were generated by the same technique used to calculate the scattered data from a disk (see the Appendix). Parameters for the point–target simulations are shown in Table I (note that only 64 sources and 64 receivers are used for the VSP case because of its greater computational requirement). The source and receiver spacings are both set at the Nyquist distance, as they are in all calculations presented in this paper. The point target, actually a disk of size much smaller than a wavelength, is given a small velocity contrast with the background medium

TABLE I. Point target simulation parameters.

Parameter	Value
Number of sources and receivers	128 (64 for VSP)
Background velocity	6000 ft/s (1829 m/s)
Target velocity	6006 ft/s (1831 m/s)
Frequency	200 Hz
Source and receiver spacing	$\lambda/2$

TABLE II. Diffraction tomography resolution.

Number of sources and receivers	128 (64 for VSP)
Source and receiver spacing	$\lambda/2$
Crosshole, target at center of imaged rectangle	
Borehole separation	20λ
Aspect ratio	3.2:1
Resolution (x,z)	$1.1\lambda \times 0.3\lambda$
Surface seismic, target 10λ below center of surface array	
Aspect ratio	3.2:1
Resolution (x,z)	$0.5\lambda \times 0.2\lambda$
VSP, target at center of imaged square	
Aspect ratio	1:1
Resolution (x,z)	$0.3\lambda \times 0.3\lambda$

(6006 ft/s in a 6000 ft/s background) to insure that strength-of-scattering effects do not influence the calculations. The diffraction tomography images were oversampled by a factor of two to allow more accuracy in measuring the resolution.

The resolution was studied for crosshole, VSP, and surface seismic geometries. Table II shows the geometrical parameters used for these cases, and the resolution measured for each. The resolution is given for a centrally located target for each geometry; the target positions are indicated in the table. Coordinates are defined such that the x axis is parallel to the surface.

For the crosshole case, as expected, the resolution is significantly poorer in the x direction because of poorer coverage in the lateral wave number K_x. In general, the resolution is approximately 1.1λ in x and 0.3λ in z, relatively independent of target location. The resolution depends on the amount of the figure-eight coverage [Fig. 2(a)] that can be covered, and will therefore depend on the length of the source and receiver lines. Larger aspect ratios (the aspect ratio is defined as the ratio of the length encompassed by sources and receivers to well separation) provide more angular coverage and hence better resolution.

The resolution achieved with surface seismic data is expected to be best in the vertical direction, and this is found to be true, with the x resolution being 0.5λ, and the z resolution 0.2λ, with the target located centrally with an aspect ratio of 3.2:1. However, while the z-direction response falls off very quickly, there are sidelobes which rise to approximately the half-amplitude level. The overall width is then approximately 1.5λ.

The VSP experiment gives a resolution of 0.3λ in both x and z, for the target at the center of the area defined by the source and receiver arrays (aspect ratio of 1:1). As required by the symmetry of this position, the resolutions in the x and z directions are equal. Hence the VSP experiment provides the best overall resolution; however, we will see that it does a poor job of reconstructing the velocity of finite-size targets.

We reiterate that, for the crosshole experiment, the area actually covered is lenticular in shape, as in Fig. 2(b). We will later show that this partial coverage of the ''figure-eight'' region available in this experiment has a great effect on the ability of the algorithm to correctly reconstruct velocities.

TABLE III. Disk target simulation parameters.

Parameter	Value
Number of sources and receivers	128
Background velocity	6000 ft/s (1829 m/s)
Frequency	200 Hz
Well separation (crosshole)	600 ft (183 m)
Well depth (crosshole)	1906 ft (581 m)
Source-receiver line length (surface seismic)	1906 ft (581 m)
Well depth and offset (VSP)	1906 ft (581 m)
Source and receiver spacing	$\lambda/2$

IV. PERFORMANCE WITH DISK TARGETS

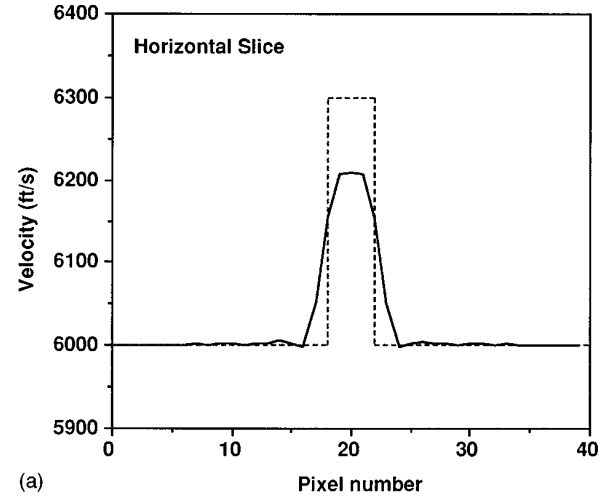
The algorithms were also tested on a variety of disk targets (cylinders in three dimensions). These tests provide insight into the range of validity of the Born approximation as an image reconstruction tool, and on the ability of the algorithms to image a larger target. It is possible to calculate exactly the field scattered from a disk (see the Appendix), thereby making the reconstruction a valid test of the Born approximation. Parameters for these simulations are shown in Table III.

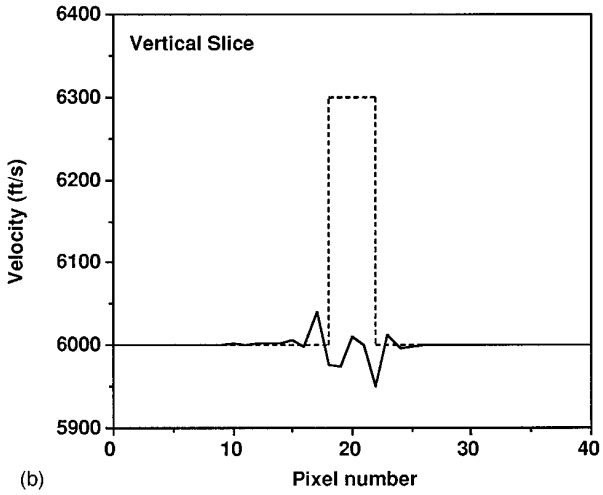
Figure 3(a)–(c) shows slices through reconstructions of a disk target of radius 1λ, for $\delta C/C = 5\%$, for the crosshole, surface seismic, and VSP geometries, respectively. For the crosshole experiment, the reconstruction of the shape and velocity are both quite good. The surface seismic experiment locates the disk boundaries, but does a very poor job of estimating the velocity. This is not unexpected since only high-frequency components of the wave-number space are available [see Fig. 2(c)], so low-frequency features like the average velocity are poorly reconstructed. This also holds true for the VSP experiment, which finds the boundaries but not the average velocity.

An understanding of the comparative properties of the disk images can be obtained from the spatial-frequency coverage diagrams of Fig. 2. The VSP and surface seismic experiments provide no coverage near zero spatial frequency, so they cannot faithfully reconstruct the average velocity of the disk. They do, however, adequately sample the high-frequency content, and therefore can reconstruct the positions of the disk boundaries. The cross-hole experiment provides the best overall coverage, and this is reflected in the target image; both the target boundary and velocity are moderately well determined.

Figure 4 shows slices through a disk of radius 2λ with a 5% velocity contrast. The definition of the disk boundaries is good, and one can see that, as expected, the vertical resolution is substantially better than the horizontal resolution. As seen before in Fig. 3(a), the velocity contrast is not completely accurately reconstructed.

To demonstrate the effects of variations in velocity contrast and target size on the validity of the Born approximation, Fig. 5 shows horizontal slices through reconstructions of disk targets of radius 1λ and 2λ, for $\delta C/C = 5\%$–20%, for the crosshole geometry. Note that even for the smallest, most weakly scattering disk the velocity is not correctly reproduced. The Born approximation has clearly failed for the larger target with a 10% velocity contrast, for which





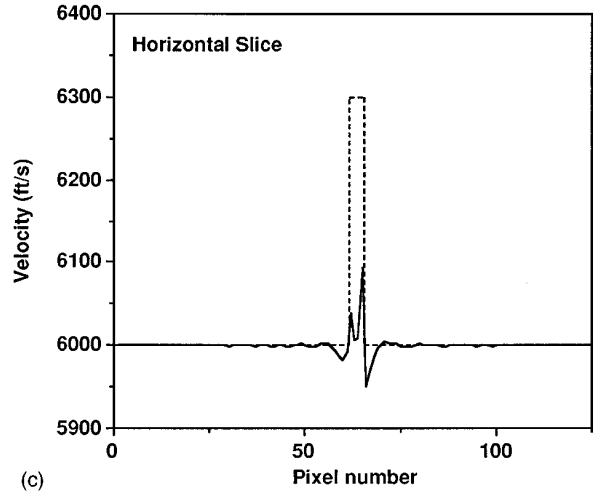
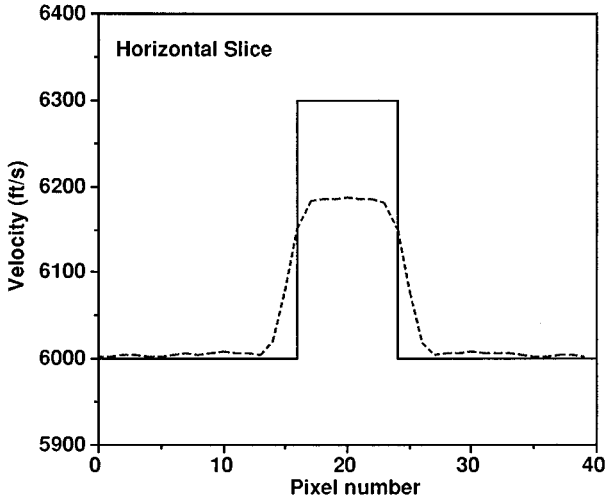


FIG. 3. (a) Horizontal slice through a reconstruction of a disk target of radius λ, crosshole experiment. In all figures, the dashed lines represent the exact velocities, and the pixel spacing is one-half wavelength. (b) Vertical slice through a reconstruction of a disk target of radius λ, surface seismic experiment. (c) Horizontal slice through a reconstruction of a disk target of radius λ, VSP experiment.

$$\frac{\delta C}{C}\frac{D}{\lambda}=0.4. \tag{6}$$

This behavior is consistent with Eq. (3). It is perhaps noteworthy that the Born approximation works so well for this



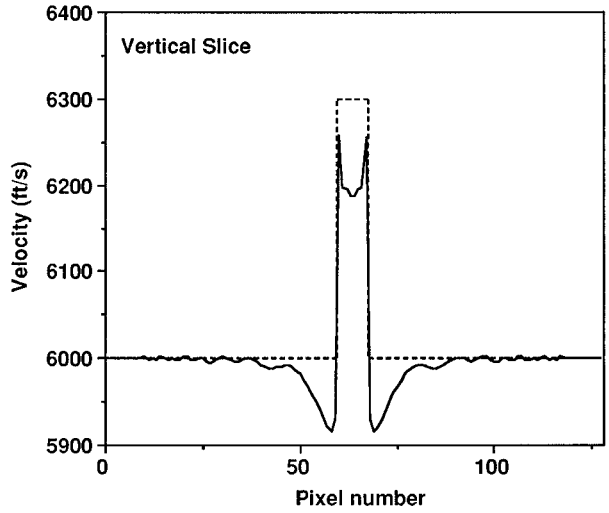


FIG. 4. Horizontal and vertical slices through a reconstruction of a disk target of radius 2λ and 5% velocity contrast, crosshole experiment.

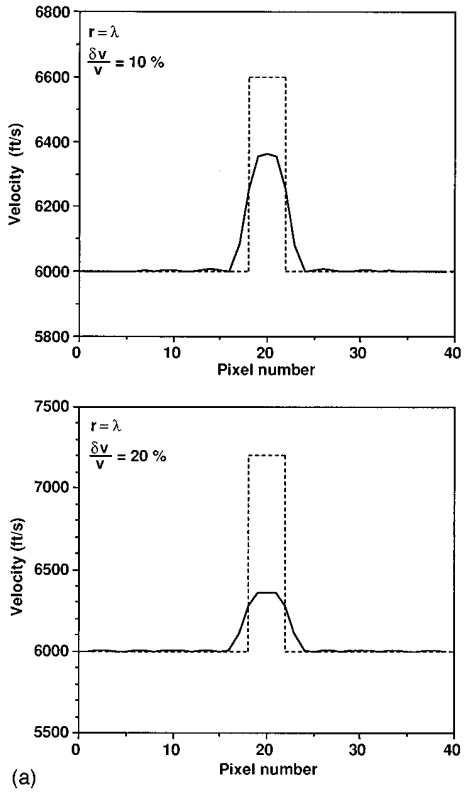
large a value of the target size and velocity contrast. For the smaller target, the accuracy of the velocity reconstruction is seen to decrease gradually as the contrast increases.

The failure to reproduce the velocity correctly, even when the Born approximation should hold, arises from the inability to obtain experimental data over all of **K** space. In fact, it is possible to predict the fraction of the actual velocity that will result from the reconstruction by integrating $\tilde{O}(\mathbf{K})$ over the appropriate lenticular-shaped coverage regions. For a disk of radius a, $\tilde{O}(\mathbf{K})$ is given by

$$\tilde{O}(\mathbf{K})=\frac{2\pi a}{|\mathbf{K}|}J_1(|\mathbf{K}|a)\left(1-\frac{C^2}{C^2_{\text{disk}}}\right), \tag{7}$$

where J_1 is the Bessel function of the first kind of order one. This can be integrated numerically over the appropriate coverage region in wave-number space to predict the reconstructed velocity.

Figure 6 shows plots of the predicted fraction of the actual object function (at the disk center) that will be obtained in the reconstruction of a disk target of radius a as a function of k_0a, for various types of coverage. The coverages for crosshole and surface seismic have been param-



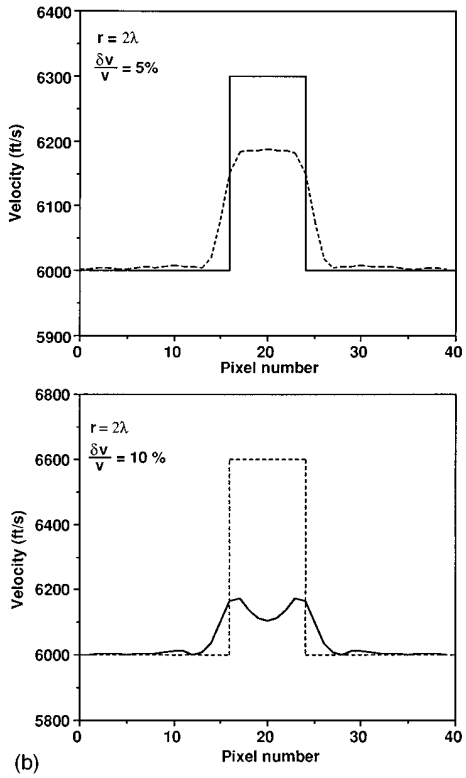


FIG. 5. Horizontal slices through reconstructions of disk targets, crosshole experiment, illustrating the breakdown of the Born approximation as the target size and contrast become larger. The target radius and contrast are shown on each plot.

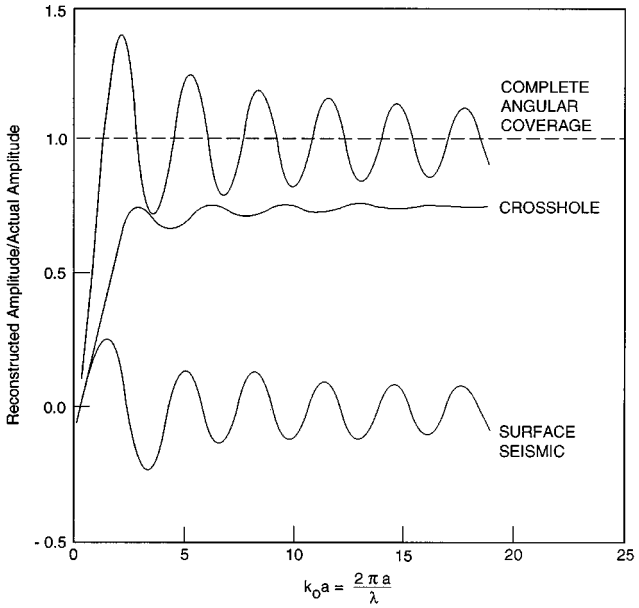


FIG. 6. Ratio of the reconstructed value of the object function to the actual value of the object function at the center of a disk target of radius a, for various **K**-space coverage types. The crosshole and surface seismic curves assume an aspect ratio of approximately 3:1.

etrized by the minimum scattering angle available, θ_m [see Fig. 1(a)]; results are shown for $\theta_m = 20°$, which corresponds to an overall aspect ratio for the crosshole experiment of approximately 3:1. The estimate for full-circle coverage approaches the correct value for large $k_0 a$, but contains oscillations of fairly large amplitude. [The oscillations arise because $\tilde{O}(\mathbf{K})$ is oscillatory and is integrated over a region with a sharp cutoff.] Because of the geometrical constraints on the source and receiver geometry (limited aspect ratio), the crosshole experiment fails to ever yield a correct value for the velocity, even for very large $k_0 a$. If $\theta_m = 0°$, with the result that the ''figure eight'' coverage is completely filled, then the correct velocity *is* reproduced for large $k_0 a$. It is easy to show that, for the lenticular region, the ratio asymptotically approaches $1 - \theta_m/(\pi/2)$ $(= 0.78$ in Fig. 6). This value agrees well with the reconstructed fraction observed in Fig. 3(a), in which the target scatters weakly enough that Born approximation violations have little effect. The surface seismic data yield a very poor estimate of the velocity; an example of this was seen in Fig. 3(b).

We speculate that, if the diffraction tomography algorithm is applied iteratively to crosshole data (which requires its extension to nonconstant background velocities), the result will converge to the correct velocity. This is because the reconstruction at each iteration will provide a fixed fraction of the object function; the magnitude of the object function will decrease with each iteration, so that the target velocity will eventually be reproduced correctly. We consider, for example, a situation in which the previous estimate of a disk target's velocity is used as the local background velocity for the next iteration of diffraction tomography. In this case, if the difference between the true target velocity and the initial background velocity is $\Delta = C_{\text{true}} - C_0$ and the fraction of the object function which is returned by each iteration is ϵ, then

$$C_{\text{true}} - C_K = (1-\epsilon)^K \Delta, \tag{8}$$

where C_{true} is the true target velocity and C_K is the estimate

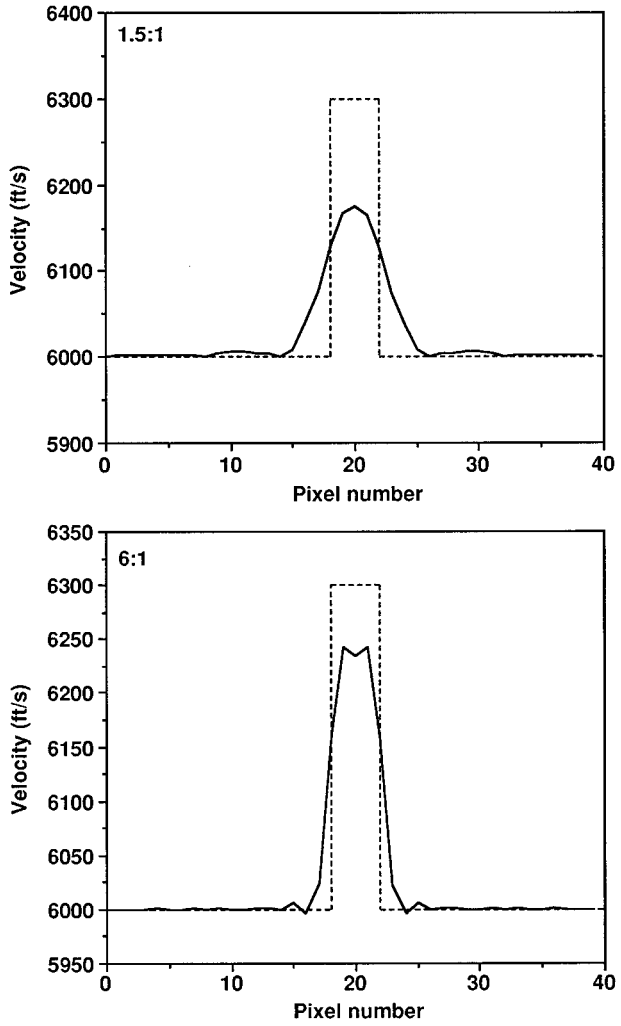


FIG. 7. Horizontal slices through reconstructions of disk targets, with aspect ratios of 1.5:1 and 6:1, using the crosshole geometry.

after K iterations. This equation is derived by assuming that the initial background velocity C_0 satisfies

$$\left|\frac{C_{\text{true}}-C_0}{C_{\text{true}}}\right| \ll 1, \tag{9}$$

and that the fraction ϵ is constant from iteration to iteration, since, as we discuss elsewhere in this paper, it is a function only of the experiment geometry and target shape. Further exploration of this issue is beyond the scope of this paper.

It is also important to study the effectiveness of the algorithm as a function of the number of sources and receivers used, which translates to aspect ratio, for fixed source and receiver spacing. Figure 7 shows plots of slices (in the x direction) through the image of a disk of radius λ and velocity contrast 5%, for $N=$ 64 and 256. The source and receiver spacings are kept at $\lambda/2$, so the aspect ratio varies from 1.5 to 6. Note the substantially improved performance of the 256 source case, especially in the accuracy with which the target velocity is reconstructed, due to the larger angular coverage.

TABLE IV. Layered medium simulation parameters.

Parameter	Value
Number of sources and receivers	128
Frequency	200 Hz
Well separation	600 ft (183 m)
Well depth	1906 ft (581 m)
Source and receiver spacing	$\lambda/2$

V. PERFORMANCE WITH LAYERS

Given that most targets of geological interest will consist of layers of variable velocity, it is important to determine the capability of diffraction tomography to image such structures. This provides a more realistic test of the applicability of the method to subsurface imaging.

The acoustic field scattered from an arbitrary layered medium with parameters which depend only on depth can be calculated in a numerically stable fashion, by first making the familiar lateral Fourier transform of the wave equation. The wave equation (in wave-number space) is then recast as a first-order Riccati differential equation which can be solved exactly in each constant-velocity layer. One uses the boundary conditions of the problem to choose the correct combination of the solutions from each layer. An inverse Fourier transform then yields the position-space Green's function.

A code was developed to implement this algorithm to provide scattered field data to test the diffraction tomography algorithm. Parameters for the calculations shown in this section, all of which are performed with the crosshole geometry, are given in Table IV.

Figure 8 shows three layers [velocities 5950, 6050, and 5950 ft/s (1814, 1844, and 1814 m/s)] embedded in a background of velocity 6000 ft/s (1829 m/s). The layers are approximately 3λ thick. Here the velocities and positions are both well determined; the velocity contrast is low enough so that the Born approximation holds ($\delta C/C \times D/\lambda = 0.16$). Note that the layer velocities are very well determined. This

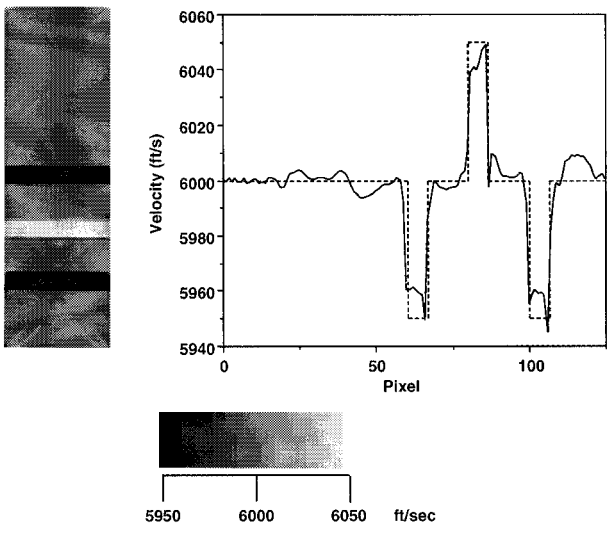


FIG. 8. Image and vertical slice, reconstruction of three low-contrast layers in a constant velocity background.

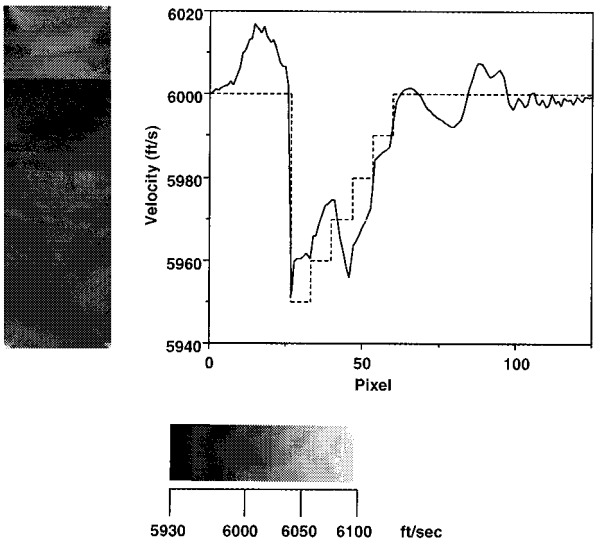


FIG. 9. Image and vertical slice, reconstruction of five low-contrast layers in a constant velocity background.

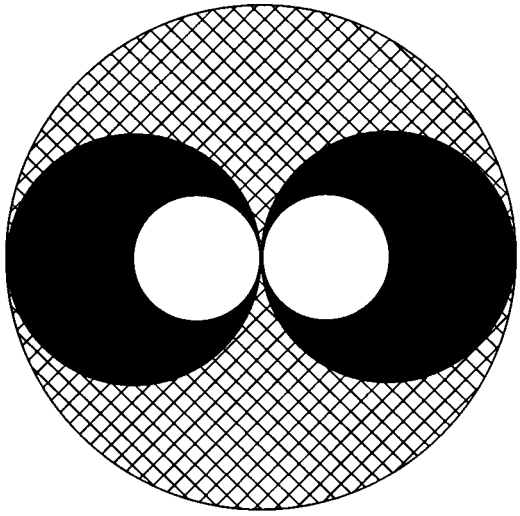


FIG. 10. **K**-space coverage with multiple frequencies, surface seismic data. The crosshatched area represents the coverage at a single frequency $f_{\max}$; the solid area shows the additional coverage obtained when frequencies from $0.5\,f_{\max}$ through $f_{\max}$ are used.

is because the spatial frequency representation of a layer contains only low-frequency components in the K_x direction; therefore, the crosshole coverage shown in Fig. 2 provides adequate sampling of the layer's frequency spectrum. Integrating the analytic expression for $\tilde{O}(\mathbf{K})$ for a layer over the lenticular coverage area for the example shown in Fig. 8 predicts that the ratio of the reconstructed object function to the actual object function would be 1.13, so it is not surprising that the velocity is reconstructed accurately in Fig. 8.

A further example, Fig. 9, shows five layers of varying velocity [5950–5990 ft/s (1814–1826 m/s)] embedded in a constant velocity [6000 ft/s (1829 m/s)] background. The performance here is clearly not as good as is the case for isolated layers; however, the general trend of the velocity variation is fairly well represented. Some anomalies are present in the background region. The area in which the velocity contrast occurs is large (the total thickness is 16λ), so that the Born approximation is not as accurate for this case.

This last example serves to point out that in geological situations likely to be encountered in nature, such as a layered medium with a more or less smoothly increasing velocity, methods based on the Born approximation may fail to provide an acceptable inversion. Such methods are more suitable for isolated targets, and need to be improved upon to be made into tools for geophysical study. As discussed previously, using traveltime tomography to provide starting velocities for iterative techniques based on the distorted wave Born approximation should widen the applicability of diffraction tomography to include realistic geological situations.

VI. EXTENDING COVERAGE IN K SPACE

It is possible to extend the region of coverage in **K** space by including multiple frequencies in the calculation (Wu and Toksöz, 1987). For example, Fig. 10 shows the extension in coverage that can be obtained for surface seismic imaging if frequencies from $1/2f_{\max}$ through $f_{\max}$ are included, where

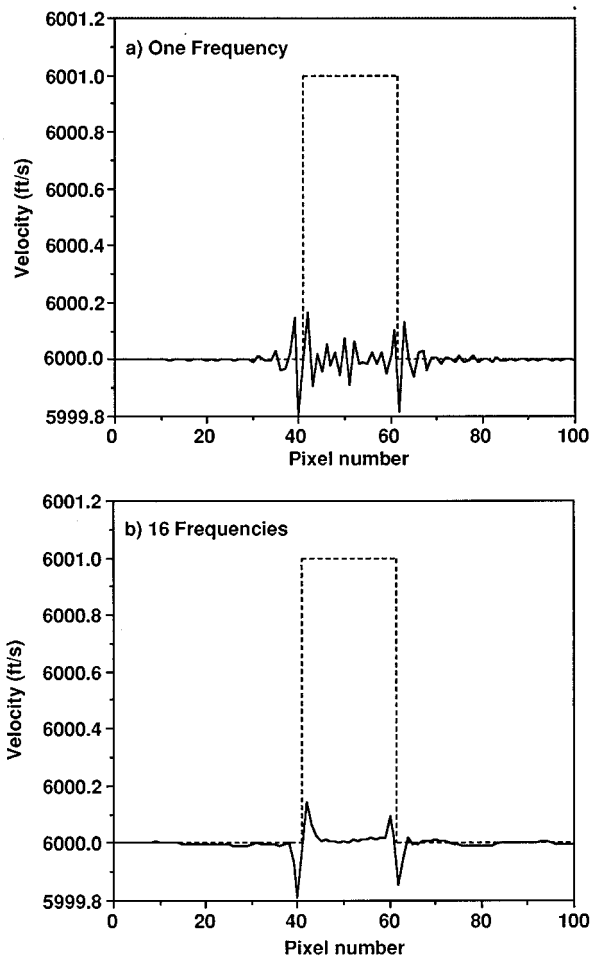


FIG. 11. Vertical slices through a surface seismic reconstruction of a disk of radius 2λ, using single (a) and multiple (b) frequencies.

TABLE V. Multiple frequency surface seismic simulation parameters.

Parameter	Value
Number of sources and receivers	128
Background velocity	6000 ft/s (1828.8 m/s)
Target velocity	6001 ft/s (1829.1 m/s)
Maximum frequency	256 Hz
Source–receiver line length	1490 ft (454 m)
Source and receiver spacing	$\lambda_{min}/2$

f_{max} is the maximum temporal frequency used in the reconstruction. Substantially more information is obtained. Figure 11 shows vertical slices through reconstructions of a disk of radius 2λ reconstructed from surface seismic data with 128 sources and receivers, both at a single frequency [f_{max}, Fig. 11(a)] and averaged over 16 frequencies [Fig. 11(b)] from $f_{max}/16$ through f_{max}. Parameters for the calculation are shown in Table V. Note that the source and receiver spacings are kept constant as the frequency is changed, at the Nyquist spacing for the highest frequency used.

The boundary of the disk is much more clearly defined when many frequencies are averaged, and the high-frequency noise is greatly reduced. Again, the velocity is very poorly determined for this geometry. This is expected, since even with the multiple frequencies there is little information near $|\mathbf{K}|=0$ because of the angular coverage limitation.

Another way to provide additional coverage is by adding data obtained from different geometries (Wu and Toksöz, 1987). A very slight improvement in the VSP image of a disk of radius 1λ was achieved by adding surface seismic data. The average velocity was still poorly reconstructed, because there were still no data near $|\mathbf{K}|=0$.

VII. SUMMARY

In this paper diffraction tomography has been shown to be effective in imaging an isolated target with low velocity contrast and size of a few wavelengths. The ability of the algorithm to determine the velocity within the target depends crucially on the configuration of sources and receivers. Only the crosshole geometry gives a reasonable estimate of the average velocity, because it is the only geometry which enables us to accurately reconstruct the low-**K** Fourier components of the target velocity. The surface seismic and VSP data work well only for locating the boundaries of the target. With limited-view-angle data, for a disk target none of the geometries can provide a completely correct determination of target velocity. Crosshole imaging is particularly well suited for determining the velocities of layer targets, as its spatial-frequency response matches the frequency content of a layered target.

These differences in performance, and the inability of even the crosshole experiment to exactly determine the target velocity, can be ascribed to the lack of complete coverage in **K** space. When there is little data near the origin in **K** space, the estimate of the target velocity is very poor. In the examples that we have studied, if only high-frequency data are available, it is only possible to determine the location of the target boundaries.

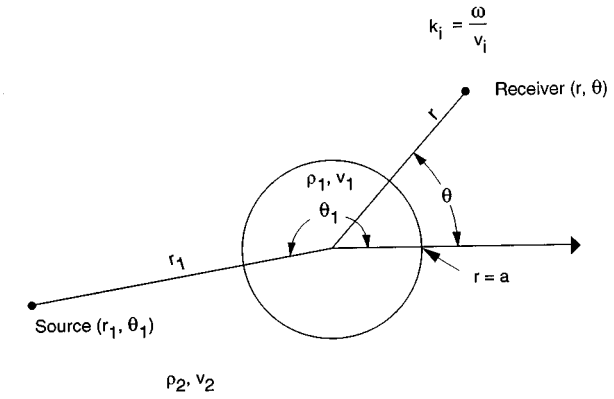


FIG. A1. Geometry and notation used in the calculation of the acoustic scattering from a disk (see the Appendix).

Additional coverage in **K** space can be obtained by using additional frequencies or by adding data from different source–receiver configurations, such as VSP or surface seismic. Use of additional frequencies tends to sharpen the boundaries in a reflection image.

When the Born approximation is valid, diffraction tomography forms an accurate imaging method for the classes of targets discussed in this paper. Thus the limitations on resolution discussed here are not just limitations of the diffraction tomography method of data analysis. They are in fact limitations of all imaging techniques that could be applied to the data.

However, diffraction tomography has the advantage over other techniques that its imaging potential is clearly displayed in terms of the sampling achieved in **K** space as shown in Devaney (1984) and Wu and Toksöz (1987). It is clear that image enhancement can be achieved by interpolating whatever sampling in **K** space is created by a given experimental geometry. A minimum-entropy approach to this is given in Lo *et al.* (1990). In addition, approaches based on the method of convex projection (Sezan and Stark, 1982) are also possible and may be able to create enhanced images.

This paper has discussed the use of diffraction tomography to image an isolated target in a uniform background host medium. As we indicated at the beginning of this paper, diffraction tomography is by no means limited to this simple case. Dickens (1992, 1994) explains how to image complex structures using a generalization of the traditional methods discussed in this paper.

APPENDIX: MATHEMATICAL FORMULATION OF DISK MODEL

Figure A1 defines the geometry and notation for the following calculation of the field scattered from a disk. [A similar calculation is presented for plane-wave illumination in Morse and Ingard (1968).] Source and receiver coordinates are given by (r_1,θ_1) and (r,θ), respectively, while the density and velocity inside and outside the disk are given by (ρ_1,v_1) and (ρ_2,v_2), respectively. The source term is the two-dimensional Green's function

$$\phi_0(|\mathbf{x}_1-\mathbf{x}|)=\frac{i}{4}H_0(k_2|\mathbf{x}_1-\mathbf{x}|), \tag{A1}$$

which can be expanded as, for $r<r_1$,

$$\phi_0(|\mathbf{x}_1-\mathbf{x}|)=\frac{i}{4}\sum_n e^{-in(\theta_1-\theta)}H_n(k_2r_1)J_n(k_2r). \tag{A2}$$

Here, H_n is the Hankel function of the first kind and J_n is the Bessel function of the first kind.

The fields outside and inside the disk can be written as

$$\phi_B(\mathbf{r})=\sum_n e^{in\theta}H_n(k_2r)B_n \tag{A3}$$

and

$$\phi_A(\mathbf{r})=\sum_n e^{in\theta}J_n(k_1r)A_n, \tag{A4}$$

respectively.

The coefficients A_n and B_n are determined by using the boundary conditions for the problem, which require that $\rho\phi$ and $\partial\phi/\partial r$ be continuous at the surface of the disk. Enforcing these boundary conditions leads to the expression for the scattered field ϕ_B

$$\phi_B=\frac{i}{4}\sum_n e^{-in(\theta_1-\theta)}H_n(k_2r_1)H_n(k_2r)\frac{\tilde{N}_B}{D}, \tag{A5}$$

where

$$\tilde{N}_B=\rho_1k_2J_n(k_1a)\dot{J}n(k_2a)-\rho_2k_1\dot{J}_n(k_1a)J_n(k_2a) \tag{A6}$$

and

$$D=\rho_2k_1H_n(k_2a)\dot{J}_n(k_1a)-\rho_1k_2\dot{H}_n(k_2a)J_n(k_1a). \tag{A7}$$

The sum in Eq. (A5) can be written as a sum over positive indices by using the index symmetry properties of the Bessel function. In Eqs. (A6) and (A7) a dot over a function denotes differentiation.

Bregman, N. D., Bailey, R. C., and Chapman, C. H. (**1989**). ''Crosshole seismic tomography,'' Geophysics **54**, 200–215.

Devaney, A. J. (**1984**). ''Geophysical diffraction tomography,'' IEEE Trans. Geosci. Remote Sensing **GE-22**, 3–13.

Devaney, A. J., and Zhang, D. H. (**1991**). ''Geophysical diffraction tomography in a layered background,'' Wave Motion **14**, 243–265.

Dickens, T. A. (**1992**). ''Study of diffraction tomography for layered backgrounds,'' 62nd Annual Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 770–773.

Dickens, T. A. (**1994**). ''Diffraction tomography for crosswell imaging of nearly layered media,'' Geophysics **59**, 694–706.

Gelius, L.-J., Johansen, I., Sponheim, N., and Stamnes, J. J. (**1991**). ''A generalized diffraction tomography algorithm,'' J. Acoust. Soc. Am. **89**, 523–528.

Gelius, L.-J. (**1995**). ''Limited-view diffraction tomography in a nonuniform background,'' Geophysics **60**, 580–588.

Harris, J. M., and Wang, G. Y. (**1996**). ''Diffraction tomography for inhomogeneities in layered background medium,'' Geophysics **61**, 570–583.

Kak, A. C., and Slaney, M. (**1988**). *Principles of Computerized Tomographic Imaging* (IEEE, New York).

Krohn, C. E. (**1990**). ''Cross-well continuity logging using seismic guided waves,'' 60th Annual Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 43–46.

Lazaratos, S. K., Rector, J. W., Harris, J. M., and Van Schaack, M. (**1991**). ''High-resolution imaging with cross-well reflection data,'' 61st Annual Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 107–110.

Lazaratos, S. K., Harris, J. M., Rector, J. W., and Van Schaack, M. (**1992**). ''High resolution cross well of a West Texas carbonate reservoir: Part 4. reflection imaging,'' 62nd Annual Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 49–53.

Lo, T., Toksöz, M. N., Xu, S., and Wu, R. S. (**1988**). ''Ultrasonic laboratory tests of geophysical tomographic reconstruction,'' Geophysics **53**, 947–956.

Lo, T., Duckworth, G. L., and Toksöz, M. N. (**1990**). ''Minimum cross entropy seismic diffraction tomography,'' J. Acoust. Soc. Am. **87**, 748–756.

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

Pratt, R. G., and Worthington, M. H. (**1988**). ''The application of diffraction tomography to cross-hole seismic data,'' Geophysics **53**, 1284–1294.

Pratt, R. G., and Goulty, N. R. (**1991**). ''Combining wave-equation imaging with traveltime tomography to form high-resolution images from cross-hole data,'' Geophysics **56**, 208–224.

Reiter, D. T., and Rodi, W. (**1996**). ''Nonlinear waveform tomography applied to crosshole seismic data,'' Geophysics **61**, 902–913.

Sezan, M. I., and Stark, H. (**1982**). ''Image restoration by the method of convex projections: Part 2—Applications and numerical results,'' IEEE Trans. Med. Imag. **MI-1**, 95–101.

Stewart, R. R., and Marchisio, G. (**1991**). ''Cross-well seismic imaging using reflections,'' 61st Annual Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 375–378.

Williamson, P. R. (**1991**). ''A guide to the limits of resolution imposed by scattering in ray tomography,'' Geophysics **56**, 202–207.

Williamson, P. R., and Worthington, M. H. (**1993**). ''Resolution limits in ray tomography due to wave behavior: Numerical experiments,'' Geophysics **58**, 727–735.

Wu, R. S., and Toksöz, M. N. (**1987**). ''Diffraction tomography and multisource holography applied to seismic imaging,'' Geophysics **52**, 11–25.

Two alternative expressions for the spherical wave expansion of the time domain scalar free-space Green's function and an application: Scattering by a soft sphere

O. Merih Buyukdura and S. Sencer Koc
Department of Electrical and Electronics Engineering, Middle East Technical University, 06531 Ankara, Turkey

(Received 27 March 1996; accepted for publication 24 July 1996)

The importance of expanding Green's functions, particularly free-space Green's functions, in terms of orthogonal wave functions is practically self-evident when frequency domain scattering problems are of interest. With the relatively recent and widespread interest in time domain scattering problems, similar expansions of Green's functions are expected to be useful in the time domain. In this paper, two alternative expressions, expanded in terms of orthogonal spherical wave functions, for the free-space time domain scalar Green's functions are presented. Although the two expressions are equivalent, one of them is seen to be more convenient for the calculation of the scattered field for a known equivalent source density, whereas the second expression is more suitable for setting up an integral equation for the equivalent source density. Such an integral equation may be setup, for example, by the application of a time domain equivalent of the T-matrix (extended boundary condition) method. © *1997 Acoustical Society of America.* [S0001-4966(97)00312-3]

PACS numbers: 43.20.Px, 43.20.Fn [ANN]

INTRODUCTION

Expansions of Green's functions in terms of orthogonal wave functions are well known to be important and useful in the freqency domain; see for instance Harrington, 1961; Bowman *et al.*, 1969, and Felsen, 1957. Such expansions should prove to be useful also in the time domain. We present two alternative expressions for the time domain free-space scalar Green's function in terms of wave functions orthogonal over a spherical surface. Both expressions are in terms of separated functions of θ, θ', ϕ, and ϕ'. The dependence on R and R', in contrast, however, is in a more ''compact'' form in one of the expressions, while it is ''separated,'' in the second expression. Here, by ''separated,'' we mean that each term in the expansion appears as a function of R *convolved* with a function of R'. In the foregoing, (R,θ,ϕ) are the familiar spherical coordinates of the point of observation and their primed counterparts are those of the source point. While the first expression is more convenient to calculate the field due to known impressed or equivalent (or induced) sources, the second is more suitable to use in a scattering formulation to set up an (for instance integral) equation for some unknown quantities which, in turn, yield the equivalent sources.

The problem is formally stated in Sec. I. In Sec. II, the ''compact'' expression in the form of an expansion in terms of spherical wave functions is found; while in Sec. III, a similar expansion which is ''separated'' is given. As a numerical example, the second expression for the Green's function is used to solve for the scattering by an acoustically soft spherical surface in Sec. IV. Brief concluding remarks are given in Sec. V.

To discuss some of the earlier related work, we cite Heyman and Devaney, 1996; Hansen, 1995; and Jones, 1986.

In Heyman and Devaney, 1996, a formulation is given to find the time domain field due to a source distribution, of compact support, in free space. The time dependence of the source is taken to be of finite duration. Admittedly, the problems for which the two formulations (those in Heyman and Devaney, 1996 and in the present work) are similar (essentially to find the time domain field due to a source distribution in free-space). Another similarity is that both solutions are in the form of an expansion in terms of Tesseral harmonics. But there, the similarity ends: In Heyman and Devaney, 1996 the space-time dependence of the source enters via an nth order linear integral operator [Eq. (3.14) therein], while in the present work, the same dependence enters via two [superposition (in space) and convolution (in time)] integrals. (Here the index n is as used in this paper, namely, the order of the associated Legandre functions; in the work cited, the letter l is used for the same index.) It is our opinion that our formulation is easier to use and can also be considered to be more appealing in that the analogy with using the frequency domain Green's function (which enjoys wide familiarity) is more straightforward.

In Hansen, 1995 the analysis deals with finding the time domain field, outside a sphere of radius a which encapsulates all the impressed sources, once the field on the spherical surface is known. Apart from considerations of probe correction and sampling, the essential achievement of this work is to find the Green's function of the first kind (it satisfies the Dirichlet boundary condition on the spherical surface $R=a$) and to use it in conjunction with an expansion of the field on the spherical surface. While the latter expansion is in terms of Tesseral harmonics, the Green's function of the first kind is expanded in terms of the complex resonant frequencies of the soft sphere (ω_{ns} therein). Although this Green's function could certainly be used to find the scattering from a soft sphere, the formulation is substantially different from

that in the present work, in that here, the problem is solved entirely in the time domain (via deconvolution).

In Jones, 1986, a formulation is given for the time domain scattering from a hard sphere. Again, this formulation is substantially different from ours, in that the differential equation for the field is Laplace transformed, the boundary conditions imposed and the time domain field found via inverse transformation.

Finally, in regards to the two latter formulations cited, we note that in the present work, the scattering from a sphere is presented only as an *example* (i) to check numerically the validity of our Green's function expressions, and (ii) to demonstrate the utility of the same. In summary, what in essence we offer to the literature are two alternative expressions of the free-space time domain Green's function expanded in terms of spherical waves.

Throughout the paper the velocity of waves in free-space is taken to be $c=1$, so that in any frequency domain expression which appears, k, the wave number is equal to ω, the angular frequency.

I. STATEMENT OF THE PROBLEM

The problem of interest is to find the scalar Green's function which satisfies

$$\left[\nabla^2-\frac{\partial^2}{\partial t^2}\right]g_0(\mathbf{R},\mathbf{R}',t)=-\delta(\mathbf{R}-\mathbf{R}')\,\delta(t), \tag{1}$$

as well as the radiation condition and causality. Here, $\mathbf{R}$ and $\mathbf{R}'$ stand for the position vectors denoting the observation and the source points, respectively and t stands for time. The Green's function is useful to find the field $\psi(\mathbf{R},t)$, in free-space, due to a source $f(\mathbf{R},t)$, in other words the solution to the inhomogeneous partial differential equation

$$\left[\nabla^2-\frac{\partial^2}{\partial t^2}\right]\psi(\mathbf{R},t)=f(\mathbf{R},t). \tag{2}$$

The solution to Eq. (2) is given by

$$\psi(\mathbf{R},t)=-\int_V\int_{\tau=-\infty}^{\infty}g_0(\mathbf{R},\mathbf{R}',t-\tau)f(\mathbf{R}',\tau)d\tau\,dv', \tag{3}$$

where V is the region occupied by the source f. Note from Eq. (3) that when viewed spatially, g_0 is indeed a Green's function in that it appears as the kernel of a superposition integral, while considered temporally, it is an *impulse response* appearing in a convolution integral.

The solution to Eq. (2) is well known in closed form (Chew, 1990):

$$g_0(\mathbf{R},\mathbf{R}',t)=\frac{\delta(t-|\mathbf{R}-\mathbf{R}'|)}{4\pi|\mathbf{R}-\mathbf{R}'|}. \tag{4}$$

Equation (4) is simply the inverse Fourier transform of the well known expression for the closed form Green's function in the frequency domain, namely,

$$G_0(\mathbf{R},\mathbf{R}',k)=\frac{e^{-jk|\mathbf{R}-\mathbf{R}'|}}{4\pi|\mathbf{R}-\mathbf{R}'|}. \tag{5}$$

Although, strictly speaking, Eq. (5) is valid only for positive values of k, the frequency (it is frequently noted in the literature that it is valid for an $e^{+j\omega t}$ time dependence), it can be used as is to be inverse-tranformed since it is already conjugate-symmetric, yielding the real Eq. (4).

II. ONE EXPANSION FOR THE GREEN'S FUNCTION

It is desirable to expand Eq. (5) in terms of functions which are orthogonal over a spherical surface. To this end, we temporarily restrict the source point to be on the z axis, i.e., choose $\mathbf{R}'=\hat{z}z'$. Another temporary restriction is that $R>R'$. Next the Green's function is expanded in terms of Tesseral harmonics:

$$g_0(R,\theta,\phi;z';t)=\sum_p\sum_{n=0}^{\infty}\sum_{m=0}^{n}a_{pnm}(R,z',t)Y_{pnm}(\theta,\phi), \tag{6}$$

where the parity index p takes on the ''values'' either e or o (standing for ''even'' and ''odd,'' respectively) and the coefficients a_{pnm} are, as yet, unknown. The Tesseral Harmonics are given by

$$Y^{e}_{onm}(\theta,\phi)=P_n^m(\cos\,\theta)\begin{matrix}\cos\,m\phi\\ \sin\,m\phi\end{matrix}, \tag{7}$$

in which P_n^m are the associated Legendre functions of the first kind; $\cos m\phi$ goes with parity e and $\sin m\phi$ goes with parity o. The Tesseral Harmonics are well known to be orthogonal over a spherical surface; this property is exploited to solve for a_{pnm} using standard techniques: Both sides of Eq. (6) are multiplied by a Tesseral harmonic of a given index (or, rather set of indices) and integrated on a spherical surface, centered at the origin, of radius R. The integration is straightforward, since the integrand is independent of ϕ and involves a delta (or impulse) function in θ. When the coefficients thus found are substituted into Eq. (6), the desired expression for the restricted Green's function is obtained:

$$\begin{aligned}g_0(R,\theta,\phi;z';t)&=\frac{1}{8\pi Rz'}\sum_{n=0}^{\infty}(2n+1)\\&\quad\times P_n\left(\frac{R^2+z'^2-t^2}{2Rz'}\right)P_n(\cos\,\theta)p\left(\frac{t-R}{z'}\right).\end{aligned} \tag{8}$$

Here, P_n are the Legendre polynomials and $p(\cdot)$ is a pulse equal to unity when its argument is in between -1 and 1 and vanishes elsewhere. To remove the restriction $R>z'$, the same procedure is repeated, this time integrating over a spherical surface with radius $R<z'$; this yields an equation very similar to Eq. (8).

To remove the restriction $\theta'=0$, use is made of the addition theorem for the Legendre polynomials under coor-

dinate rotation (Gradshteyn and Ryzhik, 1980). The details of the application of the theorem are left out for the sake of brevity; the final desired result is

$$g_0(R,\theta,\phi;R',\theta',\phi';t) = \frac{1}{8\pi RR'}\sum_{n=0}^{\infty}\left\{(2n+1)P_n\left(\frac{R^2+R'^2-t^2}{2RR'}\right)p\left(\frac{t-R_>}{R_<}\right) \times \sum_{p=e,o}\sum_{m=0}^{n}\epsilon_m\frac{(n-m)!}{(n+m)!}Y_{pnm}(\theta,\phi)Y_{pnm}(\theta',\phi')\right\}, \tag{9}$$

where $\epsilon_0=1$ and $\epsilon_m=2, m\neq 0$. Finally we note that $R_>$ stands for the larger of R and R', whereas $R_<$ for the smaller. It is seen that in Eq. (9) the dependence on R and R' is not in a separated form.

III. ANOTHER EXPANSION FOR THE GREEN'S FUNCTION

In some applications, it is desirable to express each mode in Eq. (9) as a function of R, convolved with a function of R'. To this end, we shall find the inverse Fourier transform of the frequency-domain wave function expansion of the Green's function:

$$G_0(R,\theta,\phi;R',\theta',\phi';t) = -\frac{jk}{4\pi}\sum_{n=0}^{\infty}\left\{(2n+1)j_n(kR_<)h_n^{(2)}(kR_>) \times \sum_{p=e,o}\sum_{m=0}^{n}\epsilon_m\frac{(n-m)!}{(n+m)!}Y_{pnm}(\theta,\phi)Y_{pnm}(\theta',\phi')\right\}. \tag{10}$$

For this purpose, it is necessary to find the inverse Fourier transforms of $j_n(kR_<)$ and $h_n^{(2)}(kR_>)$, the spherical Bessel and Hankel functions of the second kind, respectively. In passing, we note that Eq. (10) is also conjugate-symmetric in k, and will properly yield a real inverse transform. The transform of the Bessel functions are found easily (Abramowitz and Stegun, 1972):

$$\mathscr{F}^{-1}\{j_n(kR_<)\} = \frac{j^n}{2R_<}P_n\left(\frac{t}{R_<}\right)p\left(\frac{t}{R_<}\right). \tag{11}$$

Unfortunately, the inverse Fourier transforms of the Hankel functions do not exist *even in the distributional sense*. On the other hand, we note that what we need is not the transform of the Hankel function *per se* but only that of the product:

$$\mathscr{F}^{-1}\{-jkR_>h_n^{(2)}(kR_>)j_n(kR_<)\}. \tag{12}$$

This transform does exist, so that we formally write, in the sense defined in the Appendix,

$$\mathscr{F}^{-1}\{-jkR_>h_n^{(2)}(kR_>)\} = j^n\frac{\partial}{\partial t}\left[P_n\left(\frac{t}{R_>}\right)u(t-R_>)\right] \triangleq j^nU_n(R_>,t-R_>), \tag{13}$$

where $u(\cdot)$ stands for the unit step function. The functions U_n can also be expressed in terms of

$$U_n(x) = \frac{dP_n(x)}{dx} = -\frac{P_n^1(x)}{\sqrt{1-x^2}} = C_{n-1}^{(3/2)}(x), \tag{14}$$

in which $C_n^{(m)}$ are the Gegenbauer (or ultraspherical) polynomials. Thus, we can express the time domain Green's function as

$$g_0(R,\theta,\phi;R',\theta',\phi';t) = \frac{1}{8\pi RR'}\sum_{n=0}^{\infty}\left\{(-1)^n(2n+1) \times\left[P_n\left(\frac{t}{R_<}\right)p\left(\frac{t}{R_<}\right)\right]*U_n(R_>,t-R_>) \times\sum_{p=e,o}\sum_{m=0}^{n}\epsilon_m\frac{(n-m)!}{(n+m)!}Y_{pnm}(\theta,\phi)Y_{pnm}(\theta',\phi')\right\}, \tag{15}$$

where $*$ stands for the operation of convolution. A few of the outgoing wave functions U_n are given below:

$$U_0(R_>,t-R_>) = \delta(t-R_>), \tag{16}$$

$$U_1(R_>,t-R_>) = \delta(t-R_>)+\frac{1}{R_>}u(t-R_>), \tag{17}$$

$$U_2(R_>,t-R_>) = \delta(t-R_>)+\frac{3}{R_>}u(t-R_>) + \frac{3}{R_>^2}(t-R_>)u(t-R_>). \tag{18}$$

In general $U_n(R_>,t-R_>)$ consists of a delta function striking at $t=R_>$ and an nth order polynomial in $(t-R_>)$ beginning at $t=R_>$. A recursive relation for the functions $U_n(R_>,t-R_>)$ can be found very easily by using the recursive relations for the Gegenbauer polynomials. Note that a comparison of Eq. (9) and Eq. (15) implies

$$P_n\left(\frac{R^2+R'^2-t^2}{2RR'}\right)p\left(\frac{t-R_>}{R_<}\right) = (-1)^n\left[P_n\left(\frac{t}{R_<}\right)p\left(\frac{t}{R_<}\right)\right]*U_n(R_>,t-R_>), \tag{19}$$

which we have also confirmed numerically for a large number of orders n.

IV. SCATTERING BY A SOFT SPHERE

In this section we present as an application of the expressions given above to the formulation of the scattering by an acoustically soft sphere.

Consider a soft sphere of radius a centered at the origin insonified by an incident wave. The form of the Green's function given in Eq. (15) suggests that the incident field can be be expressed as

$$\psi^i(\mathbf{R},t)=\sum_{p=e,o}\sum_{n=0}^{\infty}\sum_{m=0}^{n}\frac{d_{nm}}{R}a_{pnm}(t) *\left[P_n\left(\frac{t}{R}\right)p\left(\frac{t}{R}\right)\right]Y_{pnm}(\theta,\phi), \tag{20}$$

where

$$d_{nm}=\frac{(-1)^n(2n+1)(n-m)!}{8\pi(n+m)!}\epsilon_m. \tag{21}$$

For a known ψ^i, or for a known impressed source, a_{pnm} are known. Particularly, for an impulsive plane wave incident from the positive z axis (from $\theta=0$), i.e., $\psi^i=\delta(t+z)$

$$\psi^i(\mathbf{R},t)=\sum_{n=0}^{\infty}\frac{4\pi d_{n0}}{R}\left[P_n\left(\frac{t}{R}\right)p\left(\frac{t}{R}\right)\right]. \tag{22}$$

Again, inspired by the form of the Green's function given in Eq. (15), let the scattered field be given by

$$\psi^s(\mathbf{R},t)=\sum_{n=0}^{\infty}\frac{4\pi d_{n0}}{R}P_n(\cos\,\theta)c_n(t) *\left[P_n\left(\frac{t}{a}\right)p\left(\frac{t}{a}\right)\right]*U_n(R,t-R), \tag{23}$$

where $c_n(t)$ are, as yet, unknown. Imposing the boundary condition on the surface of the sphere, i.e.,

$$(\psi^i+\psi^s)|_{R=a}=0, \tag{24}$$

which implies

$$\left[P_n\left(\frac{t}{a}\right)p\left(\frac{t}{a}\right)\right]+c_n(t)*\left[P_n\left(\frac{t}{a}\right)p\left(\frac{t}{a}\right)\right]*U_n(a,t-a)=0, \tag{25}$$

or

$$c_n(t)*U_n(a,t-a)=-\delta(t). \tag{26}$$

In Eq. (25) the impulsive plane wave mentioned above was assumed to be incident, hence Eq. (22) was used for ψ^i. Note that the orthogonality of the wave functions implies that they are linearly independent; the latter property enables us to interpret Eq. (24) as a *term-by-term equality*, leading to Eq. (25). Solving for $c_n(t)$ and substituting into Eq. (23) yields the scattered field. The solution of Eq. (26) can be achieved either by direct deconvolution or by using system identification techniques (Goodwin and Sin, 1984).

As a numerical example, the scattering by a soft sphere of radius $a=1$ in response to a plane wave incident from the $+z$ axis is given. The incident field is a Gaussian waveform given by $e^{-8t^2/9}$, which is roughly of duration 6. This corresponds to a scatterer of radius $a=1$ m and an incident waveform duration of 20 ns. The incident field is such that at $t=0$, it occupies a region 6 m thick situated symmetrically around $z=0$ provided that the wave velocity is that of light in vacuum. To find the scattered field in response to such an incident field, we simply convolve the (Gaussian) waveform with Eq. (23). Shown in Figure 1 are two solutions to the scattered field in the backscatter direction at a distance of $R=3$ m. The solid curve is obtained using the present formulation, while the cross marks are obtained by inverse Fourier transforming the well known frequency domain solution (Bowman *et al.*, 1969). Direct deconvolution was used to solve Eq. (26) for the time domain solution. Both solutions are obtained by including only five terms in the series expansion. The excellent agreement of the two solutions is worth noting.

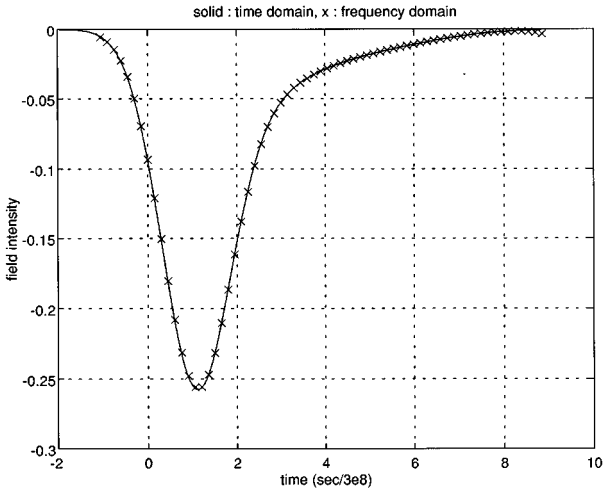


FIG. 1. Scattered field intensity versus time at $R=3$ m due to a Gaussian pulse waveform plane wave.

We have found that the solution converges rapidly if the dimension of the scatterer is small compared to *the wavelength at the highest frequency component of the incident field*. This is expected for those familiar with the solution of similar problems in the frequency domain.

V. CONCLUSION

Two expressions for the time domain free-space scalar Green's function have been presented. The expressions are in the form of an expansion in terms of wave functions which are orthogonal over a spherical surface. As opposed to one of the expansions, the other involves a radial dependence on the source and field points which is ''separated'' in that the expression is in the form of a function of R convolved with a function of R'. The ''compact'' expansion is seen to be more efficient to use to find the field when either impressed or induced sources (or both) are known in that it saves one extra convolution operation for each term in the expansion. On the other hand, the ''separated'' expansion is more suitable to obtain an equation for the unknown equivalent sources.

The utility of the expansion(s) (the two are equivalent) is demonstrated by the use of an example, namely that of acoustical scattering by a soft sphere, where the solution was performed entirely in the time domain.

Finally, we foresee that the Green's function expressions (particularly the ''separated'' version) given here can be used to find the time domain field scattered by small objects of arbitrary shape using an equivalent of the T-matrix

method (where such expansions of the Green's function are used extensively) formulated in the time domain. The meaning of ''small'' is given at the end of Sec. IV; it depends on the frequency content of the incident waveform. This, as well as a formulation of a (dyadic) Green's function useful for vector scattering problems is planned (and suggested) future work.

APPENDIX

To obtain the inverse Fourier transform of $\{-jkR_>h_n^{(2)}(kR_>)j_n(kR_<)\}$, we first state the following properties of the Fourier transform:

Theorem: Let $\mathscr{F}\{g(t)\}=G(\omega)$. Then, provided that $G(0)=0$, we have

$$\mathscr{F}\left\{\int_{-\infty}^{t} g(\tau)d\tau\right\}=\frac{1}{j\omega}G(\omega), \tag{A1}$$

or equivalently

$$\mathscr{F}^{-1}\left\{\frac{1}{j\omega}G(\omega)\right\}=\int_{-\infty}^{t} g(\tau)d\tau=u(t)*g(t). \tag{A2}$$

Corollary: If $\lim_{w\to 0}\ G(\omega)/(j\omega)^{n-1}=0$, then

$$\mathscr{F}^{-1}\left\{\frac{G(\omega)}{(j\omega)^n}\right\}=\frac{t^n}{n!}u(t)*g(t). \tag{A3}$$

Proof: (By induction)

$$\begin{aligned}\mathscr{F}^{-1}\left\{\frac{1}{j\omega}\frac{G(\omega)}{(j\omega)^n}\right\}&=u(t)*\frac{t^{n-1}}{(n-1)!}u(t)*g(t),\\ &=\left(\int_{-\infty}^{\infty}u(t-\tau)\frac{\tau^{n-1}}{(n-1)!}u(\tau)d\tau\right)*g(t),\\ &=\left(u(t)\int_0^t\frac{\tau^{n-1}}{(n-1)!}d\tau\right)*g(t),\\ &=\frac{t^n}{n!}u(t)*g(t).\end{aligned} \tag{A4}$$

It must also be noted that

$$\mathscr{F}^{-1}\{G(\omega)\}=g(t)=\delta(t)*g(t). \tag{A5}$$

Next, we write the spherical Hankel functions as (Abramowitz and Segun, 1972)

$$h_n^{(2)}(z)=-\sum_{i=0}^{n}\frac{j^n(n+i)!}{i!(n-i)!2^i}\frac{e^{-jz}}{(jz)^{i+1}}. \tag{A6}$$

Thus,

$$-jkR_>h_n^{(2)}(kR_>)=\sum_{i=0}^{n}a_{ni}\frac{e^{-jkR_>}}{(jkR_>)^i}, \tag{A7}$$

where

$$a_{ni}=\frac{j^n(n+i)!}{i!(n-i)!2^i}. \tag{A8}$$

We also note that as $z\to 0$, $j_n(z)\to z^n/(2n+1)!!$, where $(2n+1)!!=1\cdot 3\cdot 5\cdots(2n+1)$ is the product of odd integers from 1 to $(2n+1)$. Therefore we have

$$\lim_{k\to 0}\frac{1}{(jkR_>)^i}j_n(kR_<)=0 \quad \text{for } i=0,1,\ldots,n-1. \tag{A9}$$

Thus

$$\begin{aligned}&\mathscr{F}^{-1}\{-jkR_>j_n(kR_<)h_n^{(2)}(kR_>)\}\\ &=\mathscr{F}^{-1}\left\{\sum_{i=0}^{n}a_{ni}\frac{e^{-jkR_>}}{(jkR_>)^i}j_n(kR_<)\right\}\\ &=\left\{(j)^n\delta(t-R_>)+\sum_{i=1}^{n}\frac{a_{ni}}{R_>}\left[\frac{(t-R_>)^i}{i!}u(t-R_>)\right]\right\}\\ &\quad *\left\{\frac{j^n}{2R_<}P_n\left(\frac{t}{R_<}\right)p\left(\frac{t}{R_<}\right)\right\}.\end{aligned} \tag{A10}$$

After some manipulation, it can be shown that

$$\begin{aligned}&(j)^n\delta(t-R_>)+\sum_{i=1}^{n}\frac{a_{ni}}{R_>}\left[\frac{(t-R_>)^i}{i!}u(t-R_>)\right]\\ &=(j)^n\frac{d}{dt}\left[P_n\left(\frac{t}{R_>}\right)u(t-R_>)\right]\\ &=(j)^nU_n(R_>,t-R_>).\end{aligned} \tag{A11}$$

Hence, we say that

$$\begin{aligned}\mathscr{F}_n^{-1}\{-j\omega h_n^{(2)}(\omega)\}&=(j)^n\frac{d}{dt}[P_n(t)u(t-1)]\\ &=(j)^nU_n(1,t-1),\end{aligned} \tag{A12}$$

in the sense that for any function $G(\omega)$ whose inverse Fourier transform is $g(t)$, if $\omega^nG(\omega)$ remains bounded as $\omega\to 0$, than the inverse Fourier transform of the product $(-j\omega)h_n^{(2)}(\omega)G(w)$ is given by the convolution of $g(t)$ and $j^nU_n(1,t-1)$.

Abramowitz, M., and Stegun, I. E., *Handbook of Mathematical Functions* (Dover, New York, 1972).

Bowman, J. J., Senior, T. B. A., and Uslenghi, P. L. E., *Electromagnetic and Acoustic Scattering by Simple Shapes* (North-Holland, Amsterdam, 1969).

Chew, W. C., *Waves and Fields in Inhomogeneous Media* (Van Nostrand Reinhold, New York, 1990).

Felsen, L. B., ''Alternative Field Representations in Regions Bounded by Spheres, Cones, and Planes,'' IRE Trans. Antennas Propag. **AP-5**, 109–121 (1957).

Goodwin, G. C., and Sin, K. S., *Adaptive Filtering Prediction and Control* (Prentice-Hall, Englewood Cliffs, NJ, 1984).

Gradshteyn, I. S., and Ryzhik, L. M., *Table of Integrals, Series, and Products* (Academic, Orlando, 1980).

T. B. Hansen, ''Spherical Expansions of Time-Domain Acoustic Fields: Application to Near Field Scanning,'' J. Acoust. Soc. Am. **98**, 1204–1215 (1995).

Harrington, R. F., *Time-Harmonic Electromagnetic Fields* (McGraw-Hill, New York, 1961).

E. H. Heyman, and A. J. Devaney, ''Time-Dependent Multipoles and Their Application for Radiation from Volume Source Distributions,'' J. Math. Phys. (N.Y.) **37**, 682–692 (1996).

Jones, D. S., *Acoustic and Electromagnetic Waves* (Oxford U. P., Oxford, 1986).

An acoustic boundary element method using analytical/numerical matching

Ronald J. Epstein and Donald B. Bliss
Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708-0300

(Received 27 July 1995; revised 5 July 1996; accepted 8 July 1996)

Analytical/numerical matching (ANM) is a hybrid scheme combining a low-resolution global numerical solution with a high-resolution local solution to form a composite solution. ANM is applied to a harmonically oscillating body to calculate the radiated acoustic field and the associated fluid loading. The approach utilizes overlapping smoothed dipoles, and local corrections to calculate the dipole strength distribution along the surface of the body. A smoothing length scale is introduced that is larger than the smallest physical scale, and smaller than the largest physical scale. The global low-resolution solution is calculated numerically using smoothed dipole solutions to the wave equation, and converges quickly. Local corrections are done with high-resolution local analytical solutions. The global numerical solution is asymptotically matched to the local analytical solutions via a matching solution. The matching solution cancels the global solution in the near field, and cancels the local solution in the far field. The method is very robust, offering insensitivity to node location. ANM provides high-resolution calculations from low-resolution numerics with analytical corrections, while avoiding the usual subtleties involving singular integral equations, and their numerical implementation. The method is applied to calculate the radiated acoustic field and surface pressure of various flat plate configurations in two dimensions. An oscillating rigid flat plate, a forced elastic flat plate, plane-wave diffraction, and mechanical impedance calculations are addressed. [S0001-4966(97)01012-0]

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

INTRODUCTION

Boundary element methods have become a cornerstone of modern computational acoustics. These methods construct the acoustic field, radiated or scattered from a body, using a linear superposition of simple building-block solutions distributed over the body surface. The most important feature of the method is that it only requires computations on the radiating surface or boundary. In contrast, other computational methods can require calculations throughout the entire volume of the computational domain. For applications with an infinite domain, the boundary element method offers a clear advantage over methods involving a grid throughout the entire domain. Boundary element methods have been used for many different engineering applications, including: acoustics, aeroacoustics, aerodynamics, fluid dynamics, electromagnetics, and elastodynamics.

Although boundary element methods have been in use for a number of years, and have been subsequently the subject of much research and refinement, they possess some shortcomings. Typically, boundary element methods are sensitive to the location of computational node points. These are the points at which imposed boundary conditions are applied. Specific rules regarding node location must be strictly adhered to, or incorrect answers can result. In many acoustic applications, the node location must be at the center, or centroid, of each computational element. Usually, the elements must also be of regular shape and arrangement. This is at least partly because of the aforementioned node sensitivity. Problems may also arise when element sizes or orientations vary rapidly, as is the case with complex geometries. There are additional issues concerning numerical convergence and accuracy. In general, the computational burden is greater when trying to accurately resolve regions with large gradients, such as structural discontinuities, or edges. Finally, there are other difficulties and subtleties associated with implementing the method. Boundary element methods are the numerical embodiment of a singular integral equation having a kernel function whose discretization is open to interpretation, leading to the possibility of ambiguity in the meaning of computed results.

The foundations of the boundary element method lie in the early work done on wing and body aerodynamics. Some of the earliest papers written in the field are by Munk[1] and von Karman.[2] They modeled ideal fluid flows about airship hulls and airfoils by embedding a source distribution within the body, and applying boundary conditions appropriately at the surface. Strictly speaking, this is not a boundary element method but rather an embedded singularity technique; however, it embodies many of the essential aspects of the boundary element method. Later, at the Douglas Aircraft Co., Hess and Smith[3,4] pioneered the use of boundary elements for steady nonlifting bodies by using surface distributions of source elements. Around the same time, some of the first applications in acoustics appeared. Friedman and Shaw[5,6] published a paper on transient acoustic scattering, and soon after a boundary element approach to time harmonic acoustic scattering problem was developed in a paper by Goldsmith and Banaugh.[7] More recently, an aero/acoustic formulation has been developed by Farassat[8] based on an integral form of the Ffowcs Williams–Hawkings equation. Many of the theo-

retical aspects of the traditional boundary element method are covered in the works by Colton and Kress,[9] and Filippi.[10] Over the years, many other boundary-element-based techniques have been considered. Some of the most recent boundary element research includes the work of Wu,[11–13] Coyette,[14–16] and Seybert[17–19] among many others. The reader may wish to consult the review papers of Ashley *et al.*,[20] Shaw,[21–23] Hess,[24] and Habault,[25] for a more complete review of boundary element methods.

The present research involves a novel and innovative reformulation of the fundamental approach to boundary element or singularity methods, with the goal of alleviating some of the shortcomings and difficulties associated with the methods. The basis for this reformulation is a new analysis technique called analytical/numerical matching (ANM).

I. ANALYTICAL NUMERICAL MATCHING

Analytical/numerical matching is a general analysis method originally developed at Duke University by D. B. Bliss. The method was first applied to problems in vortex dynamics and rotorcraft free wake analysis.[26–32] Recently, ANM has been applied to problems involving acoustic radiation and structural-acoustic scattering from fluid-loaded structures with discontinuities.[33] In all these cases, accurate solutions were obtained with significant reductions in computational cost. The present work involves the further development of ANM and its application to the acoustic boundary element method.

ANM is a hybrid technique that combines analytical and numerical solutions by a matching procedure. ANM allows a global low-resolution numerical solution and a local high-resolution analytical solution to be combined formally by asymptotic matching to construct an accurate composite solution.

ANM finds accurate solutions to physical problems having small scales or rapid variations that challenge the accuracy of the numerical method. In ANM, an artificial smoothing of the physical problem is introduced. The smoothing length scale must be smaller than the large length scales in the problem, but larger than the scale of numerical discretization. Because the smoothing length scale is larger than the scale associated with the numerical discretization, the numerical solution of the smoothed problem is very accurate. However, the actual problem has a physical length scale smaller than the numerical discretization. The local region associated with the small scale is solved separately (usually analytically, but sometimes numerically) as a high-resolution local problem that captures the small scales and rapid variations. This local problem, because of its idealizations, becomes invalid with increasing distance from the local region of rapid change.

The numerical problem and the local problem are combined by asymptotic matching to form a composite solution. ANM utilizes a matching procedure similar to the method of matched asymptotic expansions (MAE), but otherwise it differs from MAE in several important ways. The ANM approach requires a matching solution that is similar to the local problem but is solved with the smoothing imposed. The smoothing length scale is the largest scale associated with the local region, but is still small compared to global scales. This scale separation allows these local solutions (high resolution and matching) to be solved with simplified geometry. The composite solution is then given by the low-resolution global numerical solution plus the high-resolution local solution minus the matching solution, namely:

$$\begin{aligned}\textit{Composite soln.} = {} & \textit{low-resolution numerical soln.} \\ & + \textit{high-resolution local soln.} \\ & - \textit{smoothed local matching soln.}\end{aligned}$$

In the local region, the matching solution subtracts away the local error associated with the smoothed numerical solution, leaving the local solution. Far from the local region, the local solution and the matching solution cancel, since they become identical beyond the smoothed region. For the method to work well, the smoothing must be chosen to achieve a mathematical ''overlap'' so that the transition zone between the local and numerical solutions is accurate.

II. PROBLEM FORMULATION

The following sections will present a concise review of the classical boundary element formulation, discuss the concept of smoothed acoustic singularities and finally, using the classical integral equation as a starting point, develop the form of the integral equation used in ANM with smoothed acoustic dipoles.

A. Integral equation development

Following what is called the ''indirect'' approach,[23] the integral equation will represent any physical boundary with a distribution of acoustic dipole sources whose strengths are adjusted to yield solutions imposed by the boundary conditions. Note that the ''direct'' approach[23] leaves the integral equation in terms of surface velocity and potential, and does not explicitly identify monopole and dipole source terms.

The following derivation is based on one of Green's identities[34] taking into account the appropriate boundary conditions. For clarity, only wave motions in three dimensions will be considered. The results for two-dimensional wave motions follow by simple analogy and will be given. Following Baker,[35] the velocity potential for small amplitude sound waves, in a fixed frame of reference, satisfies the acoustic wave equation

$$\nabla^2 \varphi = \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2}, \tag{1}$$

where φ is the acoustic velocity potential, c is the speed of sound, and t is time. Consider a body with a specified velocity distribution on it's surface. The Neumann boundary condition

$$\frac{\partial \varphi}{\partial n} = V_s \tag{2}$$

applies on the body surface, where V_s is a specified normal velocity.

Assuming harmonic disturbances, or taking a Fourier transform with respect to t, the velocity potential becomes, $\varphi = \hat{\varphi} e^{ikct}$, and the wave equation is written

$$\nabla^2\hat{\varphi} = -k^2\hat{\varphi}, \tag{3}$$

where k is the wave number defined by $k=\omega/c$, and ω is the angular frequency. This equation, referred to as the Helmholtz equation, reduces to Laplace's equation when $k=0$.

A useful form of Green's identity[34] is given by

$$\int\int\int_V (\hat{\varphi}\nabla^2\hat{\psi} - \hat{\psi}\nabla^2\hat{\varphi})\,dV = -\int\int_S \left(\hat{\varphi}\frac{\partial\hat{\psi}}{\partial n} - \hat{\psi}\frac{\partial\hat{\varphi}}{\partial n}\right) dS, \tag{4}$$

where S is defined as the surface bounding the volume V, and $\partial/\partial n$ denotes differentiation along the outward normal to S. This identity is valid only when $\hat{\psi}$ and $\hat{\varphi}$ are sufficiently well behaved, that is to say that their first- and second-order partial derivatives are continuous within and on the boundary S. When considering multiply connected regions, appropriate branch cuts can be introduced. Assuming $\hat{\varphi}$ and $\hat{\psi}$ satisfy Helmholtz's equation everywhere in V, the left-hand side of Eq. (4) vanishes, giving

$$\int\int_S \left(\hat{\varphi}\frac{\partial\hat{\psi}}{\partial n} - \hat{\psi}\frac{\partial\hat{\varphi}}{\partial n}\right) dS = 0. \tag{5}$$

Now, let $\hat{\psi}$ be defined as $\hat{\psi}=e^{-ikr}/r$, where

$$r = \sqrt{(x-x_0)+(y-y_0)+(z-z_0)} \tag{6}$$

denotes the distance from the location of the origin of $\hat{\psi}$ at the fixed point $p=(x_0,y_0,z_0)$ to a point of evaluation located at (x,y,z) in V. This expression for $\hat{\psi}$ is the potential for a harmonic acoustic source of strength 4π, which is the three-dimensional free-space Green's function for the Helmholtz equation. After substitution, Eq. (5) becomes

$$\int\int_S \left\{\hat{\varphi}\frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right) - \frac{e^{-ikr}}{r}\frac{\partial\hat{\varphi}}{\partial n}\right\} dS = 0, \tag{7}$$

where it is assumed that the point p is not within the volume V and therefore there are no singularities within V. If p lies inside the boundary S, and any singularities of $\hat{\varphi}$ are outside S, Eq. (7) no longer holds, because $\hat{\psi}$ becomes infinite at p. To avoid this difficulty, we apply Green's identity, Eq. (4), to the volume V bounded externally by S and internally, at p, by a small sphere bounded by the surface S_ϵ, with center at p, and small radius ϵ. Now, $\hat{\varphi}$ and $\hat{\psi}$ satisfy Helmholtz's equation everywhere in V, therefore

$$\int\int_S \left\{\hat{\varphi}\frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right) - \frac{e^{-ikr}}{r}\frac{\partial\hat{\varphi}}{\partial n}\right\} dS = \lim_{\epsilon\to 0}\int\int_{S_\epsilon} \left\{\hat{\varphi}\frac{e^{-ikr}}{r}\left(-ik-\frac{1}{r}\right) - \frac{e^{-ikr}}{r}\frac{\partial\hat{\varphi}}{\partial r}\right\} dS. \tag{8}$$

After evaluation, the integral on the right-hand side of Eq. (8) in the limit becomes $-4\pi\hat{\varphi}(p)$ if p is inside V, and $-2\pi\hat{\varphi}(p)$ if p (Ref. 35) is on the boundary S. Therefore for any point p with V we have

$$4\pi\hat{\varphi}(p) = \int\int_S \left\{\frac{e^{-ikr}}{r}\frac{\partial\hat{\varphi}}{\partial n} - \hat{\varphi}\frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right)\right\} dS, \tag{9}$$

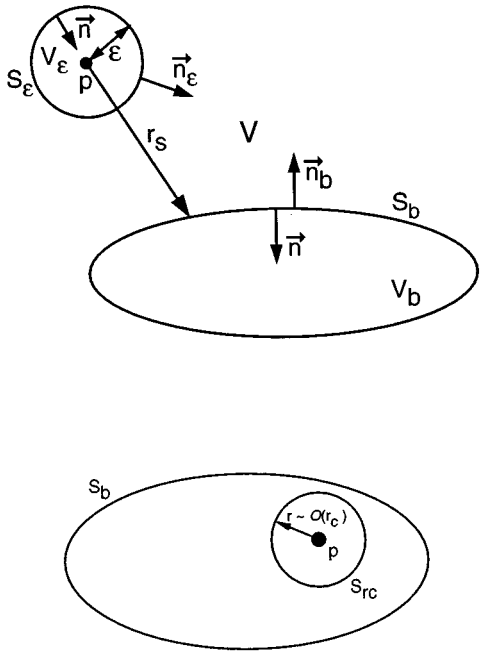


FIG. 1. Nomenclature for acoustic potential derivation: Body surrounded by an acoustic medium (top); subsurface S_{r_c} on S_b (bottom).

where r is the distance from p to the surface S.

For a body surrounded by an acoustic medium, as shown in Fig. 1, S becomes the boundary of the body S_b, and V the domain exterior to the body. Rigorously speaking, there is an additional boundary at infinity bounding V; however, the contribution from the right-hand side of Eq. (9) becomes negligible as $r\to\infty$ and this boundary need not be considered explicitly.[36] In the preceding analysis a normal vector to a surface has been defined such that it points out of the volume bound by that surface. Therefore, the normal vector on S_b points in the opposite direction than that of the normal vector of the surface bounding V.

Within the body, exterior to V, not containing p, Eq. (7) is valid. Now denote $\hat{\varphi}$ in Eq. (7) as $\hat{\varphi}_i$, the ''inner'' potential. Exterior to the body, within V, Eq. (9) is valid if the small sphere S_ϵ is taken about the point p. Again denote $\hat{\varphi}$ in Eq. (9) as $\hat{\varphi}_o$, the ''outer'' potential. Adding Eq. (7) to Eq. (9), and remembering that the normal vector to the surface S_b in each equation points in opposite directions, yields the expression

$$\hat{\varphi}(p) = \frac{1}{4\pi}\int\int_{S_b}\left\{\frac{e^{-ikr}}{r}\left(\frac{\partial\hat{\varphi}_o}{\partial n} - \frac{\partial\hat{\varphi}_i}{\partial n}\right) - (\hat{\varphi}_o - \hat{\varphi}_i)\frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right)\right\} dS, \tag{10}$$

Eq. (10) provides the value of $\hat{\varphi}(p)$ in terms of $\Delta\hat{\varphi}$ and $\Delta\partial\hat{\varphi}/\partial n$ on the boundary of V_b.

Using Stoke's theorem[37] and the definition of dipole strength, the difference between the ''external'' and ''internal'' values of $\hat{\varphi}$ can be related to the dipole source strength distribution on the surface,[36]

$$-\mu_\omega = \hat{\varphi}_o - \hat{\varphi}_i. \tag{11}$$

Similarly, by the definition of monopole source strength, the difference between the normal derivatives of $\hat{\varphi}$ can be related to the monopole source strength distribution on the surface,

$$-\hat{\sigma}_\omega = \frac{\partial \hat{\varphi}}{\partial n} - \frac{\partial \hat{\varphi}_i}{\partial n}. \tag{12}$$

Making the appropriate substitutions in Eq. (10) with Eqs. (11)–(12),

$$\hat{\varphi}(p) = -\frac{1}{4\pi}\int\int_{S_b}\left\{\hat{\sigma}_\omega \frac{e^{-ikr}}{r} - \mu_\omega \frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right)\right\} dS, \tag{13}$$

Eq. (13) gives $\hat{\varphi}(p)$ in terms of a distribution of acoustic sources and dipoles along the boundary S_b. This expression is valid for points within, or external to, the volume V_b.

To satisfy the Neumann boundary condition at the surface, the specified normal velocity V_s must be equal to the velocity induced by the integral distribution of acoustic monopoles and dipoles, leading to the expression,

$$V_s(p) = \boldsymbol{n}\cdot\left\{\frac{1}{4\pi}\int\int_{S_b}\left\{\mu_\omega \nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right)\right] - \hat{\sigma}_\omega \nabla\left[\frac{e^{-ikr}}{r}\right]\right\} dS\right\}. \tag{14}$$

The acoustic monopole source term is not always required, and valid solutions can be composed with only dipole distributions.[36] In fact, if a continuous normal velocity across the boundary is desired, then

$$\frac{\partial \hat{\varphi}_o}{\partial n} = \frac{\partial \hat{\varphi}_i}{\partial n} \rightarrow \hat{\sigma}_\omega = \frac{\partial \hat{\varphi}_i}{\partial n} - \frac{\partial \hat{\varphi}_o}{\partial n} = 0. \tag{15}$$

Therefore, writing Eq. (14) without the acoustic monopole term

$$V_s(p) = \boldsymbol{n}\cdot\left\{\frac{1}{4\pi}\int\int_{S_b}\left\{\mu_\omega \nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right)\right]\right\} dS\right\}, \tag{16}$$

where Eq. (16) is a homogenous Fredholm integral equation of the first kind relating the unknown dipole strength distribution on S_b to the specified normal velocity on S_b.

In two dimensions, the expression analogous to Eq. (16) is given by

$$V_s^{2D}(p) = \boldsymbol{n}\cdot\left\{\frac{i}{4}\int_{\Gamma_b}\left\{\mu_\omega^{2D}\nabla\left[\frac{\partial}{\partial n}(H_1^{(2)}(kr))\right]\right\} dS\right\}, \tag{17}$$

where Γ_b is a closed curve in space that surrounds the two-dimensional body. Equation (17) involves Hankel functions of the first kind, which give outgoing cylindrical waves; however, Hankel functions of the first kind rather than the second kind would arise if the harmonic form $e^{-i\omega t}$ were chosen.

The boundary element method, which is based on the above formulation, is created when the surface is discretely approximated by a contiguous set of small elements or panels. Typically, the panels are quadrilaterals on which singularities are distributed in a continuous manner; on each element, a node point is identified. In many cases the node point must be at the center of the computational element. The singularity strength can be constant over the panel or vary in some higher-order fashion. The integral equation, and the boundary conditions are evaluated at the node points. This formulation gives rise to a set of algebraic equations for the unknown singularity strength μ_ω, based on the discrete approximation of Eqs. (16) and (17). The form of the kth equation is given by

$$\left[\frac{\partial \varphi}{\partial n}\right]_k = \frac{1}{4\pi}\sum_{\text{body}} A_{kj}\mu_j. \tag{18}$$

The influence coefficients A_{kj} depend on the frequency, speed of sound and body geometry; not the boundary conditions, and therefore, are fixed for harmonic body motions.

The radiated acoustic pressure field and the surface pressure distribution can be calculated from the dipole strength distribution μ_ω, solved for discretely in Eq. (18). The radiated field is calculated by backsubstituting μ_ω into a surface distribution of dipoles, and calculating the radiation directly in the field. The surface pressure can be calculated from the unsteady momentum equation. On the surface of a flat plate this gives,

$$\Delta p = -i\rho c k \mu_\omega, \tag{19}$$

where Δp denotes the jump in pressure across the surface. Due to the symmetric nature of the flat plate problem, the surface pressure can then be written as

$$p = \mp\frac{i\rho c k \mu_\omega}{2}, \tag{20}$$

where the sign depends if evaluation is taken on an upper or lower surface.

B. Smoothed acoustic dipoles

Before preceding to develop the ANM method, the smoothed dipole, which is the basic building block of the method, will be described. The concept of smoothing singularities has its origin in the field of computational vortex dynamics.[26,30–32,38,39] The following discussion describes the development of smoothed acoustic dipoles as a natural extension of the concepts used in vortex dynamics.

In three dimensions, the harmonic monopole source solution to the acoustic wave equation can be written

$$P(x,y,z,t) = \frac{ik\rho c}{4\pi r}\sigma_\omega e^{ik(ct-r)} = ik\rho c\sigma_\omega g_\omega(\boldsymbol{r}|\boldsymbol{r}_0)e^{i\omega t}. \tag{21}$$

In Eq. (21), the Green's function is defined as

$$g_\omega(\boldsymbol{r}|\boldsymbol{r}_0) = \frac{1}{4\pi r}e^{-ikr} \tag{22}$$

and

$$r^2 = |\boldsymbol{r}-\boldsymbol{r}_0|^2 = (x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2, \tag{23}$$

where the source point is $\boldsymbol{r}_0 = (x_0, y_0, z_0)$, the point of evaluation $\boldsymbol{r} = (x,y,z)$, the source strength is σ_ω, the density is ρ, and $k = \omega/c$. Performing a simple transformation on the radial variable,

$$r^2 \Rightarrow r^2 + r_c{}^2 = |\boldsymbol{r}-\boldsymbol{r}_0|^2 + r_c{}^2 = (x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2 + r_c{}^2, \quad (24)$$

where the smoothing length scale r_c has been added radially. After transformation Eq. (21) becomes

$$P(x,y,z,t) = \frac{ik\rho c}{4\pi\sqrt{r^2+r_c{}^2}}\sigma_\omega e^{ik(ct-\sqrt{r^2+r_c{}^2})}. \quad (25)$$

This transformed form of Eq. (21) is an acoustic monopole with algebraic smoothing. It is no longer singular at the source point. The source strength is now distributed smoothly over a region. As a result, the pressure approaches a constant value as $r \to 0$. Note, however, as r gets large, namely $r \gg r_c$, the pressure tends quickly toward the potential value, the value without smoothing. In the limit when $r_c \to 0$, Eq. (21) is recovered. As shown, the deviation from the potential result depends on both the magnitude r_c and r. The smoothed monopole can be interpreted as a convolution of a singular monopole with a suitable smoothing function.[40] There is no unique type of smoothing; therefore, other types of smoothing are possible.[40,41] It follows that, the potential for a smoothed acoustic monopole can be written

$$\hat{\varphi}(x,y,z,t) = \frac{\sigma_\omega}{4\pi}\frac{e^{ik(ct-\sqrt{r^2+r_c{}^2})}}{\sqrt{r^2+r_c{}^2}}. \quad (26)$$

The algebraically smoothed dipole can be derived by differentiation with respect to the characteristic direction of dipole orientation,

$$P(x,y,z,t) = \boldsymbol{n}\cdot\nabla\left\{\frac{ik\rho c}{4\pi\sqrt{r^2+r_c{}^2}}\sigma_\omega e^{ik(ct-\sqrt{r^2+r_c{}^2})}\right\}. \quad (27)$$

After differentiation,

$$P(x,y,z,t) = -k^2\mu_\omega\frac{\rho c r}{4\pi(r^2+r_c{}^2)} \times\left(1-\frac{i}{k\sqrt{r^2+r_c{}^2}}\right)\cos\theta e^{ik(ct-\sqrt{r^2+r_c{}^2})}, \quad (28)$$

where μ_ω is the dipole strength, and θ is defined relative to the dipole orientation vector and the field point. Equation (28) is a smoothed acoustic dipole and exhibits all the properties discussed with regard to smoothed monopoles. Most importantly, when $r \to 0$ the pressure approaches a constant value, because the dipole strength is distributed. An entire family of smoothed multipole solutions can be derived by differentiating the smoothed monopole. The velocity field of the smoothed dipole is given by integrating the momentum equation in the appropriate coordinate system,

$$\rho\frac{\partial \boldsymbol{u}}{\partial t} = -\nabla P, \quad (29)$$

where $\boldsymbol{u}$ is the desired velocity field.

In general, using an inverse Fourier transform, Eq. (28) can be inverted to the time domain,

$$P(r,t) = \frac{\rho}{4\pi c}\frac{r}{r^2+r_c{}^2}\cos\theta\left\{D''\left(t-\frac{\sqrt{r^2+r_c{}^2}}{c}\right) + \frac{c}{\sqrt{r^2+r_c{}^2}}D'\left(t-\frac{\sqrt{R^2+r_c{}^2}}{c}\right)\right\}, \quad (30)$$

where D is an arbitrary dipole strength function, and $()'$ denotes differentiation with respect to the argument. Equation (30) can be interpreted as a mathematical statement of Huygen's Principle for a dipole acoustic disturbance displaced r_c from the source point.

In two dimensions, the smoothed monopole can be derived by direct integration of Eq. (25) giving

$$P(y,z,t) = -\frac{k\rho c\sigma_\omega}{4}H_0^{(2)}(k\sqrt{\hat{r}^2+r_c{}^2})e^{i\omega t}, \quad (31)$$

where $H_0^{(2)}(x)$ is a Hankel function of the second kind of order 0, and

$$\hat{r}^2 = |\boldsymbol{r}-\boldsymbol{r}_0|^2 = (y-y_0)^2 + (z-z_0)^2. \quad (32)$$

Similar to Eq. (28), the two-dimensional smoothed acoustic dipole can be derived from Eq. (31) by differentiation in the characteristic dipole direction to give

$$P(y,z,t) = -\frac{k\rho c\mu_\omega}{4} \times H_1^{(2)}(k\sqrt{\hat{r}^2+r_c{}^2})\frac{k\sqrt{y^2+z^2}}{\sqrt{y^2+z^2+r_c{}^2}}\cos\theta e^{i\omega t}, \quad (33)$$

where θ is defined relative to the dipole orientation vector and the field point. The corresponding velocity field can be found with the two-dimensional form of the momentum equation, Eq. (29).

The smoothing operation takes a discrete singularity, a doublet in this case, and replaces it with a continuous distribution having the same net strength and the same far-field behavior. As a result, the region over which the singularity is distributed is no longer divergence free, although it is irrotational. The governing equation is then an inhomogeneous form of the wave equation, with a distributed source/sink distribution on the right-hand side.

When the ANM solution procedure separates the problem into local and global parts, as illustrated in Fig. 2, the global solution and the local matching solution are not divergence-free. When the composite solution is constructed for this linear problem, the distributed source/sink effects from the global and matching solutions cancel (after all, their equal and opposite nature was the basis for decomposing the solution in Fig. 2), and the composite solution does satisfy the homogeneous wave equation.

This part of the solution procedure, which involves a decomposition by adding and subtracting smoothed singularities, is actually one of the keys to how the method works.

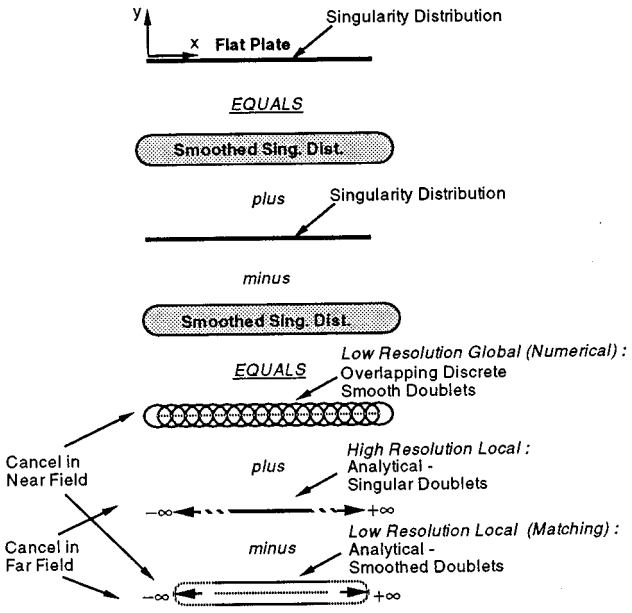


FIG. 2. A schematic representation of the ANM solution methodology for acoustics.

Understanding the role of smoothing is a subtle point, especially since the smoothing cancels out in the final answer. This process is one of the real innovations of the ANM method.

C. ANM acoustics

A distribution of smoothed dipoles, as given by Eq. (26), is added and subtracted to Eq. (16), with no net change in the equation, giving

$$V_s(p)=\boldsymbol{n}\cdot\left\{\frac{1}{4\pi}\int\int_{S_b}\left\{\mu_\omega\nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right)\right]\right\}dS + \frac{1}{4\pi}\int\int_{S_b}\left\{\mu_\omega\nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ik\sqrt{r^2+r_c^2}}}{\sqrt{r^2+r_c^2}}\right)\right]\right\}dS - \frac{1}{4\pi}\int\int_{S_b}\left\{\mu_\omega\nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ik\sqrt{r^2+r_c^2}}}{\sqrt{r^2+r_c^2}}\right)\right]\right\}dS\right\}. \tag{34}$$

Regrouping terms,

$$V_s(p)=\boldsymbol{n}\cdot\left\{\frac{1}{4\pi}\int\int_{S_b}\left\{\mu_\omega\nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ik\sqrt{r^2+r_c^2}}}{\sqrt{r^2+r_c^2}}\right)\right]\right\}dS + \left(\frac{1}{4\pi}\int\int_{S_b}\left\{\mu_\omega\nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right)\right]\right\}dS - \frac{1}{4\pi}\int\int_{S_b}\left\{\mu_\omega\nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ik\sqrt{r^2+r_c^2}}}{\sqrt{r^2+r_c^2}}\right)\right]\right\}dS\right)\right\}. \tag{35}$$

Combining the two integrals in the second term on the right-hand side gives

$$V_s(p)=\boldsymbol{n}\cdot\left\{\frac{1}{4\pi}\int\int_{S_b}\left\{\mu_\omega\nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ik\sqrt{r^2+r_c^2}}}{\sqrt{r^2+r_c^2}}\right)\right]\right\}dS + \frac{1}{4\pi}\int\int_{S_b}\left\{\mu_\omega\nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right) - \frac{\partial}{\partial n}\left(\frac{e^{-ik\sqrt{r^2+r_c^2}}}{\sqrt{r^2+r_c^2}}\right)\right]\right\}dS\right\}. \tag{36}$$

Notice, that the second term on the right-hand side of Eq. (36) becomes negligible when $r\gg r_c$, because the value of the smooth dipole kernel function approaches that of the singular kernel function and they asymptotically cancel each other for large r. The contribution of this integral becomes significant only when $r=O(r_c)$.

The far-field contributions vanish asymptotically for the integral in Eq. (36) as intended. There is no clear distinction between near and far field, only a gradual transition between them. This behavior is intentional, and it is one of the distinguishing features of the method, which is a ''matching'' method—not a ''patching'' method. The final results are independent of the choice of integration areas, as long as they extend into the far field. Similarly, and related to this, the results are independent of the degree of smoothing used, as shown in Fig. 10.

On the surface S_b where V_s is specified, there is a small region S_{r_c} about p where $r=O(r_c)$, see Fig. 1. Note, however, most of the contribution to the integral comes from the region S_{r_c} immediately about p. Therefore, the bounds of integration can be modified such that the integration is done on any surface S_{mod} which contains S_{r_c}, without changing the final result.

Due to the local nature of the integral over S_{mod}, a series expansion about p can be written for the dipole strength density

$$\mu_\omega(\boldsymbol{p}+\boldsymbol{\Delta r})=\mathscr{F}(\boldsymbol{\Delta r})\sum_{k=0}^{\infty}\frac{1}{k!}(\boldsymbol{\Delta r}\cdot\nabla)^k\mu_\omega(\boldsymbol{p}), \tag{37}$$

where $\boldsymbol{\Delta r}$ is the vector change in position on the surface, $\boldsymbol{p}$ is the vector location of p, and $\mathscr{F}(\boldsymbol{\Delta r})$ is a weighting function discussed in the next section. Essentially, $\mathscr{F}(\boldsymbol{\Delta r})$ is included because the dipole strength density may be nonanalytic approaching the edges of a surface. Only the first few terms in the series are necessary; however, at this point the entire series is retained for completeness. Substituting the series expression into Eq. (36), and changing the bounds of integration on the second integral from S_b to S_{mod} gives

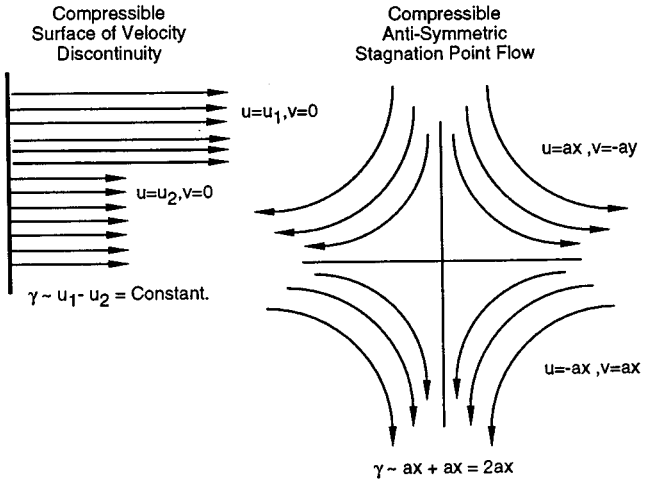


FIG. 3. Examples of simple flow solutions to the Helmholtz equation.

$$V_s(p)=\boldsymbol{n}\cdot\left\{\frac{1}{4\pi}\int\int_{S_b}\left\{\mu_\omega\nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ik\sqrt{r^2+r_c^2}}}{\sqrt{r^2+r_c^2}}\right)\right]\right\}dS+\frac{1}{4\pi}\int\int_{S_{\mathrm{mod}}}\left\{\mathscr{F}\sum_{k=0}^{\infty}\frac{1}{k!}(\Delta\boldsymbol{r}\cdot\nabla)^k\mu_\omega(\boldsymbol{p})\nabla\times\left[\frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right)-\frac{\partial}{\partial n}\left(\frac{e^{-ik\sqrt{r^2+r_c^2}}}{\sqrt{r^2+r_c^2}}\right)\right]\right\}dS\right\}. \tag{38}$$

This can then be rewritten as

$$V_s(p)=\boldsymbol{n}\cdot\left\{\frac{1}{4\pi}\int\int_{S_b}\left\{\mu_\omega\nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ik\sqrt{r^2+r_c^2}}}{\sqrt{r^2+r_c^2}}\right)\right]\right\}dS+\sum_{k=0}^{\infty}\left[\varphi_k^{\mathrm{local}}(p)-\frac{\mathscr{F}}{4\pi}\int\int_{S_{\mathrm{mod}}}\left\{\frac{1}{k!}(\Delta\boldsymbol{r}\cdot\nabla)^k\times\mu_\omega(\boldsymbol{p})\nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ik\sqrt{r^2+r_c^2}}}{\sqrt{r^2+r_c^2}}\right)\right]\right\}dS\right]\right\}, \tag{39}$$

where

$$\varphi_k^{\mathrm{local}}(p)=\frac{\mathscr{F}}{4\pi}\int\int_{S_{\mathrm{mod}}}\left\{\frac{1}{k!}(\Delta\boldsymbol{r}\cdot\nabla)^k\times\mu_\omega(\boldsymbol{p})\nabla\left[\frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right)\right]\right\}dS, \tag{40}$$

$\varphi_k^{\mathrm{local}}$ is the local potential, and in Eq. (40) it is written in terms of an integral over the surface S_{mod}. This integral represents a family of simple flow solutions to Helmholtz's equation, as shown in Fig. 3, which can be found by separation of variables. As such, the explicit integration of Eq. (40) is not required. Equation (39) is the fundamental integral equation of the ANM boundary element method. In two dimensions Eq. (39) becomes

$$V_s^{\mathrm{2D}}(p)=-\boldsymbol{n}\cdot\left(\frac{1}{2\pi}\int_{\Gamma_b}\left\{\mu_{\mathrm{2D}}\nabla\left[\frac{\partial}{\partial n}(H_1^{(2)}(k\sqrt{\hat{r}^2+r_c^2}))\right]\right\}dS+\sum_{k=0}^{\infty}\frac{\mu_{\mathrm{2D}}^{(k)}(p)}{k!}\left\{\varphi_k^{\mathrm{local}}(p)-\frac{\mathscr{F}}{2\pi}\int_{\Gamma_{\mathrm{mod}}}(x-p)^k\times\nabla\left[\frac{\partial}{\partial n}(H_1^{(2)}(k\sqrt{\hat{r}^2+r_c^2}))\right]dS\right\}\right), \tag{41}$$

where

$$\varphi_k^{\mathrm{local}}(p)=\frac{\mathscr{F}}{2\pi}\int_{\Gamma_{\mathrm{mod}}}(x-p)^k\nabla\frac{\partial}{\partial n}(H_1^{(2)}(k\hat{r}))dS. \tag{42}$$

Again $\varphi_k^{\mathrm{local}}$ represents simple flow solutions to Helmholtz's equation, the hat notation denotes two dimensions, Γ_{mod} is a contour that contains Γ_{r_c}, and the series expansion is taken about p.

It may be noted that, if the point p is not on the surface S_b, Eqs. (39)–(41) are integral representations of the normal velocity field, in terms of the normal velocity field on the surface S_b. If the point p is on S_b, Eqs. (39)–(41) become inhomogeneous Fredholm integral equations of the first kind, relating the known normal velocity V_s to the unknown values of μ_ω on S_b. At this point, the benefit of using Eqs. (39)–(41) may not be clear; however, it will be shown that due to their smooth nature these equations are advantageous for numerical implementation.

Examining the terms in Eq. (39) in greater detail reveals the relationship between the original problem, Eq. (16), and the problem cast in terms of ANM. The ANM methodology breaks the problem up into global, and local constituent parts. It will be shown that the first term on the right-hand side of the equation embodies smoothed global effects, ultimately treated numerically, while the second term is a local correction that can be found analytically.

To illustrate the underlying idea in more physical terms consider a flat plate oscillating in an infinite acoustic medium. Figure 2 shows a schematic of the ANM decomposition of the original linear problem. As a first step, equivalent smoothed singularity distributions are added to and subtracted from the original problem, as given in Eq. (34). Because these smoothed distributions exactly cancel, the net result is unchanged. However, as shown in Eq. (36), the singularity distributions can also be regrouped. The first smoothed singularity distribution, the first term on the right-hand side of Eq. (36), can be viewed as the low-resolution smoothed global problem, and represented by discrete overlapping smoothed dipoles, given by Eq. (27), as shown in Fig. 2. The second (original) distribution and the third (smoothed) distribution can be combined, as in Eq. (36). When this is done, these two distributions always cancel asymptotically at points away from the point of evaluation, because their far-field behavior is the same. Thus only the local behavior of these distributions ultimately matters. For convenience they can be evaluated as if they are part of an infinite sheet (or a semi-infinite sheet if near an edge). Using a Taylor series expansion of the singularity strength distribution in this local region, as in Eq. (37), it can be shown that

these local solutions can be expressed in terms of simple flow fields obtained from the governing equation by separation of variables.

An interesting feature of the ANM composite solution, given by Eqs. (39)–(41) is that all its constituent parts are functionally smooth. In fact, it is only vaguely reminiscent of a traditional boundary element method. The surface S_b of Eq. (16) is no longer discretely approximated by a contiguous set of boundary elements, but rather, a discrete distribution of overlapping smoothed dipoles. Therefore, the idea of a surface of contiguous elements exists only with regard to the spacing between adjacent smoothed dipoles.

In traditional methods, singularities such as dipoles are integrated over a small region to make an element. A number of these elements are then used to build up the overall surface. On each element is a node at which boundary conditions are applied. Applying the boundary conditions on the N nodes associated with the N elements leads to an Nth-order linear system that can be solved to determine the element strengths.

In the ANM approach, the numerical part of the problem is still solved as a linear system, very much in the same spirit as the traditional approach. In place of the elements, smoothed overlapping doublets are used. Because of the smoothing and overlapping, these distributed doublets do not need to first be integrated to form elements (although this could be done for higher-order accuracy). A ''discrete'' set of these smoothed doublets accurately approximates a continuous distribution. This fact gives the formulation an additional simplicity. These smoothed doublets play roughly the same role as elements in the traditional method.

The analytical local and matching solutions are employed to remove the unwanted effects of smoothing from the final answer. These solutions do involve a continuous distribution of doublets and smoothed doublets, and they can be found by an integration process somewhat similar to that used to construct elements, although a more direct method using separation of variables is generally superior. Because of the asymptotic nature of this process, these local solutions are not really elements either, but are actually corrections to the smoothed doublets.

Boundary conditions are still applied at nodes, as before. There is one node for each smoothed doublet. The strengths of the smoothed doublets are the unknowns. The strengths of the local and matching solutions are related to the smoothed doublet strengths. All this leads to a linear system for the smoothed doublet strengths, from which all other solution properties can be constructed in a direct manner.

This smoothness leads to accurate, rapidly converging numerics. In addition, the analytical aspects of the problem are well defined and well behaved. The traditional problems associated with discretizing a singular integral equation to obtain numerical solutions are avoided. In contrast to traditional discretized calculations, there is no ambiguity as to whether local averages or point values of variables, such as pressure, are being computed on panels. ANM computes local values of variables, and even handles singular behavior at edges in a formal manner that allows singularity strength to be determined. Sources of error are quantifiable, and there is a procedure for extending the method to higher-order accuracy.

D. ANM local correction

This paper presents results from two-dimensional calculations, therefore, the following section addresses the development of the ANM local correction explicitly in two dimensions. The extension to three dimensions of the following formulation is not difficult, however, for the sake of brevity it will be included with future work.

The local correction is based on the second term on the right-hand side of Eq. (41), which represents the difference between a distribution of singular dipoles and smoothed dipoles. The difference between these distributions cancels far away, therefore the far-field shape and behavior is unimportant, allowing for simplification in the evaluation of the integral. The domain of integration only needs to be geometrically correct immediately about p. In particular, when considering a flat surface, for convenience, the bounds of integration can be taken to infinity. The integration is done on an infinite flat surface (a contour in two dimensions) with p located at the origin, given by

$$\mathscr{L}_w = \boldsymbol{n}\cdot\left\{\frac{i}{4}\int_{-\infty}^{\infty}\left\{\mathscr{F}\sum_{k=0}^{\infty}\frac{\mu_\omega^{2\mathrm{D}^k}(p)}{k!}(x-p)^k\nabla \times\left[\frac{\partial}{\partial n}(H_1^{(2)}(k\hat{r}))-\frac{\partial}{\partial n}(H_1^{(2)}(k\sqrt{\hat{r}^2+r_c{}^2}))\right]\right\}dS\right\}, \tag{43}$$

where $\mathscr{L}_w$ denotes the net local correction. Any effects from the far field will cancel, leaving only the local difference between the smoothed, and singular distributions. Therefore, in practice, only the first few terms in the series need to be retained. Special consideration is made for edges, due to the additional length scale that measures from the tip of the edge to p, leading to a nonanalytic region. The integration is taken over semi-infinite bounds along the edge coordinate direction, starting at the edge, where p is located a distance l inboard of the edge

$$\mathscr{L}_w^{\mathrm{edge}} = \boldsymbol{n}\cdot\left\{\frac{i}{4}\int_{0}^{\infty}\left\{\mathscr{F}\sum_{k=0}^{\infty}\frac{\mu_\omega^{2\mathrm{D}^k}(p)}{k!}(x-p)^k\nabla \times\left[\frac{\partial}{\partial n}(H_1^{(2)}(k\hat{r}))-\frac{\partial}{\partial n}(H_1^{(2)}(k\sqrt{\hat{r}^2+r_c{}^2}))\right]\right\}dS\right\}. \tag{44}$$

The contribution of the distribution of singular dipoles in Eqs. (43) and (44), the first term in the square brackets, is called the high-resolution local solution. It contains the details of the local flow structure, including the local gradients. Looking specifically at this part of the integral, it can easily be shown that the high-resolution local solution consists of exact solutions to the Helmholtz equation on an infinite domain. This is because it is integral distributions of free-space Green's functions with varying strength distributions given by the terms in the series, which explicitly satisfy the Helm-

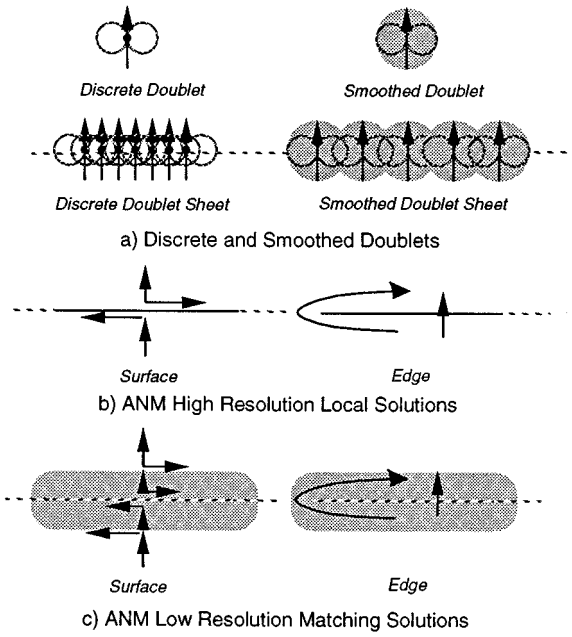


FIG. 4. Building blocks of the ANM solution: (a) discrete and Smoothed dipoles, (b) high-resolution local solutions, (c) smoothed low-resolution local solutions.

holz equation and boundary conditions. As such, by considering the terms in the Taylor series of the dipole strength distribution in Eqs. (43) and (44), the flow field about a node can be deduced in terms of simple exact solutions to the Helmholtz equation on an infinite domain, shown in Fig. 3.

Using the Taylor series as a guide, solutions to the Helmholtz equation can be found that correspond to each term in the series, and therefore represent the integrated distributions of singular Green's functions given in Eqs. (43) and (44). As noted earlier, the series can be truncated at any desired accuracy; however, depending on S_{mod}, care must be taken to assure the resulting integrals are uniformly convergent. In the present implementation, the series was truncated after the quadratic terms, thereby giving the local correction in the form of a truncated asymptotic expansion.

The contribution of the distribution of smoothed dipoles in Eqs. (43) and (44), the second term in the square brackets, is called the matching solution. It cancels the global solution locally, and the high-resolution local solution globally, thereby asymptotically matching the global solution to the local high-resolution solution. The matching solution is precisely a smoothed version of the high-resolution local solution.

The smoothing, as shown in Fig. 4, can be done by any suitable means; however, a technique called the ''split-sheet analogy'' is effective. This approach is based on the observation that the smoothing process is similar to viewing a singular solution at an offset location. In particular, consider the smoothed doublet velocity field, given by Eqs. (28) and (29), when viewed in the $z=0$ plane, which is the surface on which the downwash is evaluated. This expression can be shown to be the same as that for a discrete source and sink displaced a distance r_c above and below the $z=0$ plane, respectively, provided the doublet strength and source strength are related as $\mu=2r_c\sigma$. It follows that a sheet of smoothed doublets on $z=0$ can be replaced by corresponding sheets of discrete sources and sinks, offset above and below $z=0$ by distance r_c. For the calculation of downwash, smoothing a doublet sheet is equivalent to splitting the doublet sheet into its constituent source and sink sheets displaced an appropriate amount. This approach leads to a simple way to construct smoothed local solutions using separation of variables. Since the local solutions can be interpreted basic flows obtained by separation of variables, as shown in Fig. 3, the smoothed local solutions used for matching can also be obtained in a similar manner. However these flows must be associated with source/sink surfaces, rather than doublet surfaces. Furthermore, the downwash associated with these surfaces must be evaluated at positions offset above and below the surfaces a distance r_c. This approach, which has been verified by other means (direct integration of the kernel function), has been found to be particularly straightforward and insightful.

The net local correction is the difference between the high-resolution local solution and the matching solution, both terms in the square brackets of Eqs. (43) and (44). For a two-dimensional flat surface with coordinate x, the net local correction $\mathscr{L}_w$, at the point $x_1=p$, in terms of upwash is given by

$$\begin{aligned}\mathscr{L}_w=\boldsymbol{n}\cdot\Bigg\{\frac{i}{4}\int_{-\infty}^{+\infty}\Bigg\{\Bigg(\mu_\omega^{2\mathrm{D}}+\frac{d\mu_\omega^{2\mathrm{D}}}{dx}(x-p)\\ +\frac{d^2\mu_\omega^{2\mathrm{D}}}{dx^2}\frac{(x-p)^2}{2!}+\cdots\mathrm{HOT}\cdots\Bigg)\\ \times\nabla\Bigg[\frac{\partial}{\partial n}(H_1^{(2)}(k\hat{r}))-\frac{\partial}{\partial n}(H_1^{(2)}(k\sqrt{\hat{r}^2+r_c^2}))\Bigg]\Bigg\}dS\Bigg\}\\ \approx-\frac{\mu_\omega^{2\mathrm{D}}}{2r_c}\Bigg(1-\frac{1}{2}(kr_c)^2\Bigg)+\frac{d^2\mu_\omega^{2\mathrm{D}}}{dx^2}\frac{r_c}{4}\end{aligned}\tag{45}$$

for $kr_c\ll 1$, where the truncated Taylor series has been substituted into Eq. (43), p is taken as the local origin, and $\mathscr{T}$ is taken as unity. It is interesting to note that only the even terms in x contribute to the integral.

It is documented in the literature[42,43] that for subsonic flow about edges there is a $\frac{1}{2}$-power singularity in the tangential velocity component just inboard of an edge. Therefore, at an edge, the dipole strength distribution can be expected to be expanded in powers of $\frac{1}{2}$,

$$\mu_{\mathrm{edge}}^{2\mathrm{D}}(x)=c_0x^{1/2}+c_1x^{3/2}+O(x^{5/2}),\tag{46}$$

where c_0 and c_1 are arbitrary constants.

In three dimensions, Eq. (46) can be thought of as the leading-order term in an asymptotic expansion, because very close to an edge, the two-dimensional terms dominate the three-dimensional terms in the solution. This can be demonstrated with as simple perturbation expansion about the edge. Trouble could possibly arise at the wing tip corners, where two edges are present. Here, however, the flow field can be modeled with conical flow solutions.[42,44] Numerical experimentation has shown that the conical flow solutions are not

required for sufficient accuracy. If, however, higher accuracy is needed, higher-order wing tip corner solutions can be constructed from conical flow solutions.

Therefore, the weighting function $\mathscr{F}$ in Eq. (37) should be $\sqrt{\Delta r}$ for the edge panels (unity elsewhere). As mentioned earlier, this is because the dipole strength distribution becomes nonanalytic at the edge. At any point l on the surface near the edge, c_0 and c_1 can be determined in terms of the derivative of Eq. (46), ignoring terms of $O(x^{3/2})$, such that,

$$c_0=\frac{3}{2}\frac{\mu_{\text{edge}}^{\text{2D}}(l)}{\sqrt{l}}-\sqrt{l}\frac{d}{dx}[\mu_{\text{edge}}^{\text{2D}}(x)]_{x=l},$$
$$c_1=\frac{1}{\sqrt{l}}\frac{d}{dx}[\mu_{\text{edge}}^{\text{2D}}(x)]_{x=l}-\frac{1}{2}\frac{\mu_{\text{edge}}^{\text{2D}}(l)}{l^{3/2}}. \tag{47}$$

The series expansion of the dipole distribution for an edge, Eq. (46), can now be written in terms of the dipole strength density and its derivative as

$$\mu_{\text{edge}}^{\text{2D}}(x)=\left(\frac{3}{2}\frac{\mu_{\text{edge}}^{\text{2D}}(l)}{\sqrt{l}}-\sqrt{l}\frac{d}{dx}[\mu_{\text{edge}}^{\text{2D}}(x)]_{x=l}\right)x^{1/2}+\left(\frac{1}{\sqrt{l}}\frac{d}{dx}[\mu_{\text{edge}}^{\text{2D}}(x)]_{x=l}-\frac{1}{2}\frac{\mu_{\text{edge}}^{\text{2D}}(l)}{l^{3/2}}\right)x^{3/2}, \tag{48}$$

where x is the coordinate direction inboard of the edge. In practice, the doublet strength and its derivative are determined by numerical means, so these terms are replaced by doublet strengths at nodal locations. After substitution of Eq. (48) into Eq. (44), and evaluation [similar to Eq. (45) either by direct integration or separation of variables), the corresponding correction in terms of upwash for edge panels is found as

$$\mathscr{L}_w^{\text{edge}}(x)=\frac{c_0}{2r_c}R_{r_c}{}^{1/2}\left\{\left(1-\frac{(kR_{r_c})^2}{6}\right)\cos\frac{\theta}{2}+\frac{(kR_{r_c})^2}{15}\cos\frac{5\theta}{2}\right\} \tag{49}$$

$$+\frac{c_1}{2r_c}R_{r_c}{}^{3/2}\left\{\left(1-\frac{(kR_{r_c})^2}{10}\right)\cos\frac{3\theta}{2}+\frac{(kR_{r_c})^2}{35}\cos\frac{7\theta}{2}\right\}, \tag{50}$$

for $kR_{r_c}<1$, where x is the coordinate inboard of the edge, $R_{r_c}=\sqrt{x^2+r_c{}^2}$, and $\theta=\arctan(r_c/x)$.

The extension to a curved surface is completely straight forward. The application to curved surfaces can proceed as follows. The smoothed doublets are now positioned along a curved surface, leading to some straightforward changes in geometry and surface velocity calculations. To good accuracy the flat surface local and matching solutions can still be used, because the curvature effects are higher order in a local region, i.e., the region is ''locally flat,''[27–29] which involves a different utilization of ANM, which shows the role of curvature in local solutions to be surprisingly unimportant. Nevertheless, curvature can be included in local solutions by a completely straightforward perturbation procedure.

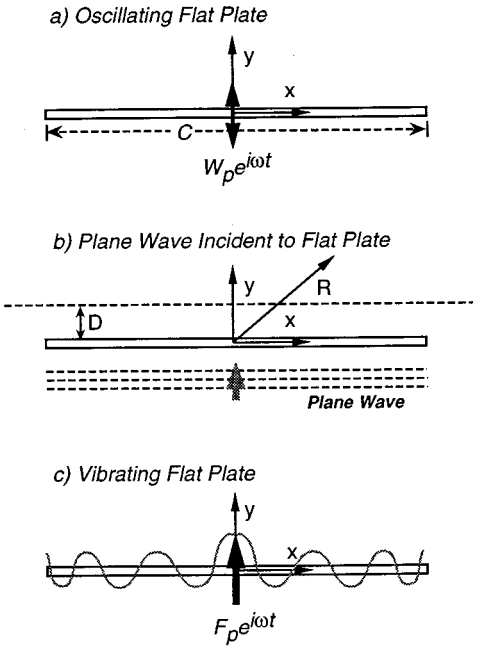


FIG. 5. Schematic diagram of the various plate geometric configurations. (a) Oscillating flat plate, (b) incident normal wave, (c) vibrating elastic flat plate.

III. NUMERICAL RESULTS

The method was implemented in two dimensions to calculate the surface pressure, and the acoustic field of an unbaffled oscillating flat plate, see Fig. 5. Later, the implementation was generalized to include calculations of acoustic scattering from a flat plate, and the acoustic field of a lightly loaded elastic plate subject to forced vibrations. The presented results include both surface pressure and far-field radiation calculations. Comparison is made with classical solutions when possible.

In previous work, the ANM methodology has been applied to both steady linear incompressible and compressible aerodynamics.[32,45] In both applications, the method has been rigorously verified by comparison to classical solutions. Throughout this developmental work, the ANM boundary element formulation proved to be a robust and versatile methodology providing accurate results.

Consider a rigid, flat plate undergoing forced oscillations in two dimensions. For the incompressible case, an exact analytical expression exists for pressure jump across the plate.[46] Figure 6 shows a calculation of the surface pressure jump for a flat plate oscillating in an incompressible flow. The calculation used 100 smoothed dipoles along the chord of the plate, with a smoothing length scale r_c of 1.6 times the dipole spacing, d_s. The jump in pressure is normalized by the density ρ, the semichord of the plate b, the frequency of oscillation ω, and the plate surface velocity W_p. The coordinate along the plate is normalized by the plate chord C. The calculation was done with increasingly

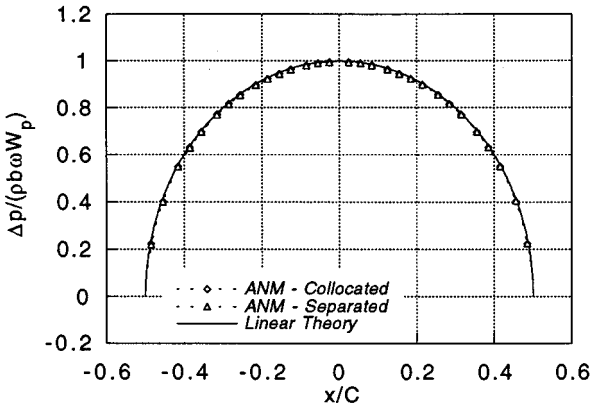


FIG. 6. Pressure jump across the surface of a two-dimensional oscillating plate in incompressible flow, $r_c/d_s=1.6$.

fewer dipoles, and was shown to successfully converge with as few as ten smoothed dipoles. The ANM result compares favorably to the exact analytical solution to the governing equations. Two cases are shown in Fig. 6. In the first case, the nodes and smoothed dipole locations were collocated; indicated in the legend of the figure as collocated. In the second case, the nodes and smoothed dipole locations are separated by $d_s/4$, one quarter the spacing between consecutive smoothed doublets, indicated in the figure as separated. It is clear from the figure that the results are independent of node/influence point positioning.

Next consider a flat rigid plate oscillating, with dimensionless frequency $kC=10$, in compressible flow. Figure 7 shows the calculation of the magnitude of the pressure along the surface of the plate. The calculation used 100 smoothed dipoles along the chord of the plate, with a smoothing length scale, $r_c/d_s=1.7$. The pressure is normalized by ρ, the speed of sound a_∞, and the plate surface velocity. Similar to the previous result, the converged calculation can be done with as few as ten smoothed dipoles, and demonstrate the characteristic ANM control point/influence point independence. The acoustic wavelength is approximately $3C/5$, and therefore the plate is not acoustically compact. This is evident by the shape of the pressure distribution. Again, two cases are shown in the figure. In both cases, the node and smoothed

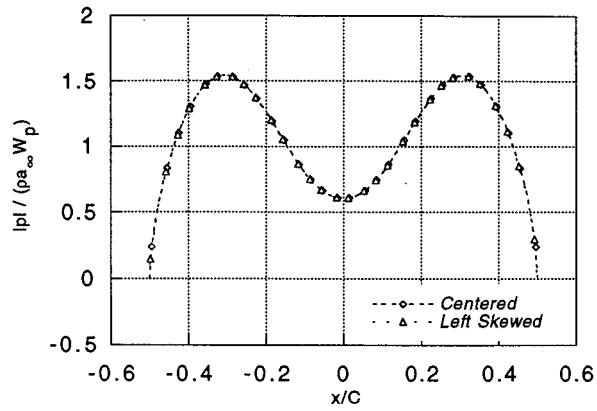
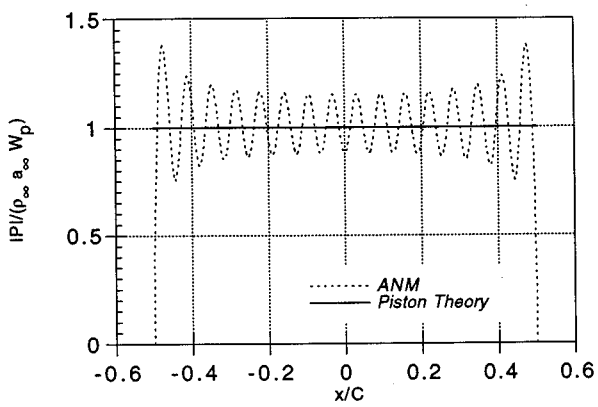


FIG. 7. Surface pressure for a two-dimensional oscillating plate in compressible flow, $kC=10$, $r_c/d_s=1.6$.



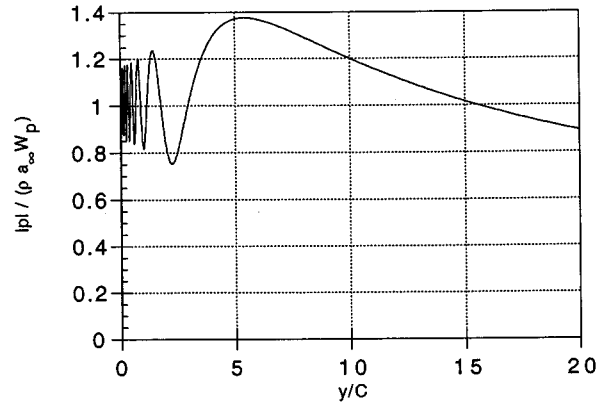


FIG. 8. Calculation for a two-dimensional oscillating plate in compressible flow, $kC=100$, $r_c/d_s=1.7$: Surface pressure (top); Pressure along axis of symmetry (bottom).

dipole locations are collocated; however, in one case they distributed evenly (symmetrically) over the plate, and in the other case, they are all skewed to the left, as indicated in the figure. Similar to the preceding calculation, it is clear that the results are independent of node/influence point positioning, even in the extreme of asymmetric skewing.

Examining a higher frequency, $kC=100$, one expects the results to approach that of piston theory,[47] which exist in the limit as $k\rightarrow\infty$. As shown in Fig. 8, the ANM calculation approaches the anticipated piston theory limit. Note the more complex structure of the surface pressure distribution. This is primarily due constructive and destructive interference of the surface pressure by waves radiated from the plate edges. The acoustic wavelength is approximately $C/16$. Due to the elevated frequencies involved in the problem, 500 smoothed dipoles were used to assure highest accuracy. However, the calculation could have been done with far fewer smoothed dipoles. The smoothing length scale used was $r_c/d_s=1.7$. Looking at the radiated acoustic field, Fig. 8 shows the magnitude of the pressure along the axis normal to the center of the plate (axis of symmetry). Here we see the complex nature of the acoustic near field, and the inverse square root fall off of the pressure in the acoustic far field, characteristic of two-dimensional acoustics.

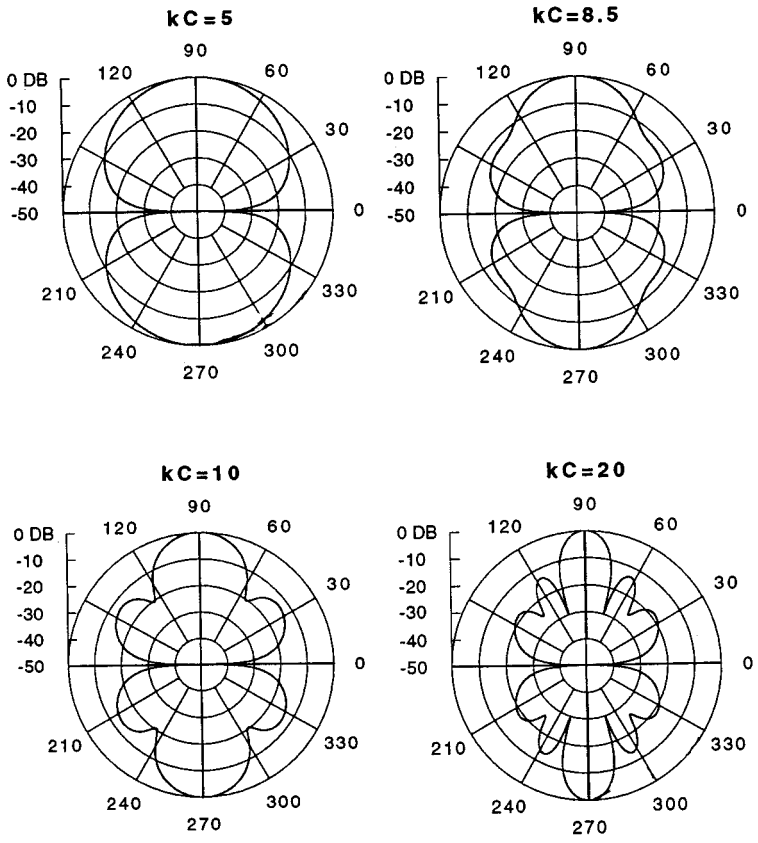


FIG. 9. Directivity patterns for a flat plate oscillating at various frequencies, $kC=5$, $kC=8.5$, $kC=10$, $kC=20$, $r_c/d_s=1.7$.

Now consider the far-field radiation for an oscillating flat plate at different frequencies, shown in Fig. 9. The radial distance from the center of the plate to the observation position is 20 chord lengths. Starting with the directivity pattern in the top left of Fig. 9, $kC=5$, the radiator at this frequency is acoustically compact, and the pattern looks like that of a simple dipole. As the frequency increases, the directivity pattern becomes increasingly more complex. In the $kC=8.5$ pattern, top right of the figure, the formation of two side radiation lobes can be easily seen. The radiator is becoming increasingly less compact, and the influence of the plate edge radiation is becoming important. At $kC=10$, the trends continue, and finally at $kC=20$ some finer scale detail is visible in the directivity pattern. These calculations were made with 100 smoothed dipoles, and a smoothing length scale $r_c/d_s=1.7$.

The smoothing length scale, r_c, is the parameter that controls the magnitude of the smoothing effect. It is desirable to make the global solution as smooth as possible, thereby removing any large spatial gradients, and improving numerical convergence. This in turn ultimately reduces the dependence on node location. If too much smoothing is used, however, the global and local solutions do not match properly, and the calculation will degrade. This is primarily because the smoothing length scale becomes so large it introduces a length scale in the local problem that is comparable to the global length scales. Furthermore, if the smoothing radius is chosen too small, the global solution is no longer smooth, resulting in numerical discreteness errors. Again, the calculation will degrade due to the influence of large spatial gradients.

There is range of smoothing length scales r_c that provide the best global-local solution match, that fall in the mathematical ''overlap'' between the global and local solutions. The smoothing length scale should be on the order of the spacing between smoothed dipoles in the global solution. Typically values of 1.4 to 2.8 times the dipole spacing quite work well. Figure 10 shows the solution overlap for the first calculation presented in this paper, shown in Fig. 6. Incompressible calculations tend to be less forgiving than compressible calculations. In Fig. 10 the net dipole strength calculated using ANM is normalized by the net dipole strength calculated using linear theory, and plotted as a function of the smoothing length scale r_c normalized by the distance between smoothed dipoles in the solution l, and 100 smoothed dipoles were used in the calculation. The figure shows that the solution is insensitive to the smoothing parameter over a range of smoothing values, thereby demonstrating the expected overlap region between the global and local solutions.

As an interesting application of the methodology, the mechanical impedance of an oscillating flat plate in two dimensions was calculated. The calculation is somewhat similar to that found in many classical acoustics textbooks for the mechanical impedance of an oscillating baffled piston in three dimensions. Quite simply, the net force on the flat plate is calculated by direct integration of the surface pressure, and then the mechanical impedance is calculated by dividing by

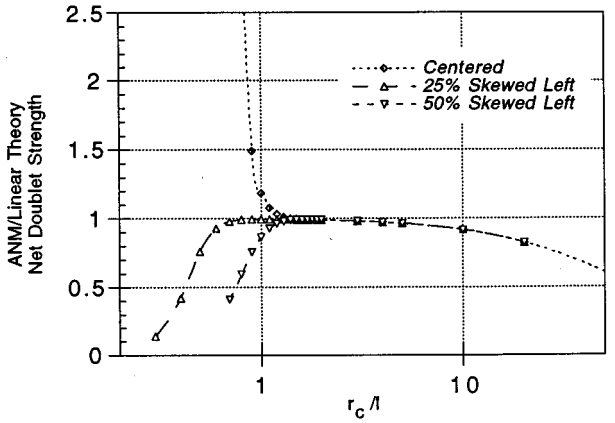


FIG. 10. Solution overlap region of a net dipole strength calculation for a rigid flat plate oscillating in incompressible flow.

the surface velocity, which is the same at all stations across the rigid plate. The impedance calculation is shown in Fig. 11, where at low frequency the apparent mass hydrodynamic loading is evident in the reactance $X(kC)$; the calculated slope agrees with low-frequency theoretical result. There is third power dipole radiation resistance evident in the low-frequency behavior of $R(kC)$. At higher frequencies, the impedance is dominated by the radiative resistance. The calculation was made with 100 smoothed dipoles, and a smoothing length scale $r_c/d_s=1.7$. The impedance is normalized by the density ρ, the speed of sound a_∞, the plate velocity W_p, and the plate chord $C=2b$.

Consider now the diffraction of an incident plane wave from a two-dimensional flat plate. As shown in Fig. 5, the flat plate is oriented such that the incident wave is parallel to the plate surface, normal to the bisecting axis of the plate. The diffracted pressure is calculated along the y_0 station, shown in Fig. 5, where $D=C$. The magnitude of the pressure along y_0 is shown in Fig. 12. The pressure is normalized by the incident value. Two hundred smoothed dipoles were used in the calculation, the frequency is $kC=40$, and the smoothing length scale is $r_c/d_s=1.7$. The shadow region behind the plate is clear, and the characteristic oscillations of the edge diffraction pattern are evident. Examining the directivity pattern, shown in Fig. 12, at a radial distance of C an increase

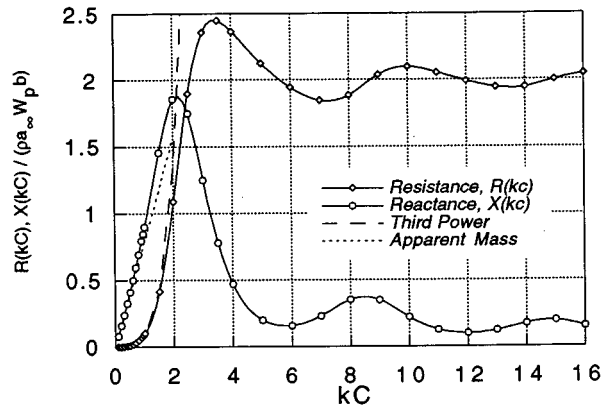
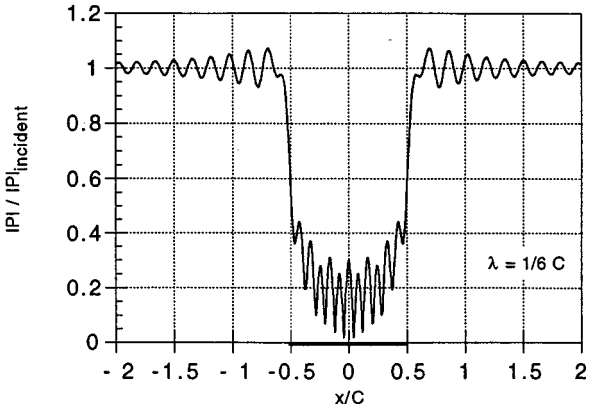


FIG. 11. The mechanical impedance for an oscillating flat plate in two dimensions, $r_c/d_s=1.7$.



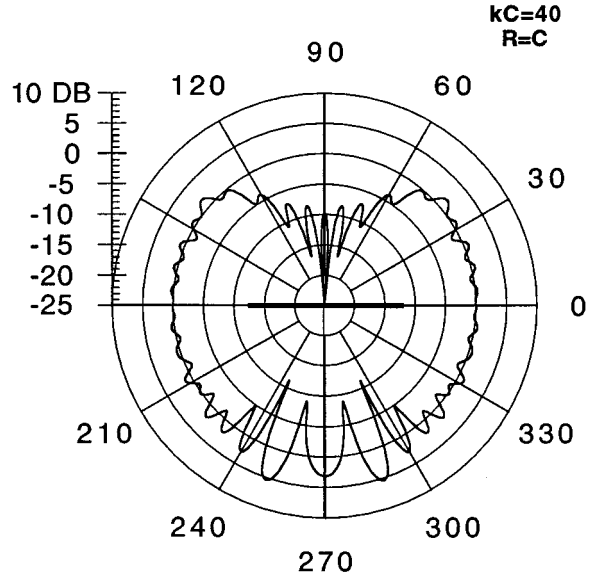


FIG. 12. Calculation of normal wave diffraction from a two-dimensional flat plate, $kC=40$, $r_c/d_s=1.7$: Pressure magnitude along axis parallel to plate (C above plate) from the diffraction of a normal incident wave (top); Directivity pattern from the diffraction of a normal incident wave, $R=C$ (bottom).

of pressure can be seen on the reflecting side of the plate, and a decrease can be seen in the shadow region. The smaller radial distance was chosen in the directivity pattern to show the shadow and defraction regions more dramatically. Toward the side of the plate, the ratio of the reflected pressure to the incident pressure is practically unity, with the exception of a small oscillatory perturbation in pressure magnitude.

Consider now the effect of forced vibrations on a two-dimensional lightly loaded elastic flat plate. The modes of vibration are governed by the transverse wave equation, and the boundary conditions are such that the ends are free, therefore no shear force or moment exist, and there is zero slope specified at the driving point. The elastic plate is driven by a harmonic point force, $F_p e^{iwt}$, at the mid-chord. Figure 5 shows a sketch of the geometry. For a specified frequency, the surface velocity of the plate can be written as a ratio of the surface velocity $w_p(x)$ to the velocity at the driving point

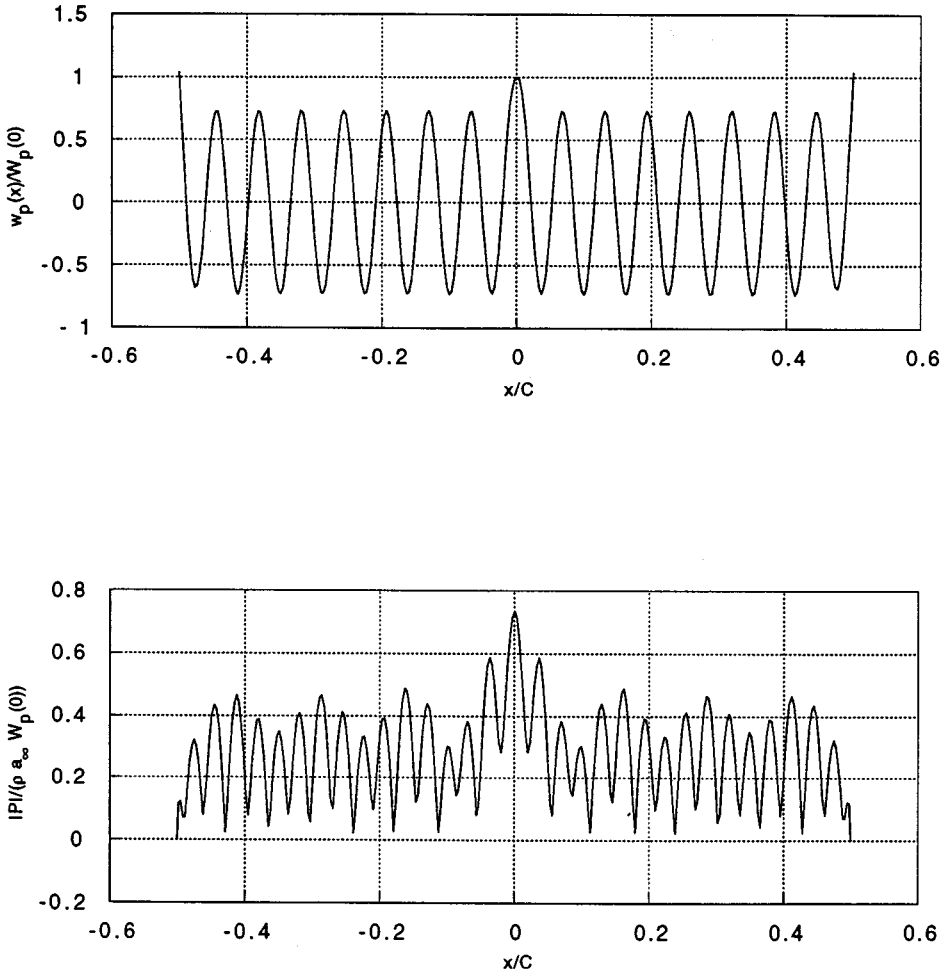


FIG. 13. Calculation for a lightly loaded elastic plate driven at the origin by a harmonic point force, $kC=50$: Surface velocity (top); Surface pressure distribution, $r_c/d_s=1.7$ (bottom).

$w_p(0)$. For a frequency of $kC=50$, the surface velocity is shown in Fig. 13. In this calculation, the flexural wave speed on the plate is subsonic, and the ratio of the plate wave speed to the acoustic wave speed is $a_p/a_\infty=0.5$. The surface pressure and acoustic response of the plate is calculated using the surface velocity as the input boundary condition. Using 300 smoothed dipoles, to assure highest resolution, and a smoothing length scale $r_c/d_s=1.8$, Fig. 13 shows the magnitude of the surface pressure distribution. The surface pressure is normalized by the density, speed of sound, and the velocity at the plate origin. Note, fewer dipoles could have been used without a compromise in accuracy. The wavelength of the surface pressure is approximately $C/8$. The directivity pattern is shown in Fig. 14, at a radial distance of 20 chord lengths from the center of the plate. The detailed structure of the radiated field is evident, and the node present along the plate axis can be clearly identified.

IV. CONCLUDING REMARKS

A boundary element method based on smoothed fundamental solutions to the Helmholtz equation has been developed and applied successfully to linear acoustics problems. The formulation is based on a technique called analytical/numerical matching, which offers a fresh point of view on traditional boundary element techniques, by solving the problem with a fundamentally different means. The ANM approach breaks the problem up into global and local con-

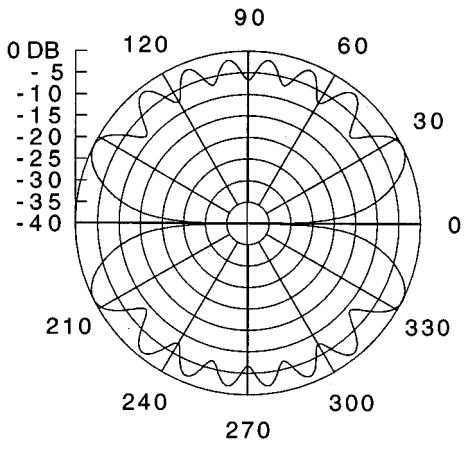


FIG. 14. Directivity pattern radiated by lightly loaded elastic plate driven at the origin by a harmonic point force, $kC=50\ r_c/d_s=1.7R=20C$.

stituent problems that are defined by the physical length scales of the problem and an additional smoothing length scale.

ANM offers a computational acoustic methodology without the traditional problems associated with the numerical solution of singular integral equations. The ANM boundary element formulation leads to accurate solutions, rapid numerical convergence, and is free from the ambiguity present in many traditional panel methods. Additionally, ANM does not have rigid rules on influence point/control point location, and therefore is suitable for complex geometric configurations, including hybrid fluid/structure calculations requiring a matching computational mesh.

In practice, ANM is a simple methodology, that offers high-resolution acoustic calculations in the framework of a unified acoustic method. In the future, one can envision the method being applied to a multitude of acoustic problems, including the design and analysis of complex acoustic configurations.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research.

[1] M. Munk, Tech. Rep. No. TR 184, NACA (unpublished).
[2] T. von Karman, Tech. Rep. No. TM 574, NACA (unpublished).
[3] J. L. Hess and A. Smith, *Calculation of Potential Flow About Arbitrary Bodies, Progress in Aeronautical Sciences* (Pergamon, New York, 1962).
[4] J. Hess and A. Smith, J. Ship Res. **8**, 22 (1964).
[5] M. Friedman and R. Shaw, J. Appl. Mech. **29**, 40 (1962).
[6] R. Shaw and M. Friedman, in *Proc. 4th U.S. National Congress of Applied Mechanics* (ASME, New York, 1962).
[7] R. Banaugh and W. Goldsmith, J. Acoust. Soc. Am. **35**, 1590 (1963).
[8] F. Farassatt, Tech. Rep. No. AIAA Paper 86-1877, AIAA (unpublished).
[9] D. Colton and R. Kress, *Integral Equation Methods in Scattering Theory* (Wiley–Interscience, New York, 1983).
[10] P. Filippi, *Theoretical Acoustics and Numerical Techniques, C.I.S.M. Courses and Lectures Number 277* (Springer-Verlag, Berlin, 1983).
[11] T. W. Wu, J. Acoust. Soc. Am. **97**, 84–91 (1995).
[12] T. W. Wu, J. Acoust. Soc. Am. **96**, 3733–3743 (1994).
[13] T. W. Wu and L. Lee, J. Sound Vib. **175**, 51 (1994).
[14] J. P. Coyette and J. L. Migeot, in *Proceedings of the International Congress on Noise Control Engineering—Inter-Noise 93*, Leuven, Belgium (Noise Control Foundation, Poughkeepsie, NY, 1993).
[15] J. P. Coyette and H. Wynendaele, in *Proceedings of the International Congress on Noise Control Engineering—Inter-Noise 94*, Yokohama, Japan (Inst. Noise Control Eng./Acoust. Soc. Japan, 1994).
[16] J. P. Coyette, in *Proceedings of the 13th International Modal Analysis Conference*, Nashville, TN (International Society for Optical Engineering (SPIE), Bellingham, WA, 1995).
[17] A. F. Seybert, *Experimental Validation of Finite Element and Boundary Element Methods for Predicting Structural Vibration and Radiated Noise* (NASA, Washington, DC, 1994).
[18] A. F. Seybert and B. Soenarko, in *Proceedings of the 2nd International Symposium on Transport Noise—Transport Noise 94* (Interpublish, Ltd., St. Petersburg, Russia, 1994).
[19] A. F. Seybert, T. W. Wu, and G. Wan, in *Proceedings of the 2nd International Congress on Recent Developments in Air and Structure Borne Sound and Vibration* (Intl. Sci. Publ., Auburn, AL, 1992).
[20] H. Ashley, S. E. Widnall, and M. T. Landahl, AIAA J. **3**, 3–16 (1965).
[21] R. P. Shaw, in *Topics in Ocean Engineering II*, edited by C. Bretschneider (Gulf, Houston, TX, 1970), pp. 143–163.
[22] R. P. Shaw, in *Developments in Boundary Element Methods I*, edited by P. Banerjee and R. Butterfield (Applied Science, Essex, U.K., 1979), Chap. 6, pp. 121–153.
[23] R. P. Shaw, in *Boundary Element Methods in Mechanics*, Vol. 3 of *Computational Methods in Mechanics*, edited by D. Beskos (North-Holland, New York, 1987), Chap. 9, pp. 417–438.
[24] J. Hess, in *Innovative Numerical Analysis for the Engineering Sciences*, edited by R. E. A. Shaw (Univ. of Virginia, Charlottesville, VA, 1980), pp. 131–143.
[25] D. M. L. Habault, in *Proceedings of the 15th International Congress on Acoustics* (Acoust. Soc. Norway, Trondheim, Norway, 1995).
[26] W. Miller, Ph.D. thesis, Duke University, Durham, NC, 1990.
[27] T. R. Quackenbush and D. B. Bliss, Vertica **14**, 313 (1990).
[28] T. R. Quackenbush, D. B. Bliss, C.-M. Lam, and A. Katz, in *Proceedings of the 46th Annual Forum of the American Helicopter Society* (Am. Helicopter Soc., Washington, DC, 1990).
[29] T. R. Quackenbush, C.-M. Lam, and D. B. Bliss, J. Am. Helicopter Soc. **39**, 14 (1994).
[30] R. J. Epstein, Master's thesis, Duke University, Durham, NC, 1991.
[31] D. B. Bliss and W. O. Miller, J. Am. Helicopter Soc. **38**, 878 (1993).
[32] R. J. Epstein and D. B. Bliss, AIAA J. **33**, 894 (1995).
[33] R. C. Loftman, Master's thesis, Duke University, Durham, NC, 1994.
[34] O. D. Kellogg, *Foundations of Potential Theory* (Dover, Mineola, NY, 1953).
[35] B. Baker and E. Copson, *The Mathematical Theory of Huygens' Principle* (Chelsea, New York, 1987).
[36] H. Lamb, *Hydrodynamics* (Dover, New York, 1962).
[37] H. Ashley and M. Landahl, *Aerodynamics of Wings and Bodies* (Dover, Mineola, NY, 1965).
[38] *Vortex Methods, Lecture Notes in Mathematics*, edited by C. Anderson and C. Greengard (Springer-Verlag, New York, 1988), proceedings, Los Angeles 1987.
[39] P. Saffman, *Vortex Dynamics, Cambridge Monographs on Mechanics and Applied Mathematics* (Cambridge U. P., New York, 1992).
[40] J. T. Beale and A. Majda, J. Comput. Phys. **58**, 188 (1985).
[41] P. J. Morris, J. Acoust. Soc. Am. **97**, 2651(A) (1995).
[42] R. T. Jones and D. Cohen, *Aerodynamics of Wings at High Speeds*, Vol. VII of *Aerodynamics of Aircraft Components at High Speeds* (Princeton U. P., Princeton, NJ, 1957).
[43] M. Van Dyke, *Perturbation Methods in Fluid Mechanics* (Parabolic, Stanford, CA, 1975).
[44] R. T. Jones, *Wing Theory* (Princeton U. P., Princeton, NJ, 1990).
[45] D. B. Bliss and R. J. Epstein, in *Proceedings of the 13th AIAA Applied Aerodynamics Conference* (American Institute of Aeronautics and Astronautics, San Diego, CA, 1995).
[46] J. Katz and A. Plotkin, *Low Speed Aerodynamics, From Wing Theory to Panel Methods, McGraw-Hill Series in Aeronautical and Aerospace Engineering* (McGraw-Hill, New York, 1991).
[47] E. H. Dowell, *A Modern Course in Aeroelasticity* (Kluwer Academic, Dordrecht, The Netherlands, 1989).

A generalized network formulation for acoustic scattering

Paul J. Titterton, Jr.
System Technology Division, SRI International, 333 Ravenswood Avenue, Menlo Park, California 94025 and Space, Telecommunications, and Radioscience Laboratory, Stanford University, Stanford, California 94305

(Received 20 December 1995; revised 21 July 1996; accepted 13 August 1996)

The discretized Helmholtz and equivalence integral equations yield the impedance representation of a generalized network formulation for acoustic scattering. Trans- and self-impedances relate the ensonification, the scattered field, and the pressure and velocity on the scatterer's surface, **S**; the scatterer's dry impedance loads **S**. This formulation suggests applying analysis tools from electrical circuit theory: network element abstraction and change of field-variable basis. The scatterer's wet admittance (impedance) corresponds to the augmented driving point admittance (impedance) from electrical network theory. The lower-branch modes of the scattering from a submerged spherical shell occur when the reactance looking into the fluid from **S** negates that looking into the shell. Change of basis on **S** transforms the impedance representation into the admittance, arbitrary-reference, and network-scattering representations. The impedance, admittance, and arbitrary-reference representations naturally separate hard, soft, and intermediate background scattering, respectively. Scattering from passive objects with limited vibrational degrees of freedom is bounded: the lower-branch modes of the submerged spherical shell approach this bound. No bound exists when the vibrational degrees of freedom are unlimited. [S0001-4966(97)00401-3]

PACS numbers: 43.20.Rz, 43.20.Tb, 43.20.Wd, 43.40.At [JEG]

INTRODUCTION

This paper presents a generalized network formulation for acoustic scattering from penetrable structures. In this conceptually simple approach, the scattering structure is treated as a load that terminates an N-port network description of the fluid, where N is the number of mesh points on the wet surface. Our primary contribution is in organizing the scattering calculation to exploit analysis tools from circuit theory; in particular, change of field-variable basis is used to explore scattering from unknown-but-passive loads, and network element abstraction is used to interpret intermediate results and determine resonance frequencies and shapes.

Change of field-variable basis manipulates an existing network representation to generate any number of algebraically equivalent scattering formulations. The field variables which describe acoustic propagation are pressure and velocity. We change basis by forming a linear combination of pressure and velocity and rewriting the acoustic equations in terms of these new field variables. Of course, the utility of this basis change depends on the particular linear combination. We discuss three different scattering formulations, each obtained from the network-impedance representation, which is derived first. Each provides a different view—in the sense of a different network analogue—of the target's scattering mechanism, which in turn provides different modeling and measurement possibilities. First, a simple basis change yields the network-admittance representation, which can be used to extend an impedance-based result previously obtained by Borgiotti.[1] Impedance and admittance formulations naturally separate ''hard-background'' and ''soft-background'' scattering, respectively. Second, a more complicated basis change separates the total acoustic scattering into that from an arbitrary-impedance structure and a residual, such as is sometimes used in electromagnetics. Third, a change to the network-scattering basis is used to bound the acoustic scattering from passive targets with limited vibrational degrees of freedom—such as targets that are represented by finite-element models or truncated modal expansions. As an example, we show that a spherical shell's lower-branch resonances[2] approach this bound. Removing the assumption of limited vibrational degrees of freedom, we show that knowledge of a target's shape and passivity is insufficient to bound its normalized scattering cross section.

Network element abstraction enables separate examination and modeling of individual network elements—that is, the fluid and scattering object can be analyzed and modeled separately. This is known as substructuring to the finite-element modeling community, where it is widely used to reduce computational effort and to integrate data from disparate models. For example, Vasudevan and McCormick[3] use substructuring to integrate lumped-element models of internal structure into a finite-element model of the nominal scatterer.

Network element abstraction also allows us to examine network properties as seen from junctions, which can provide simple interpretation of physically meaningful intermediate results. The network-impedance representation of acoustic scattering yields the wet admittance—the submerged target's surface velocity response to a source which acts as a surface pressure discontinuity—which was originally described by Borgiotti.[1] The admittance representation yields the wet impedance—the submerged target's pressure response to a source which acts as a velocity discontinuity—which was previously formulated by Gaumond and Yoder,[4]

but left uninterpreted. The wet admittance and wet impedance are directly analogous to Kuh and Rohrer's[5] augmented driving point admittance and augmented driving point impedance, which occur in electrical network theory.

Additionally, network element abstraction allows us to analyze interaction resonances by comparing the fluid and structure impedance (or admittance) matrices. Scattered-field resonance interpretation, based on impedance comparison, is found in Junger and Feit[2] for submerged spherical shells and in Muzychenko and Rybak[6] for submerged circular cylindrical shells. In these cases, where the interface between the fluid and structure is a constant coordinate surface in a separable coordinate system, and where the fluid and structure are similarly symmetric, the field-variable basis functions can be chosen to enable mode-by-mode impedance comparison. For exposition, we analyze the scattering from a spherical shell and determine the locations of the lower-branch resonances. Resolving these sharp resonances requires dense sampling in frequency, and we save significant computational effort by decreasing the sampling density away from the resonances. Generalization to arbitrary shapes is straightforward given the appropriate computational tools.

The network formulation for acoustic scattering closely resembles Harrington's[7] ''theory of loaded scatterers,'' which describes the electromagnetic scattering from a structure with a number of attached antennas, where the antenna ports are terminated electrically. Harrington's description includes network parameters relating the transmitter- and antenna-port currents to the receiver- and antenna-port voltages. An electrical circuit loads the antenna ports, and a simple coupled system of equations yields the receiver voltage. Analogously, this chapter's description of acoustic scattering includes network parameters relating the source volume flux and the target surface velocity to the scattered pressure and target surface pressure. The scattering structure loads the surface ports, and a similar coupled system of equations yields the scattered pressure.

The fluid impedance matrices can be calculated using either domain methods[8] or boundary integral methods. Particular boundary integral formulations include variational methods,[9,10] superposition methods,[11–13] and direct discretization methods.[14–16] We will use the last—and most popular—of these to calculate the fluid impedance matrices. Numerical difficulties are encountered in cases where the scattering structure has sharp or flat appendages,[17] but aside from this ''thin-shape breakdown,'' our formulation is not constrained by the scatterer's geometry. We will not address the nonuniqueness associated with the surface Helmholtz integral equation.[18] Any of the techniques for treating the critical frequencies[12,14,19] can be substituted at appropriate locations in the equations presented here.

This paper is the second in a series which investigates single-frequency circuit methods for acoustic radiation and scattering. The first[20] analyzed the scattering from a submerged spherical shell that has an internal spring-mass assembly. The internal structure was treated as a load on a network description of the submerged shell. Two analyses were performed: the first used network-element abstraction at the interface between the submerged shell and the internal structure to identify peaks in the scattered response; the second used network-scattering variables and a power flow constraint to bound the scattering from a spherical shell having *any* point-attached passive internal structure. This paper presents a formulation in which the entire scattering structure is treated as a load on a network description of the fluid. This network formulation facilitates communication between electrical engineers and acousticians, and the attendant analysis tools enable simple derivation of unique results. Subsequent papers will investigate the synthesis of passive loads to achieve a desired dynamic response.

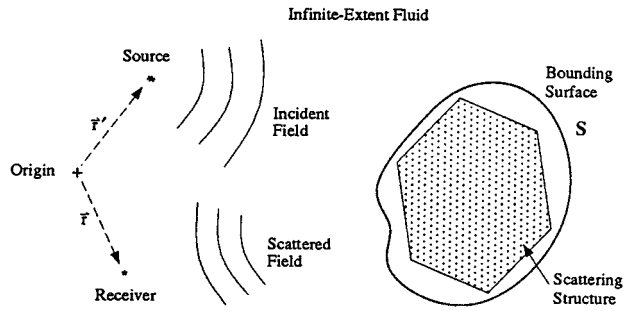


FIG. 1. Scattering configuration.

The remainder of this paper is divided into three parts: The first states the impedance representation of acoustic scattering from a general penetrable target, and the second and third apply change of field-variable basis and network element abstraction, respectively, to analyze and interpret that scattering.

I. THE IMPEDANCE REPRESENTATION FOR ACOUSTIC SCATTERING

A general scattering geometry is shown in Fig. 1: the scattering structure, shown as a polygon, is embedded in an infinite homogeneous fluid and surrounded by $\mathbf{S}$, a bounding surface that need not coincide with the surface of the structure; the field quantities associated with the interface between the fluid and scatterer are defined on $\mathbf{S}$; the scatterer is ensonified by a single source at $\mathbf{r}'$; and the scattered field is calculated at the receiver location, $\mathbf{r}$. The locations $\mathbf{r}'$ and $\mathbf{r}$ can vary over all space, external to $\mathbf{S}$. For convenience, we assume a point-monopole volume-flux source and receiver. We assume that the fluid is stationary, homogeneous, isotropic, inviscid, and isentropic, and that the microscopic, linearized equations of motion are valid external to $\mathbf{S}$. All field variations will be time harmonic with time dependence e^{iwt}.

We consider scattering from penetrable objects and refer to everything inside $\mathbf{S}$ as *the target, the scatterer, the scattering object, or the scattering structure*, even though $\mathbf{S}$ may not be coincident with the surface of the penetrable object. We assume that our scatterer can be described by a dry impedance ($\mathbf{Z}_d$) or dry admittance matrix ($\mathbf{Y}_d$) relating discretized normal velocity on $\mathbf{S}$ to discretized pressure on $\mathbf{S}$. The discretization is discussed in Appendix A. The dry impedance and admittance matrices represent the constraints imposed by the scattering structure on the relationship between the steady-state $\mathbf{p}$ and $\mathbf{v}$. Describing these matrices as

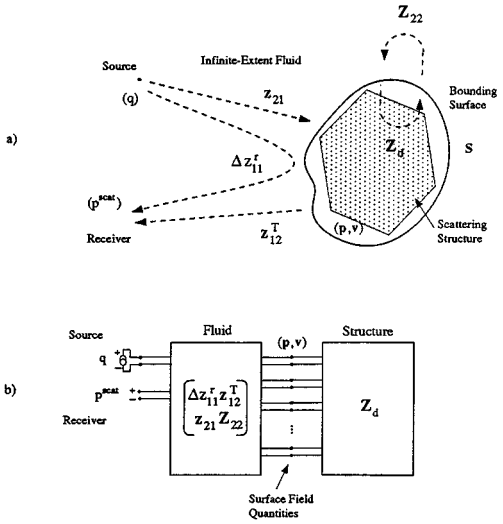


FIG. 2. Scattering configuration versus network-impedance block diagram.

dry, which follows previously published work,[1,4] is somewhat misleading. These matrices relate the **p** and **v** on **S**, regardless of conditions outside **S**.

The following coupled system of equations is the impedance representation for our target's acoustic scattering:

$$\begin{bmatrix} p^{\text{scat}} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \Delta z_{11}^r & \mathbf{z}_{12}^T \\ \mathbf{z}_{21} & \mathbf{Z}_{22} \end{bmatrix} \begin{bmatrix} q \\ \mathbf{v} \end{bmatrix}, \quad \mathbf{p} = -\mathbf{Z}_d \mathbf{v}, \tag{1}$$

where the 2×2-block matrix describes the dynamics of the fluid, **p** is the discretized total pressure on **S**, **v** is the discretized total normal velocity on **S**, q is the source strength, and p^{scat} is the scattered pressure—the difference between the total pressure at the receiving point and the pressure that would exist in the absence of the scatterer. The fluid impedance matrices are determined from first principles in Appendix A.

Figure 2(a) and (b) show the relationship between the physical scattering geometry and the network-impedance block diagram. The matrix elements have the following interpretation: the scalar Δz_{11}^r is the scattered pressure taking $q=1$ and $\mathbf{v}=0$ (i.e., Δz_{11}^r is the unit-source-strength scattering from a rigid **S**); the column vector $\mathbf{z}_{21}$ is the discretized surface pressure taking $q=1$ and $\mathbf{v}=0$ (i.e., $\mathbf{z}_{21}$ is the unit-source-strength pressure for a rigid **S**); the ith element of row vector $\mathbf{z}_{12}^T$ is the radiated field taking $q=0$ and $\mathbf{v}=\mathbf{e}_i$, where $\mathbf{e}_i$ is a vector of zeros with a one in the ith place (i.e., the elements of $\mathbf{z}_{12}^T$ are the radiated pressure for a rigid **S** with a single unit-velocity source); and finally, the ith column of $\mathbf{Z}_{22}$ is the discretized surface pressure taking $q=0$ and $\mathbf{v}=\mathbf{e}_i$ (i.e., the columns of $\mathbf{Z}_{22}$ are the surface pressure for a rigid **S** with a single unit-velocity source).

The matrix $\mathbf{Z}_{22}$ is the well-known Helmholtz impedance;[21] the other elements in the matrix which describes the fluid are used to fill out the linear relationship between field variables. The r superscript reminds us that Δz_{11}^r is the scattering from a rigid **S**. The Δ prefix reminds us that Δz_{11}^r does not relate total field quantities and cannot support a load—that is, the ratio of p^{scat} and q cannot be constrained by a physically meaningful impedance.

As stated in the introduction, the matrix description of the acoustic transmission paths in Eq. (1) is similar to the matrix description of the electromagnetic transmission paths in Harrington's[7] theory of loaded scatterers. In Appendix A, we derive Eq. (1) directly from the discretized Helmholtz integral equation and the discretized equivalence integral. Alternatively, we could follow Harrington's derivation, which emphasizes the relationship of Eq. (1) to a network formulation in which all the field variables are measured at ports (which can support loads). We start by placing spherical wave generators at $\mathbf{r}'$ and $\mathbf{r}$, the source location and the receiver location, respectively. Both generators are (i) independent sources, (ii) small compared to a wavelength, (iii) provide the same field as the point-monopole volume-flux sources of Appendix A, and (iv) do not impede the radiating field. We consider a network model for the relationship between the source-location generator, the receiver-location generator, and **S**. The port variables at the generators are total pressure and volume flux, and the port variables at **S** are **p** and **v**. In the velocity-current, force-voltage analogy, the impedance matrix for this network takes the current-like field quantities (**v** the volume flux at the source-location generator, q; and the volume flux at the receiver-location generator) to the voltage-like field quantities (**p**; the total pressure at the receiver location; and the total pressure at the source location). Because we do not consider the effect of a source at the receiver location, we can remove the associated column of the impedance matrix, and because we do not consider the radiated (or scattered) field at the source location, we can remove the associated row. With this column and row removed, loading the network with the dry impedance yields the relationship between the source strength, q, and the *total* pressure at the receiver. The scattered pressure is, of course, the difference between this total pressure and the incident field. (The incident field is the pressure that would exist with the scatterer replaced by an equal volume of the external fluid.) This incident-field term is incorporated into the impedance matrix to yield Eq. (1). With reference to Fig. 2, terminating **S** by the dry impedance yields the relationship between the source strength, q, and the *scattered* pressure at the receiver, p^{scat}.

Equation (1) is simply modified to treat multiple illumination sources and multiple receiving locations. If there is more than one source, we organize their complex amplitudes into a vector **q**, and we add the corresponding number of columns to Δz_{11}^r and $\mathbf{z}_{21}$. If there is more than one receiving location, we organize the scattered pressures into a vector $\mathbf{p}^{\text{scat}}$, and we add the corresponding number of rows to Δz_{11}^r and $\mathbf{z}_{12}^T$. Of course, if there is more than one illumination source and more than one receiving location, we assemble a vector **q** and a vector $\mathbf{p}^{\text{scat}}$, and we add the corresponding number of columns to $\mathbf{z}_{21}$, the corresponding number of rows to $\mathbf{z}_{12}^T$, and the corresponding number of columns and rows to Δz_{11}^r.

Direct substitution in Eq. (1) yields p^{scat}, **v**, and **p** in terms of q:

$$p^{\text{scat}} = \{\Delta z_{11}^r - \mathbf{z}_{12}^T [\mathbf{Z}_{22} + \mathbf{Z}_d]^{-1} \mathbf{z}_{21}\} q, \tag{2}$$

$$\mathbf{v} = \{-[\mathbf{Z}_{22} + \mathbf{Z}_d]^{-1} \mathbf{z}_{21}\} q, \tag{3}$$

$$\mathbf{p}=\{\mathbf{Z}_d[\mathbf{Z}_{22}+\mathbf{Z}_d]^{-1}\mathbf{z}_{21}\}q. \tag{4}$$

The equation for p^{scat} has a simple relation to resonant scattering theory:[22] $\Delta z_{11}^r q$ is the scattering from a rigid **S**, the hard background, and the second term is the residual due to the compliance of the structure. Additionally, if we write out $\Delta z_{11}^r=\mathbf{z}_{12}^T\mathbf{Y}_0\mathbf{z}_{21}$, which defines $\mathbf{Y}_0$, we find $p^{\text{scat}}=\{z_{12}^T[\mathbf{Y}_0-(\mathbf{Z}_{22}+\mathbf{Z}_d)^{-1}]\mathbf{z}_{21}\}q$, which is Borgiotti's[1] scattering formulation.

II. CHANGE OF FIELD-VARIABLE BASIS

Change of field-variable basis manipulates an existing network formulation by forming a linear combination of field variables and then rewriting the descriptive equations in terms of that linear combination. We define a transformation from the new surface-field variables (**a**,**b**) to the current surface-field variables (**p**,**v**):

$$\begin{bmatrix}\mathbf{p}\\ \mathbf{v}\end{bmatrix}=\begin{bmatrix}\mathbf{T}_{11} & \mathbf{T}_{12}\\ \mathbf{T}_{21} & \mathbf{T}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix}. \tag{5}$$

To rewrite the descriptive equations, we substitute Eq. (5) into Eq. (1) and replace **p** with **a** and **v** with **b** which yields

$$\begin{bmatrix}p^{\text{scat}}\\ \mathbf{a}\end{bmatrix}=\begin{bmatrix}\Delta\tilde{z}_{11} & \tilde{\mathbf{z}}_{12}^T\\ \tilde{\mathbf{z}}_{21} & \tilde{\mathbf{Z}}_{22}\end{bmatrix}\begin{bmatrix}q\\ \mathbf{b}\end{bmatrix},\quad \mathbf{a}=\tilde{\mathbf{Z}}_d\mathbf{b}, \tag{6}$$

where

$$\Delta\tilde{z}_{11}=\Delta z_{11}^r+\mathbf{z}_{12}^T\mathbf{T}_{21}[\mathbf{T}_{11}-\mathbf{Z}_{22}\mathbf{T}_{21}]^{-1}\mathbf{z}_{21}, \tag{7}$$

$$\tilde{\mathbf{z}}_{12}^T=\mathbf{z}_{12}^T\mathbf{T}_{21}[\mathbf{T}_{11}-\mathbf{Z}_{22}\mathbf{T}_{21}]^{-1}[\mathbf{Z}_{22}\mathbf{T}_{22}-\mathbf{T}_{12}]+\mathbf{z}_{12}^T\mathbf{T}_{22}, \tag{8}$$

$$\tilde{\mathbf{z}}_{21}=[\mathbf{T}_{11}-\mathbf{Z}_{22}\mathbf{T}_{21}]^{-1}\mathbf{z}_{21}, \tag{9}$$

$$\tilde{\mathbf{Z}}_{22}=[\mathbf{T}_{11}-\mathbf{Z}_{22}\mathbf{T}_{21}]^{-1}[\mathbf{Z}_{22}\mathbf{T}_{22}-\mathbf{T}_{12}], \tag{10}$$

$$\tilde{\mathbf{Z}}_d=-[\mathbf{T}_{11}+\mathbf{Z}_d\mathbf{T}_{21}]^{-1}[\mathbf{Z}_d\mathbf{T}_{22}+\mathbf{T}_{12}]. \tag{11}$$

Equation (6) shows the transformed fluid and structural impedances. Of course, we cannot admit all choices of $(\mathbf{T}_{11},\mathbf{T}_{12},\mathbf{T}_{21},\mathbf{T}_{22})$. For example, examination of Eq. (11) shows that the choice $\mathbf{T}_{11}=\mathbf{T}_{21}=0$ requires that we invert the zero matrix, as does the choice $\mathbf{T}_{11}=\mathbf{Z}_d$, $\mathbf{T}_{21}=-\mathbf{I}$. The transformations discussed in this section do not require that we invert the zero matrix.

We now examine three different bases (admittance, arbitrary reference, and network scattering) and give a brief description of the utility of each.

A. Network-admittance representation

To determine the admittance representation, we exchange **p** and **v** in Eq. (1) by choosing $(\mathbf{T}_{11},...,\mathbf{T}_{22})=(0,\mathbf{I},\mathbf{I},0)$, which yields

$$\begin{bmatrix}p^{\text{scat}}\\ \mathbf{v}\end{bmatrix}=\begin{bmatrix}\Delta z_{11}^{\text{pr}} & \mathbf{h}_{12}^T\\ \mathbf{h}_{21} & \mathbf{Y}_{22}\end{bmatrix}\begin{bmatrix}q\\ \mathbf{p}\end{bmatrix},\quad \mathbf{v}=-\mathbf{Y}_d\mathbf{p}, \tag{12}$$

where the fluid and structural dynamics matrices can be determined directly from Eqs. (7)–(11):

$$\Delta z_{11}^{\text{pr}}=\Delta z_{11}^r-\mathbf{z}_{12}^T\mathbf{Z}_{22}^{-1}\mathbf{z}_{21}, \tag{13}$$

$$\mathbf{h}_{12}^T=\mathbf{z}_{12}^T\mathbf{Z}_{22}^{-1}, \tag{14}$$

$$\mathbf{h}_{21}=-\mathbf{Z}_{22}^{-1}\mathbf{z}_{21}, \tag{15}$$

$$\mathbf{Y}_{22}=\mathbf{Z}_{22}^{-1}, \tag{16}$$

$$\mathbf{Y}_d=\mathbf{Z}_d^{-1}. \tag{17}$$

The fluid matrices can also be calculated directly from the discretized equivalence integral and surface Helmholtz integral equation using the same techniques as employed in Appendix A, in which case we arrive at simpler, but equivalent, forms for Eqs. (13)–(16).

In Eq. (12), the fluid is described by a hybrid matrix. We refer to (12) as the ''network-admittance'' representation because the fluid and structure—as seen from **S**—are described by the admittance matrices $\mathbf{Y}_{22}$ and $\mathbf{Y}_d$. This representation is useful when $\mathbf{Y}_d$ is available and $\mathbf{Z}_d$ is not—when the target has a rigid component, for example. The interpretation of the submatrices in (12) is similar to that of the submatrices in (1): $\Delta z_{11}^{\text{pr}}$ is the unit-source-strength scattering from a pressure-release **S**; $\mathbf{h}_{21}$ is the unit-source-strength velocity distribution for a pressure-release **S**; the elements of $\mathbf{h}_{12}^T$ are the radiated field for a pressure-release **S** with a single unit-pressure source; and the columns of $\mathbf{Y}_{22}$ are the surface velocity response for a pressure-release **S** with a single unit-pressure source.

Direct substitution solves for p^{scat}, **p**, and **v** in terms of q:

$$p^{\text{scat}}=\{\Delta z_{11}^{\text{pr}}-\mathbf{h}_{12}^T[\mathbf{Y}_{22}+\mathbf{Y}_d]^{-1}\mathbf{h}_{21}\}q, \tag{18}$$

$$\mathbf{p}=\{-[\mathbf{Y}_{22}+\mathbf{Y}_d]^{-1}\mathbf{h}_{21}\}q, \tag{19}$$

$$\mathbf{v}=\{\mathbf{Y}_d[\mathbf{Y}_{22}+\mathbf{Y}_d]^{-1}\mathbf{h}_{21}\}q. \tag{20}$$

Again, the equation for p^{scat} has a simple relation to resonant scattering theory:[22] $\Delta z_{11}^{\text{pr}} q$ is the scattering from a pressure-release **S**, the soft background, and the second term is the residual due to the stiffness of the structure. Additionally, if we write $\Delta z_{11}^{\text{pr}}=\mathbf{h}_{12}^T\mathbf{Z}_0\mathbf{h}_{21}$, which defines $\mathbf{Z}_0$, we find $p^{\text{scat}}=\{\mathbf{h}_{12}^T[\mathbf{Z}_0-(\mathbf{Y}_{22}+\mathbf{Y}_d)^{-1}]\mathbf{h}_{21}\}q$, which is similar in form to Borgiotti's scattering formulation, but emphasizes the fluid and structural admittances rather than fluid and structural impedances.

B. Arbitrary-reference representation

The impedance representation naturally separates the total scattering into the scattering from a rigid **S** and a residual due to the target's compliance. The admittance representation separates the total into that from a pressure-release **S** and a residual due to the target's stiffness. A rigid **S** has an infinite dry impedance, and a pressure-release **S** has a zero dry impedance. This section describes a change of basis which naturally separates the total scattering into that from an intermediate-impedance **S** and a residual due to the difference between the intermediate impedance and the actual impedance of the target. This intermediated impedance can be arbitrarily chosen and is the ''arbitrary reference'' of the section title. This section extends the separation methods of resonance scattering theory to targets that have arbitrary-impedance background (or reference) scattering.

The following transformation separates scattering from an arbitrary-reference impedance, $\mathbf{Z}_r$, from the total scatter-

ing: $(\mathbf{T}_{11},\mathbf{T}_{12},\mathbf{T}_{21},\mathbf{T}_{22})=(-\mathbf{Z}_r,\mathbf{T}_{12},\mathbf{I},\mathbf{I})$, where $\mathbf{T}_{12}$ is an arbitrary matrix with appropriate units. Choosing $\mathbf{a}$ to be the surface velocity distribution that would exist if the scatterer's impedance were $\mathbf{Z}_r$, and choosing $\mathbf{b}$ to be the difference between the actual surface velocity and $\mathbf{a}$, yields $\mathbf{T}_{12}=\mathbf{Z}_{22}$. Similar analysis is possible starting from the admittance formulation.

Solving for p^{scat}/q in the new basis yields

$$\frac{p^{\text{scat}}}{q}=\{\Delta z^r_{11}-\mathbf{z}^T_{12}[\mathbf{Z}_{22}+\mathbf{Z}_r]^{-1}\mathbf{z}_{21}\}+\mathbf{x}^T_1[\mathbf{Z}_d-\mathbf{Z}_r]\mathbf{x}_2, \tag{21}$$

where the bracketed term is the scattering from the intermediate impedance, $\mathbf{Z}_r$, the remaining term is the residual, $\mathbf{x}^T_1=\mathbf{z}^T_{12}[\mathbf{Z}_{22}+\mathbf{Z}_d]^{-1}$, and $\mathbf{x}_2=[\mathbf{Z}_{22}+\mathbf{Z}_r]^{-1}\mathbf{z}_{21}$. In order for the scattering from the intermediate impedance to equal the scattering from the actual impedance, we must have

$$\mathbf{x}^T_1[\mathbf{Z}_d-\mathbf{Z}_r]\mathbf{x}_2=0, \tag{22}$$

which is a much less stringent condition than requiring that the impedances be equal. With reference to Eq. (3), $\mathbf{x}_2$ is the surface velocity of a target that has dry impedance $\mathbf{Z}_r$, illuminated by a source at $\mathbf{r}'$. Regardless of receiver location, Eq. (22) is satisfied when the $\mathbf{Z}_d$ and $\mathbf{Z}_r$ respond to $\mathbf{x}_2$ with the same surface pressure. Given a single receiver location, $\mathbf{r}$, (22) is satisfied when $\mathbf{x}_1$ is orthogonal to the difference between those surface pressures.

A particularly interesting intermediate impedance is that of a fluid-filled $\mathbf{S}$. In this case, Eq. (22) yields exactly what is required of $\mathbf{Z}_d$ to achieve zero scattering.

C. Network-scattering representation

In this section, we assume knowledge of $\mathbf{S}$—the scatterer's shape and size—and knowledge of the scatterer's passivity, and we try find an *a priori* bound on the amplitude of the scattered field. Because passive targets cannot deliver power to the fluid, we expect some basic limitation on their ability to scatter the incident acoustic field. We provide three results: the first bounds the acoustic scattering for any passive target with *limited* vibrational degrees of freedom; the second shows that the lower-branch resonances of a spherical shell[2] approach this bound; and the third shows that if a passive target has an *unlimited* number of vibrational degrees of freedom, then the scattering cannot be bounded *a priori*.

The following transformation yields the network-scattering representation:[5]

$$(\mathbf{T}_{11},\mathbf{T}_{12},\mathbf{T}_{21},\mathbf{T}_{22})=(\mathbf{Z}^H_{22}\mathbf{R}^{-1/2},\mathbf{Z}_{22}\mathbf{R}^{-1/2},\ -\mathbf{R}^{-1/2},\mathbf{R}^{-1/2}),$$

where $\mathbf{R}=\frac{1}{2}(\mathbf{Z}^H_{22}+\mathbf{Z}_{22})$, and superscript H indicates conjugate transpose. The scattered field is

$$p^{\text{scat}}=(\Delta\tilde{z}_{11}+\tilde{\mathbf{z}}^T_{12}\boldsymbol{\Gamma}\tilde{\mathbf{z}}_{21})q, \tag{23}$$

where $\Delta\tilde{z}_{11}$ is the scattering from a target with $\mathbf{Z}_d=\mathbf{Z}^H_{22}$, and $\boldsymbol{\Gamma}=\tilde{\mathbf{Z}}_d^{-1}$. The scattering structure's passivity requires that the real part of its impedance matrix be positive semidefinite, i.e., $\text{Re}\{\mathbf{Z}_d\}\geqslant 0$. In the new basis, this condition is transformed into $\|\Gamma\|_2\leqslant 1$. This condition for passivity requires that we replace $\mathbf{v}$ in Eq. (1) with $\mathbf{A}^{\phi\phi}\mathbf{v}$, which yields power-conjugate surface variables. ($\mathbf{A}^{\phi\phi}$ is described in Appendix A.) We note that, if a network is passive, then $\text{Re}\{\mathbf{Z}_d\}\geqslant 0$. However, if we determine that $\text{Re}\{\mathbf{Z}_d\}\geqslant 0$ for some existing network, we cannot guarantee its passivity. That is, this matrix positivity condition is necessary, but not sufficient, for passivity. We also note that, if this condition is met at a particular frequency, we can always construct a passive network which has impedance $\mathbf{Z}_d$ at that frequency.

To bound the scattering from targets with limited vibrational degrees of freedom—our first result—we take the magnitude of Eq. (23) and apply Cauchy-Schwartz and the triangle inequality, which yield

$$|p^{\text{scat}}/q|\leqslant|\Delta\tilde{z}_{11}|+\|\tilde{\mathbf{z}}_{12}\|_2\|\tilde{\mathbf{z}}_{21}\|_2. \tag{24}$$

This bound applies to passive targets that have surface fields well described by $\mathbf{p}$ and $\mathbf{v}$.

The second result shows that this bound can be reached for backscattering of a plane wave from a spherical shell, in which the shell is restricted to a single mode of vibration. The fluid impedance submatrices for the sphere are calculated in Appendix B. Limiting vibration to the lth spherical harmonic and applying Eq. (24), we find

$$\left|\frac{p^{\text{scat}}}{q}\right|\leqslant\frac{(2l+1)}{k}\left\{\left|\frac{j'}{h'}-\frac{1}{2}\frac{|h'|^2}{h'^2}\right|+\frac{1}{2}\right\}, \tag{25}$$

where $h=h^{(2)}_l(ka)$ is the lth spherical Hankel function of the second kind evaluated at ka, $j=j_l(ka)$ is the lth spherical Bessel function evaluated at ka, k is the wave number, and a is the diameter of the spherical $\mathbf{S}$. From Junger and Feit[2] Eq. (11.30), at the lower-branch resonance, the spherical shell scattering magnitude is written

$$\left|\frac{p^{\text{scat}}}{q}\right|=\frac{(2l+1)}{k}\left|\frac{j'}{h'}-\frac{|h'|^2}{h'^2}\right|. \tag{26}$$

It is easy to show analytically that this magnitude is never greater than the bound; that this magnitude approaches the bound when $|j|\ll|h|$; and that this magnitude equals the bound for certain phase conditions immediately off resonance.

Finally, the third is an interesting theoretical result: We provide a constructive demonstration that, for any target shape and any constant, M, there exists a passive dry impedance which yields $|p^{\text{scat}}/q|>M$. We postulate a particular dry impedance for the target: a sphere is embedded within $\mathbf{S}$, and the remainder is filled with the external fluid. The first N harmonic modes of the embedded sphere's dry impedance are chosen to be the conjugate match to the impedance looking into the fluid from the surface of the sphere, i.e., $[\mathbf{Z}_d]^{ll}=[\mathbf{Z}_{22}]^{ll}*$. For mode numbers greater than N, we assume that the sphere is rigid. We make no attempt to realize a scatterer which achieves this dry impedance. For plane-wave backscattering, this configuration yields

$$\frac{p^{\text{scat}}}{q}=\frac{i}{k}\sum_{n=0}^{\infty}\left\{(2n+1)(-1)^n\left[\frac{j'_n}{h'_n}\right]\right\}-\frac{i}{k}\sum_{n=0}^{N}\left\{(2n+1)\times(-1)^n\left[\frac{j'^2_n+y'^2_n}{h'^2_n}\right]\right\}, \tag{27}$$

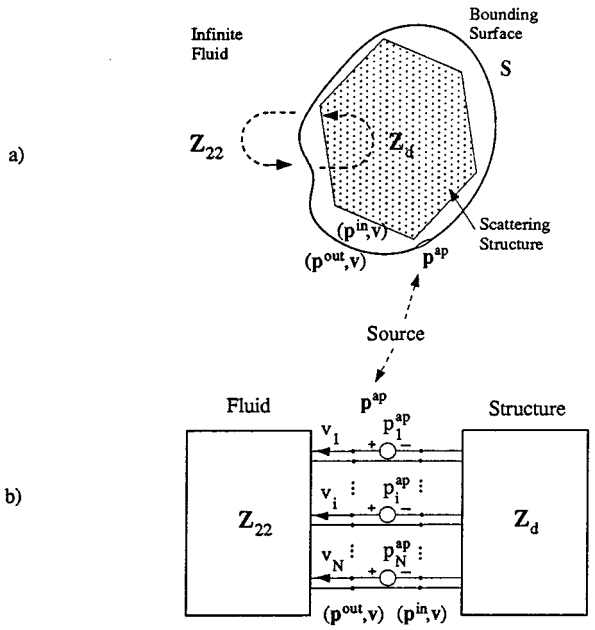


FIG. 3. Physical and network interpretations for the wet admittance.

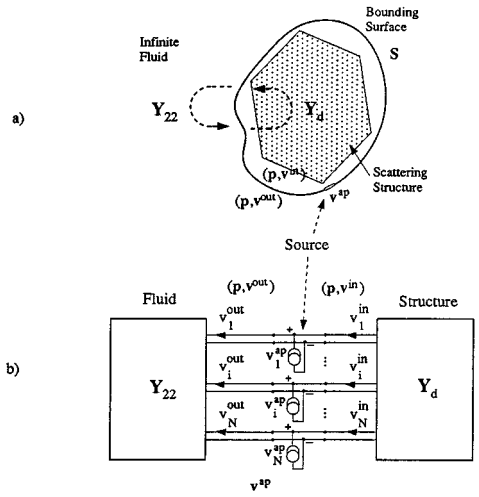


FIG. 4. Physical and network interpretations for the wet impedance.

where p^{scat} is the far-field pressure, referred to 1 m. The first term is the scattering from a rigid sphere, and the second is the contribution from the conjugate-matched modes. For $n \gg ka$ the second summand becomes $(2n+1)(-1)^n$. Thus as we increase N, p^{scat} grows without bound, which shows that knowledge of a three-dimensional target's shape and passivity, without restriction on the surface field's vibrational degrees of freedom, does not lead to an *a priori* bound on the scattering.

III. NETWORK ELEMENT ABSTRACTION

Network element abstraction has two main features: it permits us (i) to examine network properties as seen from junctions, and (ii) to examine each network element separately. In this section, we use the first feature to interpret the ''wet admittance,'' $\mathbf{Y}_w=[\mathbf{Z}_{22}+\mathbf{Z}_d]^{-1}$, and the ''wet impedance,'' $\mathbf{Z}_w=[\mathbf{Y}_{22}+\mathbf{Y}_d]^{-1}$, which occur frequently in Eqs. (2)–(4) and Eqs. (18)–(20), respectively. We use the second feature to examine resonances caused by the interaction of a spherical shell with the surrounding fluid.

A. The wet admittance and wet impedance

Figure 3(a) and (b) show physical and symbolic network interpretations of the wet admittance, which is obtained by breaking into the network at the $(\mathbf{p},\mathbf{v})$-port and considering the effect of a discontinuity in pressure as we follow the normal across $\mathbf{S}$. This discontinuity, which we call the applied pressure, $\mathbf{p}^{\text{ap}}$, is the difference between the pressure affecting the structure in the equation $\mathbf{p}=-\mathbf{Z}_d\mathbf{v}$ and the pressure affecting the fluid in $\mathbf{p}=\mathbf{Z}_{22}\mathbf{v}$. The pressure affecting the structure is $\mathbf{p}^{\text{in}}$ (for now), and the pressure affecting the fluid is $\mathbf{p}^{\text{out}}=\mathbf{p}^{\text{in}}+\mathbf{p}^{\text{ap}}$. We assume that the surface velocity is continuous as we follow the normal across $\mathbf{S}$, and we consider the applied pressure to be the only source. It is not clear how to physically realize $\mathbf{p}^{\text{ap}}$, although any force source applied to $\mathbf{S}$ could approximate its effect. Regardless of the realizability of the source, this section provides an interpretation of the quantity $[\mathbf{Z}_{22}+\mathbf{Z}_d]^{-1}$. In Fig. 3(a), the applied pressure is shown covering only a small disk. Of course, this can be generalized to any applied pressure. The equations

$$\mathbf{p}^{\text{out}}=\mathbf{Z}_{22}\mathbf{v}; \quad \mathbf{p}^{\text{in}}=-\mathbf{Z}_d\mathbf{v}; \quad \mathbf{p}^{\text{out}}=\mathbf{p}^{\text{in}}+\mathbf{p}^{\text{ap}} \tag{28}$$

combine to give $\mathbf{v}=\mathbf{Y}_w\mathbf{p}^{\text{ap}}$. The ith column of $\mathbf{Y}_w$ is the discretized normal velocity on $\mathbf{S}$ for $\mathbf{p}^{\text{ap}}=\mathbf{e}_i$.

The symbolic representation, a network analogy, shows two impedances in series with a vector of pressure sources. The wet admittance is the admittance seen by that source and is completely analogous to the ''augmented driving point admittance'' [23] from electrical network theory.

Figure 4(a) and (b) show physical and symbolic network interpretations of the wet impedance, which has derivation and interpretation similar to those for the wet admittance. We break into the network at the $(\mathbf{p},\mathbf{v})$-port and consider the effect of a discontinuity in the velocity as we follow the normal across $\mathbf{S}$. This discontinuity, which we call the applied velocity, $\mathbf{v}^{\text{ap}}$, is the difference between the velocity on the surface of the structure in the equation $\mathbf{v}^{\text{in}}=-\mathbf{Y}_d\mathbf{p}$ and the velocity on the fluid in $\mathbf{v}^{\text{out}}=\mathbf{v}^{\text{in}}+\mathbf{v}^{\text{ap}}=\mathbf{Y}_{22}\mathbf{p}$. We assume that the pressure is continuous as we follow the normal across $\mathbf{S}$, and we consider the applied velocity to be the only source. As with $\mathbf{p}^{\text{ap}}$, it is not clear how to physically realize $\mathbf{v}^{\text{ap}}$, although we could consider either injecting and extracting fluid at some point on $\mathbf{S}$ or applying a film to $\mathbf{S}$ that has a controllable rate of change of thickness. Again, regardless of the realizability of the source, this section provides an interpretation of the quantity $[\mathbf{Y}_{22}+\mathbf{Y}_d]^{-1}$. The above equations combine to give $\mathbf{p}=\mathbf{Z}_w\mathbf{v}^{\text{ap}}$, where $\mathbf{Z}_w$ is the immersed object's pressure response to an applied normal velocity.

The network analogy shows two admittances in parallel with a vector of velocity sources. The wet impedance is the impedance seen by that source and is completely analogous to the ''augmented driving point impedance'' [23] from electrical network theory.

We should note that there is no agreed-upon definition of ''wet impedance.'' Borgiotti[1] defines a wet impedance,

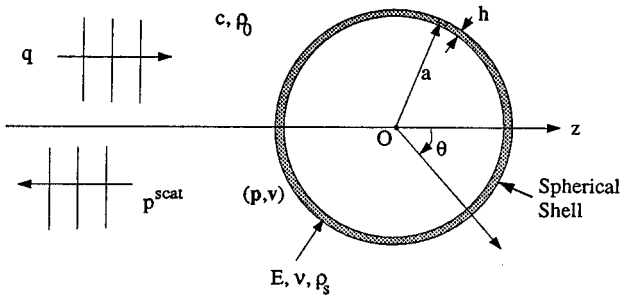


FIG. 5. Scattering configuration for a spherical shell.

$\mathbf{Y}_w^{-1}$, which he does not interpret, and Gaumond and Yoder[4] write down the wet impedance from this paper, but do not provide an interpretation. In this paper, $\mathbf{Z}_w \neq \mathbf{Y}_w^{-1}$ because the matrices relate different quantities. $\mathbf{Z}_w$ takes the surface velocity discontinuity to the resulting total pressure, and $\mathbf{Y}_w$ takes the surface pressure discontinuity to the resulting total normal velocity. The interpretations of $\mathbf{Y}_w^{-1}$ and $\mathbf{Z}_w^{-1}$ are not as appealing as those for $\mathbf{Y}_w$ and $\mathbf{Z}_w$. The ith column of $\mathbf{Y}_w^{-1}$ is the surface pressure discontinuity that is required to support an $\mathbf{e}_i$ surface velocity, and ith column of $\mathbf{Z}_w^{-1}$ is the surface velocity discontinuity that is required to support an $\mathbf{e}_i$ surface pressure.

B. Interaction resonances

The frequency and shape of interaction resonances can be determined by comparing $\mathbf{Z}_{22}$ to $\mathbf{Z}_d$. For a spherical shell, Junger and Feit[2] use this method to find the lower-branch resonances. These resonances are sufficiently narrow that a coarse frequency sampling does not capture the peaks in the scattered response.

For a target with arbitrary $\mathbf{S}$, the scattered pressure is written as in Eq. (2):

$$p^{\text{scat}} = \{\Delta z_{11}^r - \mathbf{z}_{12}^T[\mathbf{Z}_{22} + \mathbf{Z}_d]^{-1}\mathbf{z}_{21}\}q.$$

The inverted factor (the wet admittance) will be relatively large at frequencies where the matrix $\mathbf{Z}_{22}+\mathbf{Z}_d$ is nearly singular. The singular vectors associated with the small singular values of this matrix are the resonance shapes of the submerged structure. The entire second term will have a large contribution when the vectors $\mathbf{z}_{12}^T$ and $\mathbf{z}_{21}$ have components corresponding to these shapes.

We apply this analysis technique to the scattering from the spherical shell shown in Fig. 5. The scattering configuration's material parameters, fluid impedance matrices, and dry impedance matrix are given in Appendix B. Because the structure is lossless, $\text{Re}\{\mathbf{Z}_d\}=0$; because we always have some radiation from a sphere, no matter how high the spatial frequency of the field quantities, $\text{Re}\{\mathbf{Z}_{22}\}>0$. Given that $\mathbf{Z}_{22}$ and $\mathbf{Z}_d$ are diagonal,

$$p^{\text{scat}} = \left\{ \Delta z_{11}^r - \sum_n \frac{[\mathbf{z}_{12}^T]^n[\mathbf{z}_{21}]^n}{[\mathbf{Z}_{22}]^{nn} + [\mathbf{Z}_d]^{nn}} \right\} q, \tag{29}$$

and finding frequencies at which $\mathbf{Z}_{22}+\mathbf{Z}_d$ is nearly singular is greatly simplified. The second term's contribution to p^{scat} will be largest where the fluid and shell reactances of an individual mode are equal and opposite, i.e., $\text{Im}\{[\mathbf{Z}_{22}]^{nn}\} = -\text{Im}\{[\mathbf{Z}_d]^{nn}\}$, and $\text{Re}\{[\mathbf{Z}_{22}]^{nn}\}$ is small. This phenomenon is illustrated in Fig. 6. The top plot shows $\text{Im}\{[\mathbf{Z}_{22}]^{77}\}$, $-\text{Im}\{[\mathbf{Z}_d]^{77}\}$, and $\text{Re}\{[\mathbf{Z}_{22}]^{77}\}$. The fluid impedance is mass-like at lower ka and lossy at higher ka. The dry impedance of the shell is spring-like at lower ka and mass-like at higher ka. An arrow marks the point at which the ''fluid mass'' matches the ''shell spring,'' and the corresponding peak in the form function is indicated. The middle plot is similar to the top for $n=21$. For the frequency range shown in this plot, the peaks in the form function correspond to reactance matching in the $n=2$ through $n=22$ spherical harmonics. These peaks, of course, form the lower branch of the denominator zeros on the frequency-versus-mode-number dispersion plot for the submerged spherical shell.

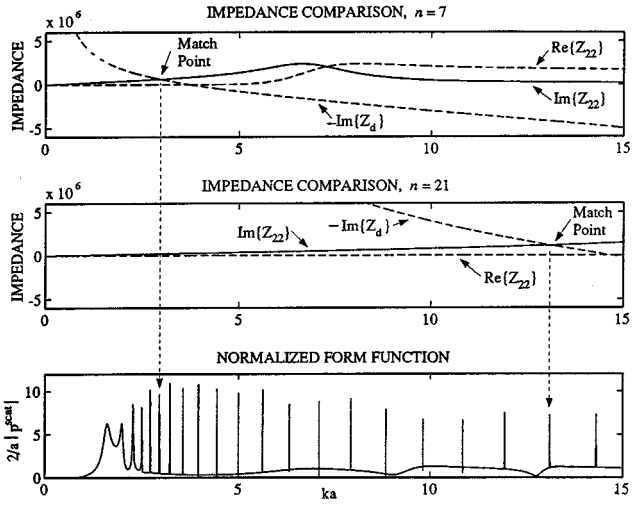


FIG. 6. Impedance comparison for $n=7,21$ and the spherical shell form function.

These peaks are particularly narrow and are plotted here with a frequency step of $\Delta ka = 10^{-8}$. Off resonance, the scattered field is plotted with $\Delta ka = 10^{-2}$. Without knowledge of the resonance frequencies, derived by impedance comparison, this relatively simple scattering calculation would be infeasible.

IV. SUMMARY AND CONCLUSIONS

This paper presents a generalized network formulation for acoustic scattering. The network formulation organizes the scattering calculation in a manner that allows us to exploit single-frequency circuit-theory analysis tools, in particular, change of field-variable basis and network element abstraction.

We change field-variable basis to determine the admittance, arbitrary-reference, and network-scattering representations. We use the network-scattering representation to determine bounds on the scattering from passive targets with limited spatial degrees of freedom, and we show that the lower-branch resonances of a spherical shell approaches these bounds. We also use a conjugate-matched load impedance to show that, for a target with unlimited spatial degrees

of freedom, we can construct a passive impedance which gives a form function greater than any bound.

We use network element abstraction to find the lower-branch resonances of a submerged spherical shell. Because these modes lose energy to radiation very slowly, the peaks are extremely sharp.[2] We show that being able to solve for the resonance frequencies allows us to increase the sampling density in their neighborhood and capture peaks that otherwise would be missed by a coarse frequency sampling. Network element abstraction is also used to interpret the wet admittance and wet impedance.

Many of these results are available through other means—the admittance and arbitrary-reference representations can be derived from first principles, and the intermediate matrix interpretations and resonance identification are also available directly. Our network formulation provides a unifying, conceptually simple, and analytically supportive framework for treating all these problems.

We have shown that network formulations are useful in analyzing acoustic scattering for: (i) developing formulations which emphasize different field quantities, (ii) interpreting intermediate results, (iii) determining features in the scattered response by means of impedance (or admittance) comparison, and (iv) examining power flow by means of the network scattering variables.

The utility of this formulation is not in improving the numerical aspects of scattering calculations, but in understanding properties of the scattering system once the calculations are complete, making more intuitive the properties of systems under design, and altering the formulation so that different models and measurements can be incorporated. The formulation also facilitates communication between electrical engineers and acousticians, and the attendant analysis tools enable simple derivation of unique results.

ACKNOWLEDGMENTS

I am grateful to G. Leonard Tyler and Al Bahr for instructive conversations throughout the preparation of this material. This work was supported in part by SRI International.

APPENDIX A: DERIVATION OF THE DISCRETIZED ACOUSTIC INTEGRAL EQUATIONS AND FORMATION OF THE FLUID IMPEDANCE MATRICES

In summary, the fluid-impedance submatrices which appear in Eq. (1) are given by

$$\mathbf{Z}_{22}=[\tfrac{1}{2}\mathbf{A}^{\psi\phi}-\mathbf{D}]^{-1}j\omega\rho_0\mathbf{G}, \tag{A1}$$

$$\mathbf{z}_{21}=[\tfrac{1}{2}\mathbf{A}^{\psi\phi}-\mathbf{D}]^{-1}\mathbf{g}^{(\mathbf{r}')}, \tag{A2}$$

$$\mathbf{z}_{12}^{T}=\mathbf{g}^{(\mathbf{r})T}[\tfrac{1}{2}\mathbf{A}^{\psi\phi}-\mathbf{D}]^{-T}\mathbf{A}^{\phi\phi}, \tag{A3}$$

$$\Delta z_{11}^{r}=\mathbf{g}^{(\mathbf{r})T}[\tfrac{1}{2}\mathbf{A}^{\psi\phi}-\mathbf{D}]^{-T}\mathbf{A}^{\phi\phi}[j\omega\rho_0\mathbf{G}]^{-1}[\tfrac{1}{2}\mathbf{A}^{\psi\phi}+\mathbf{D}] \times[\mathbf{A}^{\psi\phi}]^{-1}\mathbf{g}^{(\mathbf{r}')}, \tag{A4}$$

where ρ_0 is the fluid density.

The components of these submatrices are given by

$$[\mathbf{D}]^{li}=\oint_{\mathbf{S}}\left\{\psi^{l}(\mathbf{r}_1)\oint_{\mathbf{S}}\{(\nabla\mathcal{G}(\mathbf{r},\mathbf{r}_1)\cdot\hat{n})\phi^{i}(\mathbf{r})\}d\mathbf{S}\right\}d\mathbf{S}, \tag{A5}$$

$$[\mathbf{G}]^{li}=\oint_{\mathbf{S}}\left\{\psi^{l}(\mathbf{r}_1)\oint_{\mathbf{S}}\{\mathcal{G}(\mathbf{r},\mathbf{r}_1)\phi^{i}(\mathbf{r})\}d\mathbf{S}\right\}d\mathbf{S}, \tag{A6}$$

$$[\mathbf{A}^{\psi\phi}]^{li}=\oint_{\mathbf{S}}\psi^{l}(\mathbf{r})\phi^{i}(\mathbf{r})d\mathbf{S}, \tag{A7}$$

$$[\mathbf{A}^{\phi\phi}]^{li}=\oint_{\mathbf{S}}\phi^{l}(\mathbf{r})\phi^{i}(\mathbf{r})d\mathbf{S}, \tag{A8}$$

$$[\mathbf{g}^{(\mathbf{r}')}]^{l}=\oint_{\mathbf{S}}\{\psi^{l}(\mathbf{r})j\omega\rho_0\mathcal{G}(\mathbf{r},\mathbf{r}')\}d\mathbf{S}, \tag{A9}$$

where $[\cdot]^{li}$ indicates the (l,i)th matrix element, and

$$\mathcal{G}(\mathbf{r},\mathbf{r}')=\frac{\exp(-jk|\mathbf{r}-\mathbf{r}'|)}{4\pi|\mathbf{r}-\mathbf{r}'|} \tag{A10}$$

is Green's function for an infinite-extent fluid medium, $\mathbf{r}'$ is the source location, and $\mathbf{r}$ is the (Green's function) receiving location. We assume that the pressure and normal velocity on $\mathbf{S}$, from Fig. 2, are well approximated by

$$p(\mathbf{r})\approx\sum_{i=1}^{N}p^{i}\phi^{i}(\mathbf{r});\quad \mathbf{v}(\mathbf{r})\cdot\hat{n}\approx\sum_{i=1}^{N}v^{i}\phi^{i}(\mathbf{r}), \tag{A11}$$

where $\mathbf{r}$ is constrained to lie on $\mathbf{S}$, and from Eq. (1), $[\mathbf{p}]^{l}=p^{l}$ and $[\mathbf{v}]^{l}=v^{l}$. The ψ_i that appear in Eqs. (A5)–(A9) are method-of-moments testing functions.[24]

The remainder of this Appendix provides a detailed derivation of the fluid impedance matrices, starting from the first-order, microscopic, linearized equations of fluid motion.

1. The equivalence integral

The physical assumptions from Sec. I allow us to write the following first-order acoustic equations:[25]

$$\nabla\cdot\mathbf{v}(\mathbf{r})+\frac{j\omega}{\rho_0c^2}p(\mathbf{r})=\delta(\mathbf{r}-\mathbf{r}')q, \tag{A12}$$

$$\nabla p(\mathbf{r})+j\omega\rho_0\mathbf{v}(\mathbf{r})=0, \tag{A13}$$

where the right-hand side of this system is the independent source, and the left-hand side describes the dynamics of the fluid. Equation (A12) is derived from the continuity equation and the equation of state for the fluid. Equation (A13) is derived from the conservation of momentum. The velocity of the fluid at $\mathbf{r}$ is $\mathbf{v}(\mathbf{r})$, the dynamic pressure at $\mathbf{r}$ is $p(\mathbf{r})$, ρ_0 is the fluid's steady-state density, c is the speed of sound, ω is the angular frequency of variation, and j is $\sqrt{-1}$. The steady-state density and speed of sound are constant outside of $\mathbf{S}$. The source is a volume-flux-per-unit-volume monopole with amplitude q at $\mathbf{r}'$. These equations hold in all space external to $\mathbf{S}$.

To derive the equivalence integral, Green's theorem is applied to the field quantities in the above system.[26] The volume of integration is bounded on the interior by $\mathbf{S}$ and on the exterior by the far-field sphere. The Sommerfeld radia-

tion condition implies that the surface integral over the far-field sphere is zero, which yields one form of the equivalence integral:

$$\oint_{\mathbf{S}}\{-p_2(\mathbf{r})\mathbf{v}_1(\mathbf{r})+p_1(\mathbf{r})\mathbf{v}_2(\mathbf{r})\}\cdot\hat{n}\ d\mathbf{S}$$
$$=p_2(\mathbf{r}_1)q_1-p_1(\mathbf{r}_2)q_2, \tag{A14}$$

where the field quantities $p_1,\mathbf{v}_1$ and $p_2,\mathbf{v}_2$—corresponding to sources q_1 and q_2, respectively—are any two solutions of Eqs. (A12) and (A13) outside and on **S**. The source locations, here denoted $\mathbf{r}_1$ and $\mathbf{r}_2$, instead of $\mathbf{r}'$, can be anywhere outside **S**. The two solutions can correspond to arbitrarily different physical situations inside **S**. In the rest of this appendix, we will refer to the physical situation denoted by subscript 1 as ''Config. 1'' and the physical situation denoted by subscript 2 as ''Config. 2.'' For example, Config. 1 could have the scattering object inside **S**, while Config. 2 could have **S** a rigid surface.

Let us assume, for the moment, that either both configurations have a rigid **S** or both have a pressure-release **S**. The boundary condition for a rigid surface is $\mathbf{v}\cdot\hat{n}=0$, and the boundary condition for a pressure-release surface is $p=0$. In both cases, Eq. (A14) becomes

$$p_1(\mathbf{r}_2)q_2=p_2(\mathbf{r}_1)q_1. \tag{A15}$$

Assuming that $q_1=q_2$, we can write

$$p_1(\mathbf{r}_2)=p_1^{(\mathbf{r}_2)}(\mathbf{r}_1), \tag{A16}$$

which is simply a statement of reciprocity for physical configurations with rigid or pressure-release **S**. On the rhs, the subscript-2 field variable has been replaced by subscript 1 because the configurations are identical except for source location, while the source location is given by the superscript since it does not coincide with that implied by the subscript.

Equation (A16) also holds when both configurations have **S** filled exclusively with external fluid. In this case, (A16) becomes a statement of reciprocity for the free-space Green's function.

Substituting (A16) into (A14) gives

$$\oint_{\mathbf{S}}\{-p_2(\mathbf{r})\mathbf{v}_1(\mathbf{r})+p_1(\mathbf{r})\mathbf{v}_2(\mathbf{r})\}\cdot\hat{n}\ d\mathbf{S}$$
$$=p_2(\mathbf{r}_1)q_1-p_1^{(\mathbf{r}_2)}(\mathbf{r}_1)q_2, \tag{A17}$$

which, we have shown, holds for Config. 1 rigid, pressure-release, or fluid-filled. The source location for both configurations is $\mathbf{r}_2$, and the receiving location for both configurations is $\mathbf{r}_1$. Thus if we take $q_1=q_2=1$, the rhs becomes a difference of total fields. We usually choose the physical situations inside **S** to be different—for example, if we choose Config. 1 to be the fluid-filled **S** and Config. 2 to be the scattering structure, the rhs is the pressure scattered from the structure.

The method of moments[24] is used to discretize this equation. Substituting Eq. (A11) into (A17) yields

$$-\mathbf{p}_2^T\mathbf{A}^{\phi\phi}\mathbf{v}_1+\mathbf{v}_2^T\mathbf{A}^{\phi\phi}\mathbf{p}_1=p_2(\mathbf{r}_1)q_1-p_1^{(\mathbf{r}_2)}(\mathbf{r}_1)q_2, \tag{A18}$$

where the boldface lowercase letters are vectors, and the boldface uppercase letters are matrices. The quantities **p** and **v** are column vectors of the discretized pressure and normal velocity. The row vector $\mathbf{p}^T$ is the transpose of **p**.

2. The surface Helmholtz integral equation

The Helmholtz integral equation (HIE) can be derived from Eq. (A17) by choosing the fluid-filled **S** for Config. 1. The Config. 1 field quantities are then known everywhere:

$$p_1(\mathbf{r})=j\omega\rho_0\mathcal{G}(\mathbf{r},\mathbf{r}_1)q_1,\quad \mathbf{v}_1(\mathbf{r})=-\boldsymbol{\nabla}_{\mathbf{r}}\mathcal{G}(\mathbf{r},\mathbf{r}_1)q_1, \tag{A19}$$

where $\mathcal{G}$, Green's function for an infinite-extent fluid, is given in Eq. (A10), $\mathbf{r}_1$ is the source location, and $\mathbf{r}$ is the receiving location. Substituting (A19) into (A17) yields

$$\oint_{\mathbf{S}}\{p_2(\mathbf{r})\boldsymbol{\nabla}\mathcal{G}(\mathbf{r},\mathbf{r}_1)+j\omega\rho_0\mathcal{G}(\mathbf{r},\mathbf{r}_1)\mathbf{v}_2(\mathbf{r})\}\cdot\hat{n}\ d\mathbf{S}$$
$$=p_2(\mathbf{r}_1)-j\omega\rho_0\mathcal{G}(\mathbf{r}_1,\mathbf{r}_2)q_2. \tag{A20}$$

This equation was derived by considering $\mathbf{r}_1$ outside **S** and then applying Green's theorem. If we consider $\mathbf{r}_1$ inside **S**, the first term on the rhs does not appear, which implies that the rhs is discontinuous as $\mathbf{r}_1$ follows the normal across **S**. The value of the rhs for $\mathbf{r}_1$ on **S** is resolved by applying a three-dimensional version of the principal value theorem,[27] which yields

$$\oint_{\mathbf{S}}\{p_2(\mathbf{r})\boldsymbol{\nabla}\mathcal{G}(\mathbf{r},\mathbf{r}_1)+j\omega\rho_0\mathcal{G}(\mathbf{r},\mathbf{r}_1)\mathbf{v}_2(\mathbf{r})\}\cdot\hat{n}\ d\mathbf{S}$$
$$=\frac{\alpha}{4\pi}p_2(\mathbf{r}_1)-j\omega\rho_0\mathcal{G}(\mathbf{r}_1,\mathbf{r}_2)q_2, \tag{A21}$$

where α is the solid angle subtended by the surface at the point $\mathbf{r}_1\in\mathbf{S}$. For example, if **S** has no edges or vertices, then $\alpha=2\pi$ at every point on **S**.

To discretize the HIE, we start by substituting the approximation for the surface field quantities from (A11) into (A21), which gives

$$\sum_i\ \oint_{\mathbf{S}}\{(\boldsymbol{\nabla}\mathcal{G}(\mathbf{r},\mathbf{r}_1)\cdot\hat{n})p_2^i\phi^i(\mathbf{r})+j\omega\rho_0\mathcal{G}(\mathbf{r},\mathbf{r}_1)v_2^i\phi^i(\mathbf{r})\}d\mathbf{S}$$
$$=\frac{\alpha}{4\pi}\sum_i p_2^i\phi^i(\mathbf{r}_1)-j\omega\rho_0\mathcal{G}(\mathbf{r}_1,\mathbf{r}_2)q_2. \tag{A22}$$

To fully discretize this equation, we define the testing functions, $\{\psi^j(\mathbf{r}_1);\ i=1-N,\ \mathbf{r}_1\in\mathbf{S}\}$, where N is the number of field-variable basis functions, and integrate each testing function in turn against the entire expression. If every ψ_i has finite area on **S**, then the different values of α at the edges and vertices of **S** do not contribute to the integration against the testing functions, and we can take $\alpha=2\pi$. This yields

$$\mathbf{D}\mathbf{p}_2+j\omega\rho_0\mathbf{G}\mathbf{v}_2=\tfrac{1}{2}\mathbf{A}^{\psi\phi}\mathbf{p}_2-\mathbf{g}^{(\mathbf{r}_2)}q_2, \tag{A23}$$

where **D**, **G**, $\mathbf{A}^{\psi\phi}$, and $\mathbf{g}^{(\mathbf{r}_2)}$ are given in Eqs. (A5)–(A9). If the testing functions *do not* extend over a finite area on **S**—i.e., if we choose one or more of the ψ_i to be Dirac delta functions—and if the testing locations are chosen on edges or vertices, then Eq. (A23) can be simply modified to ac-

count for the different values of α. In the rest of this appendix, we assume that every ψ_i has finite area or that the testing locations are chosen such that $\alpha=2\pi$.

We assume that the range and domain bases provide an adequate approximation to the integral equations above. As an example of the bases used in a standard code, the combined Helmholtz integral equation formulation (CHIEF)[21] uses a subsectional basis for $\{\phi\}$ and point-matching for $\{\psi\}$. Each ϕ^i in the subsectional basis is unity over a patch on $\mathbf{S}$ and zero everywhere else, and each ψ^j in the point-matching basis is a Dirac delta, which CHIEF places at the center of each patch. Point-matching is by far the most popular choice for $\{\psi\}$, because it simplifies the second integration in Eq. (A23). Bases other than point-matching are most often used when $\mathbf{S}$ is a constant coordinate surface in a separable coordinate system. In Appendix B, where $\mathbf{S}$ is a sphere, choosing the spherical harmonics for $\{\phi\}$ and $\{\psi\}$ results in diagonal $\mathbf{D}$, $\mathbf{G}$, and $\mathbf{A}^{\phi\psi}$ matrices, which allow harmonic-by-harmonic examination of the fluid impedance matrices.

3. Fluid impedance matrices

In Eq. (1) of the the text, the relationships between the field quantities $\mathbf{p}$, $\mathbf{v}$, q, and p^{scat} define the fluid impedance matrices. As stated in the introduction, many different acoustic modeling methods could be used to determine these matrices. This section uses Eqs. (A18) and (A23), the discretized equivalence integral and the discretized HIE.

In Eq. (1), the ith column of $\mathbf{Z}_{22}$ is defined as the discretized surface pressure, $\mathbf{p}$, given $q=0$ and discretized surface velocity $\mathbf{v}=\mathbf{e}_i$—i.e., $\mathbf{Z}_{22}$ relates the surface pressure and velocity when there are no active sources external to $\mathbf{S}$. $\mathbf{Z}_{22}$ is found quite simply from the discretized HIE. We take Config. 2, here unsubscripted, to be the penetrable scatterer. (Recall that the HIE is derived by taking Config. 1 to be the fluid-filled $\mathbf{S}$.) We assume that $\mathbf{r}_2$ is inside $\mathbf{S}$, which places the source outside the region of integration in Green's theorem and eliminates the second term on the rhs of (A23). The discretized HIE can be rewritten

$$\mathbf{p}=[\tfrac{1}{2}\mathbf{A}^{\psi\phi}-\mathbf{D}]^{-1}j\omega\rho_0\mathbf{G}\mathbf{v}$$

$$\Rightarrow\mathbf{Z}_{22}=[\tfrac{1}{2}\mathbf{A}^{\psi\phi}-\mathbf{D}]^{-1}j\omega\rho_0\mathbf{G}, \tag{A24}$$

which is the well-known Helmholtz impedance.

In Eq. (1), the column vector $\mathbf{z}_{21}$ is the discretized surface pressure; $\mathbf{p}$, given $q=1$ and $\mathbf{v}=0$—i.e., $\mathbf{z}_{21}$ is the unit-source-strength pressure on a rigid $\mathbf{S}$. $\mathbf{z}_{21}$ is also determined using the discretized HIE. In this case, we assume that Config. 2 is the rigid $\mathbf{S}$, here indicated by subscript r. The $\mathbf{v}=0$ boundary condition yields the following relationship between $\mathbf{p}_r$ and q_r:

$$\mathbf{p}_r=[\tfrac{1}{2}\mathbf{A}^{\psi\phi}-\mathbf{D}]^{-1}\mathbf{g}^{(\mathbf{r}')}q_r\Rightarrow\mathbf{z}_{21}=[\tfrac{1}{2}\mathbf{A}^{\psi\phi}-\mathbf{D}]^{-1}\mathbf{g}^{(\mathbf{r}')}, \tag{A25}$$

where $\mathbf{p}_r$ is the pressure that would exist on a rigid body, and q_r is the source corresponding to that pressure. The superscript on g indicates the location of the source point. This relationship for $\mathbf{p}_r$ and q_r is a discretized version of the rigid-body Green's function.

The ith element of row vector $\mathbf{z}_{12}^T$ is the radiated pressure, given $q=0$ and $\mathbf{v}=\mathbf{e}_i$. $\mathbf{z}_{12}^T$ is determined using the discretized equivalence integral, Eq. (A18). In this case, we assume that Config. 1 is the rigid $\mathbf{S}$, indicated by subscript r. Config. 2 is the penetrable scatterer, here unsubscripted. We take the Config. 2 source interior to $\mathbf{S}$, which eliminates the second term on the rhs. This yields

$$\mathbf{p}_r^T\mathbf{A}^{\phi\phi}\mathbf{v}=p(\mathbf{r})q_r, \tag{A26}$$

where $p(\mathbf{r})$ is the penetrable scatterer radiated pressure at $\mathbf{r}$, the rigid $\mathbf{S}$ source point. Substituting from Eq. (A25), with source at $\mathbf{r}$, we get

$$p(\mathbf{r})=[[\tfrac{1}{2}\mathbf{A}^{\psi\phi}-\mathbf{D}]^{-1}\mathbf{g}^{(\mathbf{r})}]^T\mathbf{A}^{\phi\phi}\mathbf{v}$$

$$\Rightarrow\mathbf{z}_{12}^T=\mathbf{g}^{(\mathbf{r})T}[\tfrac{1}{2}\mathbf{A}^{\psi\phi}-\mathbf{D}]^{-T}\mathbf{A}^{\phi\phi}, \tag{A27}$$

which gives $p(\mathbf{r})$ for any $\mathbf{v}$.

And finally, the scalar Δz_{11}^r is the scattered pressure, given $q=1$ and $\mathbf{v}=0$—i.e., the Δz_{11}^r is the unit-source-strength scattered pressure from a rigid $\mathbf{S}$. Here, Δz_{11}^r is determined using both the discretized HIE and equivalence integral. In the discretized equivalence integral, Config. 1 is the fluid-filled $\mathbf{S}$, indicated by subscript i for ''incident field,'' and Config. 2 is the rigid $\mathbf{S}$, indicated by subscript r. The source for Config. 1 is at $\mathbf{r}'$, the source location, and the source for Config. 2 is at $\mathbf{r}$, the receiving location. This yields

$$-\mathbf{p}_r^T\mathbf{A}^{\phi\phi}\mathbf{v}_i=p_r(\mathbf{r})q_i-p_i^{(\mathbf{r}')}(\mathbf{r})q_r. \tag{A28}$$

To find the rigid-body scattered pressure, we set $q_i=q_r$ and note that the scattered field is defined as $p_r^{\text{scat}}=p_r-p_i$. Thus

$$p_r^{\text{scat}}=-\left(\frac{1}{q_r}\right)\mathbf{p}_r^T\mathbf{A}^{\phi\phi}\mathbf{v}_i=\mathbf{z}_{12}^T\mathbf{v}_i. \tag{A29}$$

Now we have to find $\mathbf{v}_i$, the incident velocity, and we are done. The incident velocity can be found from geometrical arguments, or from the discretized HIE with Config. 2 the fluid-filled S. With Config. 2 indicated by subscript i, the latter yields

$$\mathbf{D}\mathbf{p}_i+j\omega\rho_0\mathbf{G}\mathbf{v}_i=\tfrac{1}{2}\mathbf{A}^{\psi\phi}\mathbf{p}_i-\mathbf{g}^{(\mathbf{r}')}q, \tag{A30}$$

where the $\mathbf{p}_i$ and $\mathbf{v}_i$ are in the $\{\phi\}$ basis, and $\mathbf{g}^{(\mathbf{r}')}q$ is the incident pressure projected into the $\{\psi\}$ basis. To solve for $\mathbf{v}_i$ in terms of q, we must reconcile this difference in discretization. Assuming that the incident pressure is reasonably approximated as a sum over $\phi^i(\mathbf{r})$, the two representations are related by $\mathbf{g}^{(\mathbf{r}')}q=\mathbf{A}^{\psi\phi}\mathbf{p}_i$. Substituting into Eq. (A30) yields,

$$\mathbf{v}_i=-[j\omega\rho_0\mathbf{G}]^{-1}[\tfrac{1}{2}\mathbf{A}^{\psi\phi}+\mathbf{D}][\mathbf{A}^{\psi\phi}]^{-1}\mathbf{g}^{(\mathbf{r}')}q \tag{A31}$$

and then, substituting back into (A29), we find that

$$\Delta z_{11}^r=\mathbf{g}^{(\mathbf{r})T}[\tfrac{1}{2}\mathbf{A}^{\psi\phi}-\mathbf{D}]^{-T}\mathbf{A}^{\phi\phi}[j\omega\rho_0\mathbf{G}]^{-1}[\tfrac{1}{2}\mathbf{A}^{\psi\phi}+\mathbf{D}]$$

$$\times[\mathbf{A}^{\psi\phi}]^{-1}\mathbf{g}^{(\mathbf{r}')}. \tag{A32}$$

This completes the solution for the submatrices in Eq. (1) of the text.

Far-field source and receive locations (directions) can be included using the standard plane-wave assumptions. We allow $|\mathbf{r}'|$ and $|\mathbf{r}|$ to recede to ∞ while keeping the incident field amplitude constant at the origin. The source strength is redefined to be this incident field amplitude, and the scattered pressure is redefined to be the far-field pressure referred to 1 m.

APPENDIX B: FLUID AND STRUCTURAL NETWORK MATRICES FOR THE SPHERICAL SHELL

This Appendix calculates the fluid impedance matrices and structural admittance matrix for the scattering configuration shown in Fig. 5. In the numerical example, we take $a=5$ m, $h=0.15$ m, $c=1480$ m/s, $\rho_0=1000$ kg/m^3, $E=19.6\times10^{10}$ N/m^2, $\nu=0.3$, and $\rho_s=7668.7$ kg/m^3.

1. Fluid impedance matrix for spherical S

For the spherical geometry, with axisymmetric field variation, we choose $\{\phi_n\}$ and $\{\psi_n\}$,

$$\phi_n=\psi_n=\sqrt{\frac{2n+1}{4\pi}}\,P_n(\cos(\phi)), \tag{B1}$$

where P_n is the Legendre polynomial of degree n. This choice diagonalizes the fluid and structure network matrices[28] and enables analytic calculation of the matrix elements.

We substitute Eq. (B1) into (A5)–(A9), which yields

$$[\mathbf{D}]^{lj}=-\frac{ik^2a^4}{2}\{j_l'(ka)h_l(ka)+j_l(ka)h_l'(ka)\}\delta_{lj}, \tag{B2}$$

$$[\mathbf{G}]^{lj}=-ika^4 j_l(ka)h_l(ka)\delta_{lj}, \tag{B3}$$

$$[\mathbf{A}^{\psi\phi}]^{lj}=a^2\delta_{lj}, \tag{B4}$$

$$[\mathbf{A}^{\phi\phi}]^{lj}=a^2\delta_{lj}, \tag{B5}$$

$$[\mathbf{g}^{(\mathbf{r}')}]^{l}=a^2(-i)^l\sqrt{4\pi(2l+1)}\,j_l(ka), \tag{B6}$$

$$[\mathbf{g}^{(\mathbf{r})}]^{l}=a^2 i\omega\rho_0(-i)^l\sqrt{\frac{(2l+1)}{4\pi}}\,j_l(ka), \tag{B7}$$

where the Cauchy principal value is used to calculate $\mathbf{D}$, k is the wave number, a is the radius, $j_l(ka)$ is the lth spherical Bessel function evaluated at ka, $h_l(ka)$ is the lth spherical Hankel function of the second kind evaluated at ka, and δ_{lj} is the Kronecker delta. The prime superscript on j and h indicates derivative with respect to the argument. The forms of $\mathbf{g}^{(\mathbf{r}')}$ and $\mathbf{g}^{(\mathbf{r})}$ are determined by the source and receiving location. The source is taken to be the amplitude of the incident plane wave at the origin, and the receiving location (direction) is taken to be the backscattered pressure in the far field, referred to 1 m.

Using the Wronskian,[29] we can write

$$[\tfrac{1}{2}\mathbf{A}^{\phi\psi}-\mathbf{D}]^{ll}=(ka)^2 ia^2 j_l(ka)h_l'(ka), \tag{B8}$$

where we see that both $[\frac{1}{2}\mathbf{A}^{\psi\phi}-\mathbf{D}]$ and $\mathbf{G}$ are singular at the roots of the spherical Bessel functions. These critical frequencies[21] correspond to resonant conditions on the interior fluid volume, in which case the relationship between $\mathbf{p}$ and $\mathbf{v}$ in the surface HIE is not one-to-one. We note that these matrix singularities cancel in the analytic formulation of $\mathbf{Z}_{22}$, $\mathbf{z}_{21}$, and $\mathbf{z}_{12}^T$.

We form the first two impedance matrices using Eqs. (A1) and (A2):

$$[\mathbf{Z}_{22}]^{ll}=-ic\rho_0\frac{h_l(ka)}{h_l'(ka)}, \tag{B9}$$

$$[\mathbf{z}_{21}]^{l}=(2l+1)(-i)^l\sqrt{\frac{4\pi}{2l+1}}\,\frac{-i}{(ka)^2h_l'(ka)}, \tag{B10}$$

$$[\mathbf{z}_{12}^T]^{l}=\frac{c\rho_0}{k}\sqrt{\frac{2l+1}{4\pi}}\,\frac{(-i)^l}{h_l'(ka)}, \tag{B11}$$

$$\Delta z_{11}^r=\frac{-i}{k}\sum_l(2l+1)(-1)^l\frac{j_l'(ka)}{h_l'(ka)}. \tag{B12}$$

2. Spherical shell admittance

The shell model used to determine the spherical-shell dry admittance includes shear as well as membrane stresses. We include no waves traveling from the interior surface to the exterior surface. The force on and motion of the shell are represented by the force on and the motion of the midplane. From Junger and Feit,[2] using the basis functions defined above, the admittance of the shell is given as

$$Y_{lj}=\frac{i\omega\mathscr{E}A_l}{A_lD_l-B_lC_l}\,\delta_{lj}, \tag{B13}$$

where

$$A_l=\Omega^2-(1+\beta^2)(\nu+\lambda_l-1), \tag{B14}$$

$$B_l=\beta^2(\nu+\lambda_l-1)+(1+\nu), \tag{B15}$$

$$C_l=-\lambda_l[\beta^2(\nu+\lambda_l-1)+(1+\nu)], \tag{B16}$$

$$D_l=\Omega^2-2(1+\nu)-\beta^2\lambda_l(\nu+\lambda_l-1), \tag{B17}$$

and

$$\mathscr{E}=-\frac{a^2(1-\nu^2)}{Eh},\quad \Omega^2=\left(\frac{a\omega}{c_p}\right)^2,\quad \beta^2=\frac{h^2}{12a^2},$$

$$c_p=\left(\frac{E}{(1-\nu^2)\rho_s}\right)^{1/2},\quad \lambda_l=l(l+1).$$

Note that the radius parameter used in the shell model is that of the shell's middle surface, and that the radius parameter used in the fluid model is that of the shell's outer surface. In order to match a previously published analysis,[30] we take those radii to be the same.

[1] G. V. Borgiotti, ''Effect of compliance on the scattering of an elastic object immersed in fluid: A general formulation,'' J. Acoust. Soc. Am. **87**, 1055–1061 (1990).
[2] M. C. Junger and D. Feit, *Sound, Structures, and Their Interaction* (MIT, Cambridge, MA, 1986).
[3] R. Vasudevan and J. M. McCormick, ''Substructuring techniques for acoustic radiation from complex submerged structures,'' in *Numerical Techniques in Acoustic Radiation* (American Society of Mechanical Engineers, New York, 1989).
[4] C. F. Gaumond and T. Yoder, ''Determination of structural impedance from scattering data,'' J. Acoust. Soc. Am. **97**, 1415–1422 (1995).
[5] E. S. Kuh and R. A. Rohrer, *Theory of Linear Active Networks* (Holden-Day, San Francisco, CA, 1967).
[6] V. V. Muzychenko and S. A. Rybak, ''Low-frequency resonance scattering of sound by finite cylindrical shells (review),'' Sov. Phys. Acoust. **34(4)**, 325–333 (July–August 1988).
[7] R. F. Harrington, ''Theory of loaded scatterers,'' Proc. Inst. Electr. Eng. **111**(4), 617–623 (1964).
[8] D. S. Burnett, ''A three-dimensional acoustic infinite element based on a prolate spheroidal multipole expansion,'' J. Acoust. Soc. Am. **96**, 2798–2816 (1994).
[9] J.-P. Coyette and H. Wynendaele, ''A synthesis of available boundary element formulations for solving acoustic and elasto-acoustic problems,'' in *Proceedings of the Eleventh International Modal Analysis Conference*, pp. 1380–1388, 1993.
[10] X. Zeng and J. Bielak, ''Stability assessment of a unified variational boundary integral method applicable to thin scatterers and scatterers with corners,'' Comput. Methods Appl. Mech. Eng. **111**, 305–321 (1994).
[11] L. G. Copley, ''Integral equation method for radiation from vibrating bodies,'' J. Acoust. Soc. Am. **41**, 807–816 (1967).
[12] K. A. Cunefare, G. Koopmann, and K. Brod, ''A boundary element method for acoustic radiation valid for all wavenumbers,'' J. Acoust. Soc. Am. **85**(1), 39–48 (1989).
[13] D. T. Wilton, I. C. Mathews, and R. A. Jeans, ''A clarification of the nonexistence problems with the superposition method,'' J. Acoust. Soc. Am. **94**, 1676–1680 (1993).
[14] H. A. Schenck, ''Improved integral formulation for acoustic radiation problems,'' J. Acoust. Soc. Am. **44**, 41–58 (1968).
[15] G. T. Schuster and L. C. Smith, ''A comparison among four direct boundary integral methods,'' J. Acoust. Soc. Am. **77**, 850–864 (1985).
[16] T. W. Wu and A. F. Seybert, ''A weighted residual formulation for the chief method in acoustics,'' J. Acoust. Soc. Am. **90**, 1608–1619 (1991).
[17] R. Martinez, ''The thin shape breakdown (TSB) of the Helmholtz integral equation,'' J. Acoust. Soc. Am. **90**, 2728–2738 (1991).
[18] L. G. Copley, ''Fundamental results concerning integral representations in acoustic radiation,'' J. Acoust. Soc. Am. **44**, 28–32 (1968).
[19] A. J. Burton and G. F. Miller, ''The application of integral equation methods to the numerical solution of some exterior boundary value problems,'' Proc. R. Soc. London **323**, 201–210 (1971).
[20] P. J. Titterton, Jr., ''Network analysis of scattering from an internally loaded spherical shell: resonance identification and frequency-by-frequency limits on the scattered field,'' J. Acoust. Soc. Am. **98**, 1667–1672 (1995).
[21] G. W. Benthien and D. Barach, *CHIEF Users Manual* (Naval Ocean Systems Center, February 1986).
[22] M. F. Werby and G. Gaunaurd, ''Transition from soft to rigid behavior in scattering from submerged thin elastic shells,'' Acoust. Lett. **9**(7), 89–93 (1986).
[23] R. A. Rohrer, ''The scattering matrix: Normalized to complex n-port loads,'' IEEE Trans. Circuit Theory **12(2)**, 223–230 (1965).
[24] R. F. Harrington, *Field Computation by Moment Methods* (Krieger, Malabar, FL, 1968).
[25] L. D. Landau and E. M. Lifshitz, ''Fluid Mechanics,'' volume 6 of *Course of Theoretical Physics* (Pergamon, New York, 1959), Chap. 1.
[26] R. F. Harrington, *Time Harmonic Electromagnetic Fields* (McGraw-Hill, New York, 1961), pp. 116–120.
[27] A. J. Poggio and E. K. Miller, ''Integral equation solutions of three dimensional scattering problems,'' in *Computational Techniques for Electromagnetics*, edited by R. Mittra (Hemisphere, New York, 1987), Chap. 4, pp. 159–264.
[28] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).
[29] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).
[30] G. C. Gaunaurd, H. Huang, and W. Wertman, ''Acoustic scattering by elastic spherical shells that have multiple massive internal components attached by compliant mounts,'' J. Acoust. Soc. Am. **94**, 2924–2935 (1993).

New phenomena on King integral with dissipation

Gee-Pinn James Too
Department of Naval Architecture and Marine Engineering, National Cheng Kung University, Tainan, Taiwan 70101, Republic of China

(Received 5 October 1995; revised 1 June 1996; accepted 8 August 1996)

Sound beams generated by an arbitrary baffled projector are considered. A wave equation which considers thermoviscous effects is derived in terms of nondimensional velocity potential. Then, an improved King integral solution that includes the thermoviscous effect is developed here. The solution shows an interesting phenomenon: that the dissipation effect contributes exponential decay in the propagation spectrum. In addition, the effect also contributes a propagation mode in the evanescent spectrum. The dissipation effects on sound beams are investigated numerically on axis and off axis with propagation spectrum and evanescent spectrum, respectively. © *1997 Acoustical Society of America.* [S0001-4966(97)02912-3]

PACS numbers: 43.20.Rz, 43.20.Tb, 43.38.Ar [JEG]

INTRODUCTION

In this paper, a dissipated model in terms of a nondimensional velocity potential is obtained. A solution including thermoviscous effects is obtained by considering appropriate boundary conditions. An axisymmetric wave propagation problem is considered in the present study. In the earlier study, the King integral,[1] which is a Hankel transform expression, is a well-known solution for a linear wave equation without dissipation. Dissipation models with a dissipation solution have been well studied in the early works.[2–4] In addition, Harris[5] and Tjotta *et al.*[6,7] both developed a dissipation model by solving the Helmholtz equation. Therefore, their solutions are developed in the frequency domain. In the present study, another improved King integral that considers the thermoviscous effect is obtained. It is found in the solution that there is a propagation mode in the evanescent spectrum. This feature will be derived and discussed in the present study.

I. LINEAR WAVE EQUATION INCLUDING THERMOVISCOUS EFFECTS

A linear wave equation including thermoviscous effects in terms of nondimensional velocity potential is derived. In order to describe thermoviscous effects, the fundamental equations of fluid motion should be used to derive the wave equation. The fluid in which waves propagate is assumed to be irrotational, viscous, gravityless, and initial quiescent with a uniform pressure P_0, uniform density ρ_0, and an ambient speed of sound c_0. The equations of motion governing waves in this kind of fluid become[8] momentum equation:

$$\rho \frac{D\mathbf{V}}{Dt} = -\nabla P, \tag{1}$$

continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{V}) = 0, \tag{2}$$

state equation:

$$P = P(\rho, s), \tag{3}$$

heat transfer equation:

$$\rho T \frac{\partial s}{\partial t} = \kappa \Delta T, \tag{4}$$

where η, μ, κ are coefficients of bulk viscosity, shear viscosity, and thermal conductivity, respectively.

In the case of energy dissipation, the equation of state of the medium generally cannot be assumed to be adiabatic. From Eqs. (3) and (4), the approximate equation of state is obtained[9] by keeping linear terms:

$$P' = c_0^2 \rho' - \kappa \left(\frac{1}{c_v} - \frac{1}{c_p} \right) \nabla \cdot \mathbf{V}, \tag{5}$$

where C_v and C_p are specific heat at constant volume and pressure, respectively. In addition, P' and ρ' represent acoustic pressure and acoustic density, respectively.

Away from boundary layers and sound sources, the flow is assumed to be irrotational, $\nabla \cdot \phi = 0$, which is the condition for existence of a velocity potential,

$$\mathbf{V} = \nabla \phi \tag{6}$$

Combining Eqs. (1), (2), (5), and (6) leads to a linear thermoviscous wave equation in terms of velocity potential ϕ:

$$\nabla^2 \phi + \frac{D}{c_0^2} \frac{\partial}{\partial t} (\nabla^2 \phi) - \frac{1}{c_0^2} \frac{\partial^2 \phi}{\partial t^2} = 0, \tag{7}$$

where

$$D = \left[\eta + \left(\frac{4}{3} \right) \mu + \kappa \left(\frac{1}{c_v} - \frac{1}{c_p} \right) \right] \rho_0^{-1}$$

is the sound diffusivity.[10]

Considering a monochromatic source operating in a frequency ω_0, a set of nondimensional variable defined by

$$\begin{aligned} &\nabla^* = \nabla / k, \quad R^* = kR, \quad t^* = \omega_0 t, \\ &\mathbf{x}^* = kx, \quad \mathbf{V}^* = \frac{\mathbf{V}}{c_0}, \quad k = \frac{\omega_0}{c_0}, \end{aligned} \tag{8}$$

are introduced into Eq. (7), which leads to the following dimensionless equation, in which * is omitted for brevity:

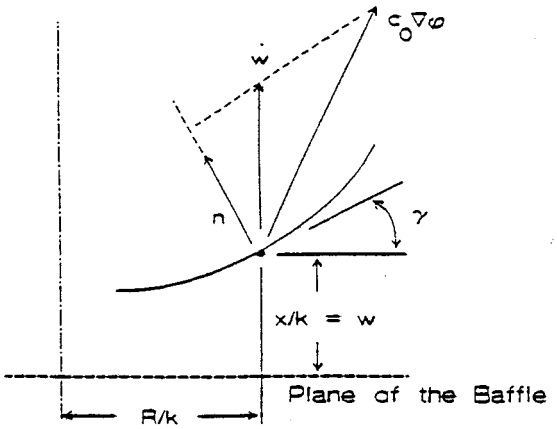


FIG. 1. Continuity of normal velocity at the projector–fluid interface.

$$\nabla^2\phi+\varsigma\,\frac{\partial}{\partial t}\,(\nabla^2\phi)-\frac{\partial^2\phi}{\partial t^2}=0, \tag{9}$$

where ζ is nondimensional dissipated coefficient$=D\omega_0/c_0^2$.

This is a linear wave equation with thermoviscosity in terms of nondimensional velocity potential.

II. BOUNDARY CONDITION

Let $w(R,t)$ denotes the normal displacement of the projector face, as shown in Fig. 1. The continuity condition on the projector face is that the normal component of the fluid particle and of projector motion must be equal, that is,

$$c_0\left(\cos\,\gamma\,\frac{\partial\phi}{\partial x}=-\sin\,\gamma\,\frac{\partial\phi}{\partial R}\right)_{x/k=w}=\dot{w}\,\cos\,\gamma, \tag{10}$$

where x, R, t are the nondimensional quantities defined in the previous section and

$$\gamma=\tan^{-1}\left(k\,\frac{\partial w}{\partial R}\right).$$

If $f(R)$ be an amplitude shading function which may be complex. A general presentation of monochromatic oscillation at frequency ω_0 and acoustic Mach number ϵ is

$$\dot{w}=\frac{1}{2i}\,\epsilon c_0 f(R)\exp(it)+\text{c.c.}, \tag{11}$$

$$w=-\frac{1}{2k}\,\epsilon c_0 f(R)\exp(it)+\text{c.c.} \tag{12}$$

Because w is $O(\epsilon)$, the surface rotation γ may be replaced by its tangent. For the same reason, the left-hand side of Eq. (10) can be expanded in Taylor series about the undeformed location of the projector face, $x=0$, and truncated at first order, which leads to a boundary condition at the fluid–baffle interface,

$$c_0\left(\frac{\partial\phi}{\partial x}\right)_{x=0}=\dot{w}. \tag{13}$$

III. KING INTEGRAL WITH THERMOVISCOSITY

Equations (9) and (13) are the linear equation including thermoviscous effect for an arbitrary baffled projector. In the nondissipative case, where ζ is set to be zero, the solution may be expressed in the form of the King integral. The Hankel transform of the amplitude shading function is

$$F_n=\int_0^\infty Rf(R)J_0(nR)dR+\text{c.c.} \tag{14}$$

The King integral is the inverse Hankel transform given by

$$\phi=\frac{\epsilon}{2}\int_0^\infty G_n\,\exp(it-\mu_n x)J_0(nR)dn+\text{c.c.}, \tag{15}$$

where

$$\mu_n=\begin{cases} i(1-n^2)^{1/2}, & n<1,\\ (n^2-1)^{1/2}, & n>1,\end{cases}$$

$$G_n=nF_n/\mu_n.$$

Note that the domain of transverse wave number $n<1$ corresponds to axisymmetric propagating spectrum in a circular waveguide, whereas $n>1$ corresponds to evanescent spectrum.

In the dissipative case, the solution of Eqs. (9) and (13) has a form similar to Eq. (15). Let

$$\phi=\int_0^\infty g(x)\exp(it)J_0(nR)dn+\text{c.c.}, \tag{16}$$

where $g(x)$ is a function to be determined. This approach is found to give the same solution as the frequency domain approach which is described in the Appendix.

Substituting Eq. (16) into Eq. (9) leads to a second-order ordinary differential equation for $g(x)$,

$$g''(x)-\left(n^2-\frac{1}{1+i\zeta}\right)g(x)=0, \tag{17}$$

where each prime denotes differentiation with respect to x.

The general solution of Eq. (17) is

$$g(x)=A\,\exp\left[-\left(\mu_n+\frac{i\zeta}{2\mu_n}\right)x\right]+B\,\exp\left[\left(\mu_n+\frac{i\zeta}{2\mu_n}\right)x\right]. \tag{18}$$

Coefficient B is set to be zero in order to satisfy the radiation condition for increasing positive x. Coefficient A is selected in order to satisfy the boundary condition Eq. (13), which gives

$$A=i\,\frac{\epsilon}{2}\,H_n, \tag{19}$$

where

$$H_n=nF_n/(\mu_n+i\zeta/2\mu_n). \tag{20}$$

In the case $\zeta\ll 1$, H_n is approximately G_n. Therefore, an improved King integral with dissipation is as shown:

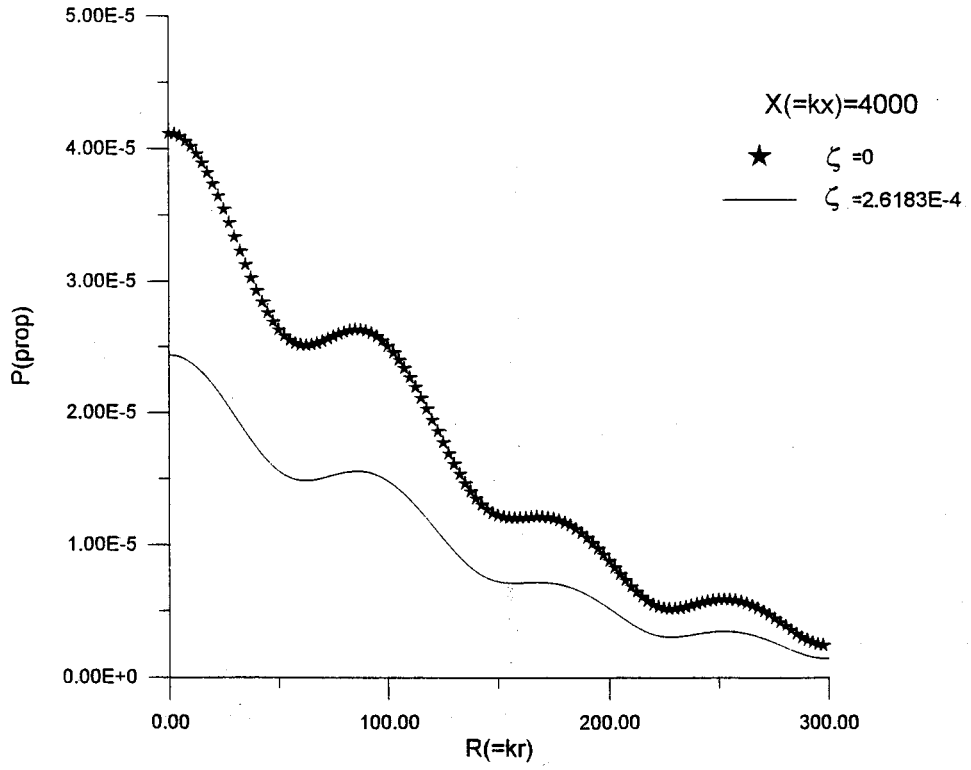


FIG. 2. Variation of pressure amplitude along the transverse direction in the propagation spectrum.

$$\phi = i\frac{\epsilon}{2}\int_0^\infty G_n \exp(it)\exp\left[-\left(\mu_n + \frac{i\zeta}{2\mu_n}\right)x\right]J_0(nR)dn + \text{c.c.} \tag{21}$$

According to pressure and velocity potential relationship, one can easily obtain

$$\frac{P}{\rho_0 c_0^2} = -\frac{\partial\phi}{\partial t} = \frac{\epsilon}{2}\int_0^\infty G_n \exp(it)\exp\left[-\left(\mu_n + \frac{i\zeta}{2\mu_n}\right)x\right] \times J_0(nR)dn + \text{c.c.} \tag{22}$$

Also, the velocity in axial direction can be expressed as

$$\frac{V_x}{c_0} = \frac{\partial\phi}{\partial x} = \frac{\epsilon}{2i}\int_0^\infty nF_n \exp(it)\exp\left[-\left(\mu_n + \frac{i\zeta}{2\mu_n}\right)x\right] \times J_0(nR)dn + \text{c.c.} \tag{23}$$

It is interesting to observe that for $n<1$, where μ_n is imaginary, the thermoviscosity term contributes an exponential decay. On the other hand, if $n>1$, so that μ_n is real, the thermoviscosity term imparts propagation features to the evanescent spectrum. More specifically, modes in the evanescent spectrum now have the appearance of waves that propagate with axial wave number $\zeta/2(n^2-1)^{1/2}$ and whose amplitude decays in the axial direction at rate $(n^2-1)^{1/2}$. This effect is insignificant because of the smallness of ζ, except possible in the vicinity of $n=1$.

IV. NUMERICAL EVALUATION

There are two problems in the evaluation of the integral for pressure. One is the singularity involved at $n=1$, the other is the difficulty to integrate over an infinite domain. To overcome these two difficulties, the integral is carried out by dividing the integral domain into two subdomains. They are the $0<n<1$ domain for propagation spectrum and the $n>1$ domain for evanescent spectrum. Then, in each domain the technique of change variables are performed. Then Eq. (22) can be rewritten as

$$\frac{P}{\rho_0 c_0^2} = \frac{P_{\text{prop}}}{\rho_0 c_0^2} + \frac{P_{\text{ev}}}{\rho_0 c_0^2}, \tag{24}$$

where

$$\frac{P_{\text{prop}}}{\rho_0 c_0^2} = \frac{\epsilon}{2}\int_0^1 \frac{nF_n}{\mu_n}\exp(it)\exp\left[-\left(\mu_n + \frac{i\zeta}{2\mu_n}\right)x\right] \times J_0(nR)dn + \text{c.c.} \tag{25}$$

represents contribution from $0<n<1$ domain, and

$$\frac{P_{\text{ev}}}{\rho_0 c_0^2} = \frac{\epsilon}{2}\int_1^\infty \frac{nF_n}{\mu_n}\exp(it)\exp\left[-\left(\mu_n + \frac{i\zeta}{2\mu_n}\right)x\right] \times J_0(nR)dn + \text{c.c.} \tag{26}$$

represents contribution from $n>1$ domain.

For $0<n<1$ domain, let

$$n = \cos\sigma, \tag{27}$$

then

$$\mu_n = i\sin\sigma, \quad dn = -\sin\sigma d\sigma. \tag{28}$$

Substituting Eqs. (27) and (28) into Eq. (25) leads to

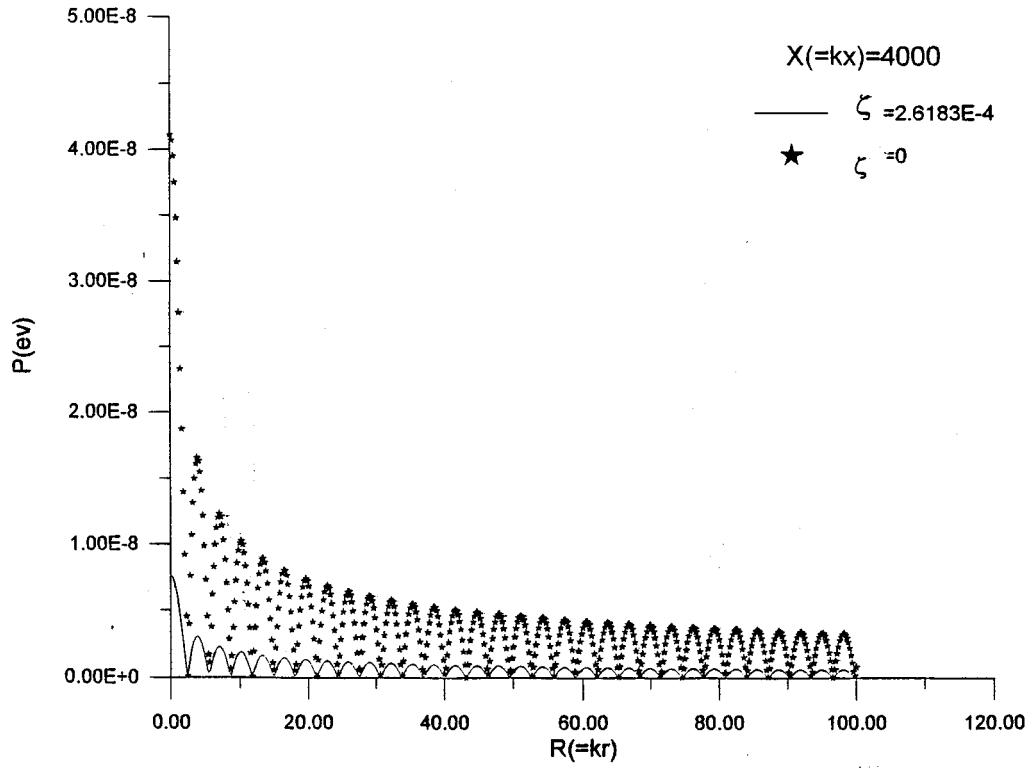


FIG. 3. Variation of the pressure amplitude along the transverse direction in evanescent spectrum.

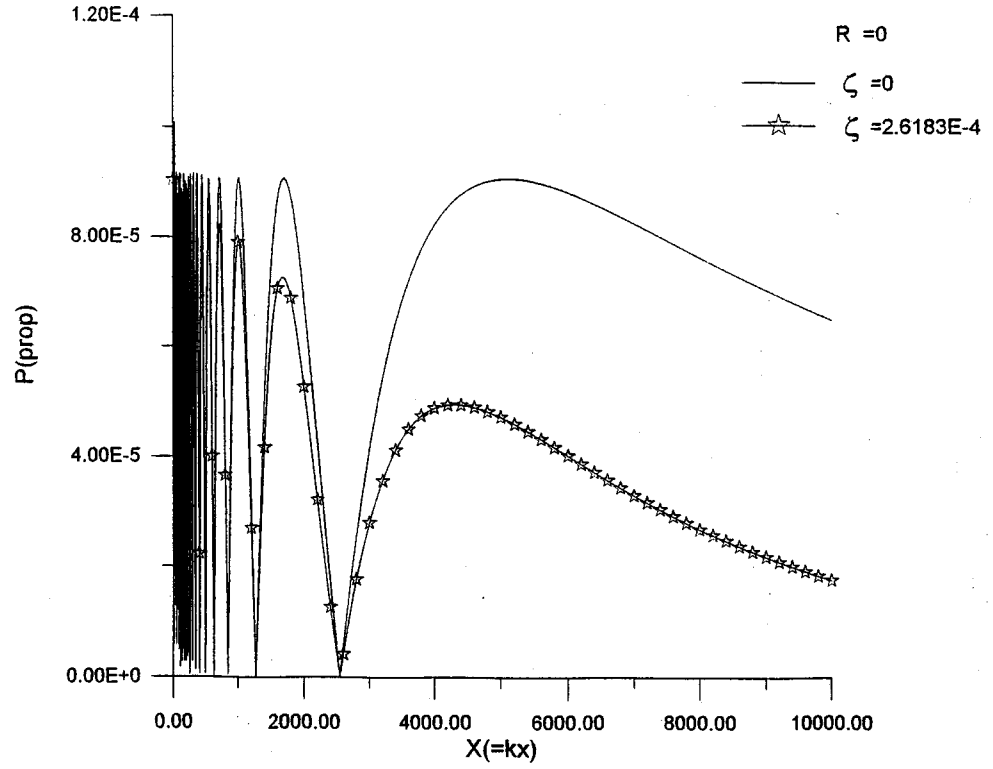


FIG. 4. Variation of the pressure amplitude along the axial direction in the propagation spectrum.

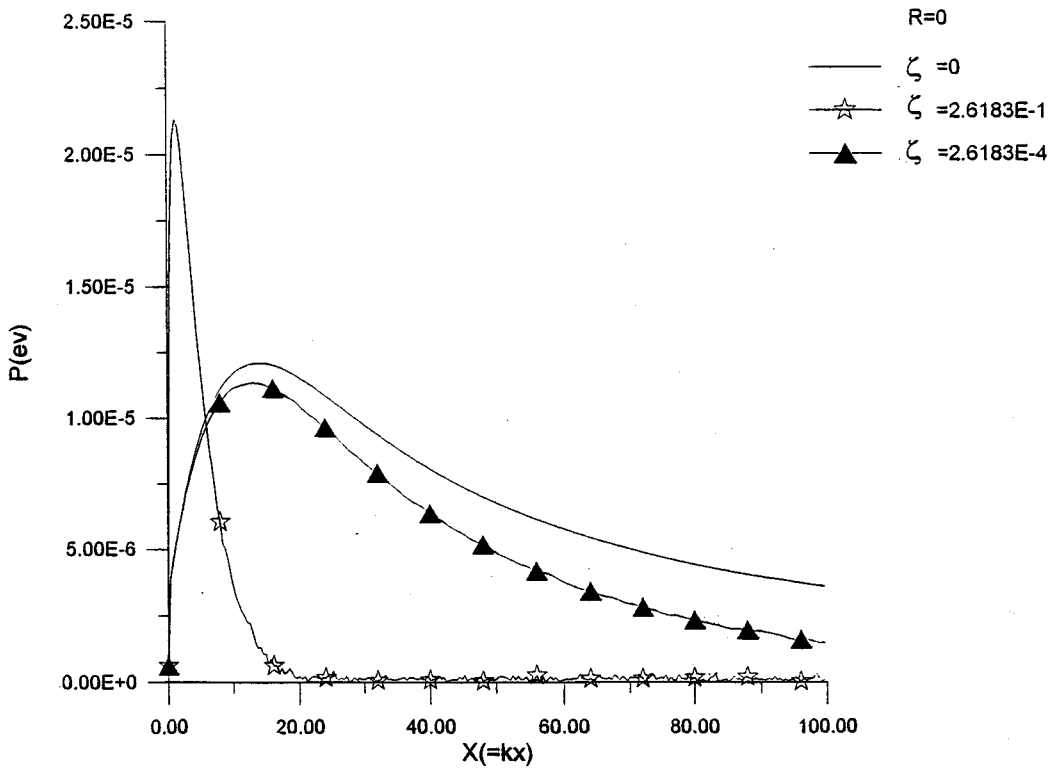


FIG. 5. Variation of the pressure amplitude along the axial direction in the evanescent spectrum.

$$\frac{P_{\text{prop}}}{\rho_0 c_0^2} = \frac{i\epsilon}{2} \int_{\pi/2}^{0} nF_n \exp(it)\exp[-i \sin(\sigma)x] \times \exp\left(-\frac{\zeta}{2 \sin \sigma} x\right) J_0(R \cos \sigma) d\sigma + \text{c.c.} \quad (29)$$

Equation (29) can be evaluated numerically.

For $n>1$ domain, let

$$n = \sec \sigma, \quad (30)$$

then

$$\mu_n = \tan \sigma, \quad dn = -\tan \sigma \sec \sigma d\sigma. \quad (31)$$

Substituting Eqs. (30) and (31) into Eq. (26) leads to

$$\frac{P_{\text{ev}}}{\rho_0 c_0^2} = -\frac{\epsilon}{2} \int_{0}^{\pi/2} nF_n \exp(it)\exp(-i \tan \sigma x) \times \exp\left(i \frac{\zeta}{2 \tan \sigma} x\right) J_0(R \sec \sigma)\sec \sigma d\sigma + \text{c.c.} \quad (32)$$

Equation (32) can be evaluated easily.

In the examples shown in the following figures, the amplitude shading function $f(R)$ is assumed to be uniform which is set to be $f(R)=1$. Then, it can be shown by solving Eq. (14) that

$$nF_n = kaJ_1(nka). \quad (33)$$

The parameters which are used in the following examples are $ka=179$ and acoustic Mach number $\epsilon=4.5286e-5$.

Figure 2 shows the variation of pressure amplitude along the transverse direction at $x=4000$ in the propagation spectrum. In the figure, it shows a stronger effect on the axis than off the axis. Figure 3 shows the variation of pressure along the transverse direction at $x=4000$ in the evanescent spectrum. It also shows a stronger dissipative effect on axis than off axis. From Figs. 2 and 3, it shows the propagation spectrum is dominant in overall pressure value. In addition, the evanescent spectrum shows more variation in the transverse direction. Figure 4 shows the variation of pressure amplitude along the axial direction on axis in the propagation spectrum. It shows that the amplitude decreases in axial direction due to dissipated effects. Figure 5 shows the variation of pressure amplitude in the axial direction on axis in the evanescent spectrum. It is interesting to find that there is a peak value close to the circular piston in the evanescent spectrum. The location of the peak value may change due to the dissipation coefficient. Also, one can see that with a higher dissipation coefficient (such as: $\zeta=2.6183e-1$), it could have higher peak values. It is a valuable phenomena that the higher dissipation is not necessary to reduce the peak pressure amplitude in the evanescent spectrum. In addition, it is found that the evanescent spectrum is important, being very close to the circular piston.

V. CONCLUSION

In the present study, a linear nondimensional wave equation is developed in terms of velocity potential. Then, an improved King integral solution including thermovicous effects is obtained. Particularly, in the solution, the spectrum which is considered to be evanescent in the nondissipated

model when $n>1$, is found to have a propagation mode with axial wave number $\zeta/2(n^2-1)^{1/2}$ due to the dissipation effects. This effect is noticeable in the evanescent spectrum ($n>1$), however, when it is compared to the overall pressure, it is insignificant. In addition, if one can find a medium with a dissipation coefficient $\zeta \approx O(1)$, then, the effect might be significant, especially being very close to the circular piston.

APPENDIX: ALTERNATIVE APPROACH IN THE FREQUENCY DOMAIN

Equation (9) is rewritten as follows:

$$\nabla^2\phi+\zeta\,\frac{\partial}{\partial t}\,(\nabla^2\phi)-\frac{\partial^2\phi}{\partial t^2}=0.$$

Considering a monochromatic source, one lets

$$\phi(x,R,t)=\psi(x,R)e^{it}+\text{c.c.} \tag{A1}$$

Substituting Eq. (A1) into Eq. (9) leads to

$$(1+i\zeta)\nabla^2\psi+\psi=0. \tag{A2}$$

Inverse Hankel transform and Hankel transform of ψ is shown, respectively,

$$\psi(R,x)=\int_0^\infty \tilde{\psi}(n,x)J_0(nR)n\,dn, \tag{A3}$$

$$\tilde{\psi}(n,x)=\int_0^\infty \psi(R,x)J_0(nR)R\,dR. \tag{A4}$$

Equation (A3) is substituted into the Laplace operator in Eq. (A2) which leads to

$$\begin{aligned}\nabla^2\psi&=\frac{\partial^2\psi}{\partial R^2}+\frac{1}{R}\,\frac{\partial\psi}{\partial R}+\frac{\partial^2\psi}{\partial x^2}\\&=\int_0^\infty\tilde{\psi}(n,x)\,\frac{\partial^2 J_0(nR)}{\partial R^2}\,n\,dn\\&\quad+\int_0^\infty\tilde{\psi}(n,x)\,\frac{\partial J_0(nR)}{\partial R}\,n\,dn\\&\quad+\int\,\frac{\partial^2\tilde{\psi}(n,x)}{\partial x^2}\,J_0(nR)n\,dn.\end{aligned} \tag{A5}$$

The identity of the Bessel function is used, which is

$$\frac{\partial^2 J_0(nR)}{\partial R^2}+\frac{1}{R}\,\frac{\partial J_0(nR)}{\partial R}=-n^2J_0(nR). \tag{A6}$$

Therefore Eq. (A5) can be rewritten as

$$\begin{aligned}\nabla^2\psi&=\int_0^\infty(-n^2)\tilde{\psi}(n,x)J_0(nR)n\,dn\\&\quad+\int_0^\infty\,\frac{\partial^2\tilde{\psi}(n,x)}{\partial x^2}\,J_0(nR)n\,dn.\end{aligned} \tag{A7}$$

Substituting Eq. (A7) and Eq. (A3) into Eq. (A2) leads to

$$\begin{aligned}&(1+i\zeta)\left(\int_0^\infty(-n^2)\tilde{\psi}(n,x)J_0(nR)n\,dn\right.\\&\left.\quad+\int_0^\infty\,\frac{\partial^2\tilde{\psi}(n,x)}{\partial x^2}\,J_0(nR)n\,dn\right)\\&\quad+\int_0^\infty\tilde{\psi}(n,x)J_0(nR)n\,dn=0.\end{aligned} \tag{A8}$$

Each term in Eq. (A8) is operated by the Hankel transform and is divided by $(1+i\zeta)$. Then Eq. (A8) becomes

$$\frac{\partial^2\tilde{\psi}(n,x)}{\partial^2 x^2}-\left(n^2-\frac{1}{1+i\zeta}\right)\tilde{\psi}(n,x)=0, \tag{A9}$$

where ''n'' respresents an individual nondimensional transverse wave number which is considered as a constant for each transverse wave. Therefore, Eq. (A9) is accurately an ordinary differential equation. Also, Eq. (A9) has the same form as Eq. (17) except $\tilde{\psi}(n,x)$ instead of $g(x)$. This proves that the two approaches give the same results.

[1] L. V. King, ''On the acoustic radiation field of the piezo-electric oscillator and effect of viscosity on transmission,'' Can. J. Res. **11**, 135 (1934).
[2] K. E. Froysa, J. N. Tjotta, and S. Tjotta, ''Linear propagation of a pulsed sound beam from a plane or focusing source,''J. Acoust. Soc. Am. **93**, 80–92 (1993).
[3] D. T. Blackstock, ''On plane, spherical, and cylindrical sound waves of finite amplitude in lossless fluids,''J. Acoust. Soc. Am. **36**, 217–219 (1964).
[4] D. T. Blackstock, ''Thermoviscous attenuation of plane, periodic, finite-amplitude sound waves,''J. Acoust. Soc. Am. **36**, 534–542 (1964).
[5] G. R. Harris, ''Review of transient field theory for a baffled planar piston,''J. Acoust. Soc. Am. **70**, 10–20 (1981).
[6] J. N. Tjotta and S. Tjotta, ''An analytical model for the nearfield of a baffled piston transducer,''J. Acoust. Soc. Am. **68**, 334–339 (1980).
[7] G. S. Garrett, J. N. Tjotta, and S. Tjotta, ''Nearfield of a large acoustic transducer. Part III: general results,''J. Acoust. Soc. Am. **75**, 769–779 (1984).
[8] S. Goldstein, *Lectures in Fluid Mechanics* (Interscience, London, 1957), Vol. II, Chap. 4.
[9] O. V. Rudenko and S. I. Soluyan, *Theoretical Foundations of Nonlinear Acoustics*, translated by R. T. Beyer (Plenum, New York, 1977).
[10] J. Naze Tjotta and S. Tjotta, ''Nonlinear equations of Acoustics,'' *Frontiers of Nonlinear Acoustics* (1990), pp. 80–97.
[11] A. D. Pierce, *Acoustics—An Introduction to Its Physical Principles and Application* (McGraw-Hill, New York, 1981), Chap. 10.

Optimal selection of parameters for the angular spectrum approach to numerically evaluate acoustic fields

Ping Wu, Rymantas Kazys, and Tadeusz Stepinski
Uppsala University, Department of Technology, Circuits and Systems, Box 534, S-751 21 Uppsala, Sweden

(Received 24 January 1996; revised 30 July 1996; accepted 13 August 1996)

Optimal selection of parameters is presented for the angular spectrum approach (ASA) to the numerical calculation of acoustic fields radiated by planar transducers and linear arrays with and without focusing. The parameters include spatial sampling interval, discretization size of a source plane in which a source is located, and the angular range over which the plane waves decomposed from the source using the discrete Fourier transform (DFT) are chosen to superimpose and construct the fields. The concept of instantaneous frequency is applied to the Fourier transformed Green's function to determine the angular range of the plane waves used for the construction with minimal spatial aliasing error. Based on the minimization of spatial frequency and spatial aliasing errors in a constructed field, optimal selection of the parameters is worked out. The ASA with the optimal selection is then applied to computing the fields radiated by planar transducers and linear arrays into water and a layered (oil/water) medium. The results demonstrate that the methods proposed in this paper to select the parameters for the ASA allow the aliasing errors to effectively eliminate and thus yield the best computational accuracy for the parameters selected. [S0001-4966(97)00501-8]

PACS numbers: 43.20.Rz, 43.20.Bi [JEG]

INTRODUCTION

The angular spectrum approach (ASA) has been widely used to calculate acoustical fields because it can be implemented by taking advantage of a powerful numerical computation tool—the fast Fourier Transform (FFT) algorithm,[1–7] or the discrete Hankel transform.[8,9] It is well known, nevertheless, that the numerical implementation of the ASA results in several types of errors. Two major ones are spatial frequency and spatial aliasing errors. They have been investigated by a few researchers. The effect of the spatial frequency aliasing has been studied by the authors and this type of aliasing error can be completely eliminated in the case of planar sources by using the sinc-Fourier transform introduced in Ref. 1, even if the spatial sampling frequency does not meet the requirement of the Nyquist criterion. The spatial aliasing is another type of error which seriously affects the computation accuracy, and it was coped with more than ten years ago by Williams and Maynard[5] and recently by Orofino and Pedersen[3,4] and Christopher and Parker.[8] Most recently, the authors[2] have given a thorough investigation to this spatial aliasing error based on the evaluation of 2-D fields.

The spatial aliasing mainly results from the undersampling of the Fourier transformed Green's function in the spatial frequency domain.[2,5] Although it can be reduced by use of a large discretization size of a source plane, the spatial aliasing cannot be avoided as long as all the decomposed plane waves are used to construct a field. The reason for this is that, if the Fourier transformed Green's function is viewed as the ''signal'' in the spatial frequency domain, the ''instantaneous frequency'' of the ''signal'' can be as large as infinite.[2] When the discretization size is fixed, however, the spatial aliasing error has been found to be sensitive to the angular range of the angular spectrum over which the decomposed plane waves are used for constructing the field.[2–4,8] The investigation in Ref. 2 revealed that the smaller the angular range used, the smaller the spatial aliasing error. Nevertheless, the larger angular range can result in a better convergence of the ASA's solution to the exact solution, provided that no spatial aliasing error exists. Spatial sampling frequency determines not only the spatial sampling interval but also the spatial frequency range, or the angular range, of the angular spectrum. Although the aliasing error in the spectrum can be eliminated with the sinc-Fourier transform, the high spatial sampling frequency for the enough large angular range is necessary to guarantee the good convergence of the ASA's solution. In the case of the 2-D DFT, as can be easily imagined, simultaneous use of high sampling frequency and large discretization size results in tremendously increasing the size of the 2-D matrix used for performing the DFT. Very often the matrix size is constrained by the size of a computer memory. Meanwhile, the computation obviously becomes more time consuming as the 2-D matrix increases in size. This implies that these two parameters may not be large enough simultaneously to eliminate the aliasing errors in the computation, and thus the calculated results may unavoidably be distorted if all the decomposed plane waves are used to construct a field. The extent of the distortion is very sensitive to the spatial sampling interval and the discretization size and, when they are fixed, to the angular range used to construct the field.[2] The main goal of this paper is to present the way to determine the parameters for the ASA to reach the best computation accuracy.

For this purpose, the methods used for the analysis of the ASA in the case of 2-D fields[2] are extended to the present case of 3-D fields. Based on the investigation of the characteristics of the ASA as a function of the angular range, optimal selection of parameters for the ASA to calculating 3-D fields is fulfilled. The optimization is applied to the ASA for

the calculations of acoustic fields radiated from rectangular planar sources and linear arrays into homogeneous and layered media.

I. ANGULAR SPECTRUM APPROACH AND ITS NUMERICAL IMPLEMENTATION

An acoustical (pressure) field radiated by a planar piston source or an array of planar piston source elements in an infinite baffle with time-harmonic excitation $e^{-j\omega t}$($e^{-j\omega t}$ is dropped in sequel) can be calculated according to the integral,[5,10,11]

$$p(x,y,z) = -jk\rho c \int\int_S v(x',y',0)\frac{e^{jkR}}{2\pi R}dx'dy', \qquad (1)$$

where $v(x,y,0)$ is the normal velocity distribution on the source plane at $z=0$, S is the area of the source, and $R=\sqrt{(x-x')^2+(y-y')^2+z^2}$. This is the well-known Rayleigh integral. Application of the convolution theorem to the Rayleigh integral yields the representation of the pressure field as an angular spectrum of plane waves,[5,12]

$$\begin{aligned} p(x,y,z) &= \frac{k\rho c}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} V(k_x,k_y) \\ &\quad \times \frac{\exp[j(xk_x+yk_y+zk_z)]}{k_z}dk_xdk_y \\ &\qquad\qquad \text{for } z \geqslant 0, \\ &= \frac{k^2\rho c}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} V(n_x,n_y) \\ &\quad \times \frac{\exp[jk(xn_x+yn_y+zn_z)]}{n_z}dn_xdn_y, \qquad (2) \end{aligned}$$

where k is the wave number, $V(k_x,k_y)$ is the 2-D Fourier transform of $v(x,y,0)$, and

$$\begin{aligned} k_z &= \sqrt{k^2-k_x^2-k_y^2} = k\sqrt{1-n_x^2-n_y^2}, \quad k^2 \geqslant k_x^2+k_y^2, \\ &= j\sqrt{k_x^2+k_y^2-k^2} = jk\sqrt{n_x^2+n_y^2-1}, \quad k^2 < k_x^2+k_y^2. \end{aligned}$$

Here $n_x = k_x/k = \cos\theta_x$, $n_y = k_y/k = \cos\theta_y$, and $n_z = \sqrt{1-n_x^2-n_y^2} = \cos\theta_z$. The direction angles of the decomposed wave components, θ_x, θ_y, and θ_z, form the orientation of the wave vector $\mathbf{k}$ with respect to the x,y, and z coordinate axes. The direction cosines n_x, n_y, and n_z are related to the wave vector in a manner $\mathbf{k} = k(n_x, n_y, n_z)$. It is obvious that there exists a pole at $k_z = 0$ (or $n_z = 0$) in Eq. (2). When the integrand is sampled for the DFT, therefore, the resultant sum becomes infinite whenever the values of any sample points give $k_z = 0$.

Numerical implementation of the ASA is conducted in the following way:

(1) Select a source plane which is a square with a length D on each side, which is several times larger than the maximal dimension of the source, and in the middle of the selected plane the source is located (Fig. 1).

(2) Determine a spatial sampling frequency F_s or its reciprocal, $\Delta d = 1/F_s$ (the sampling interval) for the proper sampling of the source plane having the normal velocity distribution $v(x,y,0)$. The source plane is discretized into a 2-D $N \times N$ matrix, where $N = D/\Delta d$ is the number of the samples along each side of the plane. The number of the samples for the planar source with the dimension $2A \times 2B$ is assumed to be $L_x \times L_y$, where $L_x = 2A/\Delta d$ and $L_y = 2B/\Delta d$. (see Fig. 1). When L_x and L_y both are odd, the source center is at the origin of the coordinates. When either L_x or L_y or both are even, the center of source is at $(\Delta d/2, 0)$, or $(0, \Delta d/2)$, or $(\Delta d/2, \Delta d/2)$, respectively, which introduces the half sample length phase shift error in the angular spectrum.[1] Correction of the phase shift error will be taken into account in the numerical implementation.

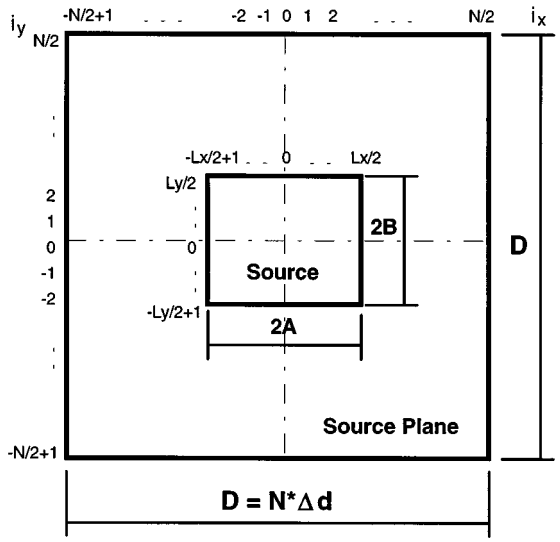


FIG. 1. Geometry and notation used for the numerical calculation of angular spectra and acoustic fields.

(3) Perform the DFT of $v(i_x\Delta d, i_y\Delta d, 0)$, where $i_x, i_y = -N/2+1, \ldots, N/2$ and obtain the discrete angular spectrum $V_D(m_x\Delta f, m_y\Delta f)$, where $m_x, m_y = -N/2+1, \ldots, N/2$ and Δf is the sampling interval in the spatial frequency domain, which is equal to $1/D$ or $1/N\Delta d$.

(4) Make a complex summation of the decomposed $N \times N$ plane waves at point (x,y,z) with the amplitudes $V_D(m_x\Delta f, m_y\Delta f)$ and the propagation directions $(n_x, n_y, n_z) = (\lambda m_x\Delta f, \lambda m_y\Delta f, \sqrt{1-(\lambda m_x\Delta f)^2-(\lambda m_y\Delta f)^2}$, $m_x, m_y = -N/2+1, \ldots, N/2$, and finally get the acoustical field $p_D(x,y,z)$ radiated by the source.

With introducing the discrete sinc-Fourier transform (DSFT) and the phase correction factor to get rid of the spatial frequency aliasing and phase shift errors,[1] the above described procedure for the numerical implementation can be formulated as

$$\begin{aligned} &V_D(m_x\Delta f, m_y\Delta f) \\ &= \operatorname{sinc}\left(\frac{\pi m_x}{N}, \frac{\pi m_y}{N}\right)\exp\left(\frac{j2\pi(m_x\phi_x+m_y\phi_y)}{N}\right) \\ &\quad \times \Delta d^2 \sum_{i_x=-N/2+1}^{N/2}\ \sum_{i_y=-N/2+1}^{N/2} v(i_x\Delta d, i_y\Delta d, 0) \\ &\quad \times \exp\left[\frac{-j2\pi(m_xi_x+m_yi_y)}{N}\right], \qquad (3) \end{aligned}$$

$$p_D(i_x\Delta d, i_y\Delta d, z)$$
$$= \frac{\Delta f^2 \rho c}{\lambda} \sum_{m_x=-N/2+1}^{N/2} \sum_{m_y=-N/2+1}^{N/2} \frac{V_D(m_x\Delta f, m_y\Delta f)}{\sqrt{1/\lambda^2-(m_x\Delta f)^2-(m_y\Delta f)^2}}$$
$$\times \exp\left[j2\pi \left(\frac{i_x m_x + i_y m_y}{N} + z\sqrt{1/\lambda^2-(m_x\Delta f)^2-(m_y\Delta f)^2} \right) \right], \quad (4)$$

where $\mathrm{sinc}(x,y)=(\sin x/x)(\sin y/y)$ is the 2-D sinc function and ϕ_x and ϕ_y are the phase correction factors, which are given by

$$\phi_x = \begin{cases} 1/2 & \text{when } L_x \text{ is even,} \\ 0 & \text{when } L_x \text{ is odd,} \end{cases}$$

and

$$\phi_y = \begin{cases} 1/2 & \text{when } L_y \text{ is even,} \\ 0 & \text{when } L_y \text{ is odd.} \end{cases}$$

Obviously, the angular spectrum in Eq. (3) can be obtained by multiplying the FFT of $v(i_x\Delta d, i_y\Delta d, 0)$ with a 2-D sinc function and a phase correction term.

II. CHARACTERISTICS OF THE ASA AS A FUNCTION OF ANGULAR RANGE

Taking advantage of the DSFT algorithm, we can eliminate the spatial frequency aliasing error and obtain a nonaliasing angular spectrum. The remaining error is mainly caused by the spatial aliasing. In order to make clear why and how the spatial aliasing error occurs in the numerical computation and to show the dependence of the error on the parameters used, we shall study Eq. (2) and its numerical implementation, Eq. (4). The approach used in Ref. 2 to deal with the problem in the 2-D case is extensively employed here in the present 3-D case.

Transforming the rectangular coordinates to the cylindrical ones in the double integral in Eq. (2) and using variable integration bounds, we get

$$p(x,y,z;\eta) = \frac{k^2\rho c}{(2\pi)^2} \int_0^{\eta} \int_0^{2\pi} V(n_\rho, \varphi) \times \frac{\exp[jk(xn_\rho \cos\varphi + yn_\rho \sin\varphi + z\sqrt{1-n_\rho^2})]}{\sqrt{1-n_\rho^2}} \times n_\rho \, d\varphi \, dn_\rho, \quad (5)$$

where $\eta \geqslant 0$, $n_\rho = \sqrt{n_x^2+n_y^2}$, and $\varphi = \arctan(n_y/n_x)$. It can be seen from Eq. (5) that, when the source is given, the field $p(x,y,z)$ at point (x,y,z), superimposed from decomposed plane waves, depends only on the variable η. The η radially determines the angular range over which the decomposed plane waves are used for constructing the field, and ranges from zero to infinity. Apparently, when $\eta \leqslant 1$, $n_z = \sqrt{1-n_\rho^2}$ is real and $p(x,y,z;\eta)$ only contains homogeneous plane waves with wave vector $\boldsymbol{k} = k(n_x, n_y, n_z)$. When $\eta > 1$, the exponential in the integral becomes $\exp(-k|n_z|z)$ since the n_z is imaginary for $n_\rho > 1$ and evanescent plane waves are present. How the $p(x,y,z;\eta)$ varies with η is the so-called characteristics of the ASA as a function of angular range. Here we observe the characteristics by taking advantage of the DFT for the numerical implementation discussed in the preceding section. For a given spatial sampling frequency F_s which gives the spatial frequency range of an angular spectrum, $-F_s/2 < f_x, f_y \leqslant F_s/2$, and according to the relations $n_x = \lambda f_x$, $n_y = \lambda f_y$, the direction cosines n_x and n_y run over a square region $[-\lambda F_s/2 \leqslant n_x, n_y \leqslant \lambda F_s/2]$, but $n_\rho = \sqrt{n_x^2+n_y^2}$ ranges from 0 to $\eta_{\max} = \sqrt{2}\lambda F_s/2$ (see Fig. 2). The equivalent of Eq. (5) in the rectangular coordinates can be expressed as

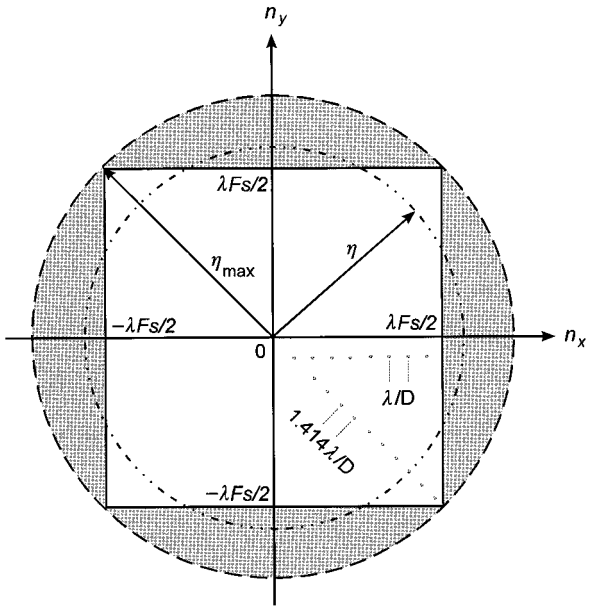


FIG. 2. Geometry and notation for investigating the characteristics of the DFT-based ASA as a function of angular range by using Eqs. (5′) and (6).

$$p(x,y,z;\eta) = \frac{k^2\rho c}{(2\pi)^2} \int_{-\eta}^{\eta} \left(\int_{-\sqrt{\eta^2-n_x^2}}^{\sqrt{\eta^2-n_x^2}} V(n_x, n_y) \times \frac{\exp[jk(xn_x + yn_y + z\sqrt{1-n_x^2-n_y^2})]}{\sqrt{1-n_x^2-n_y^2}} dn_y \right) \times dn_x, \quad (5')$$

where $0 \leqslant \eta \leqslant \eta_{\max}$. The area over which the integration is performed with respect to n_x and n_y in the spectrum domain needs to be specially specified. Referring to Fig. 2, for $\eta \leqslant \lambda F_s/2$, the area is within and on the circle with radius η, and for $\lambda F_s/2 < \eta \leqslant \sqrt{2}\lambda F_s/2$ the area is the one within and on the circle with radius η excluding the shaded part outside the square since $V(n_x, n_y)$ are not defined in the shaded part.

The numerical implementation of Eq. (5′), accordingly, can be written as

$$p_D(i_x\Delta d, i_y\Delta d, z; \eta) = \frac{\Delta f^2 \rho c}{\lambda} \sum_{m_x=-M}^{M} \sum_{m_x=-\mathrm{int}[\sqrt{M^2-m_x^2}]}^{\mathrm{int}[\sqrt{M^2-m_x^2}]} (\cdot), \quad (6)$$

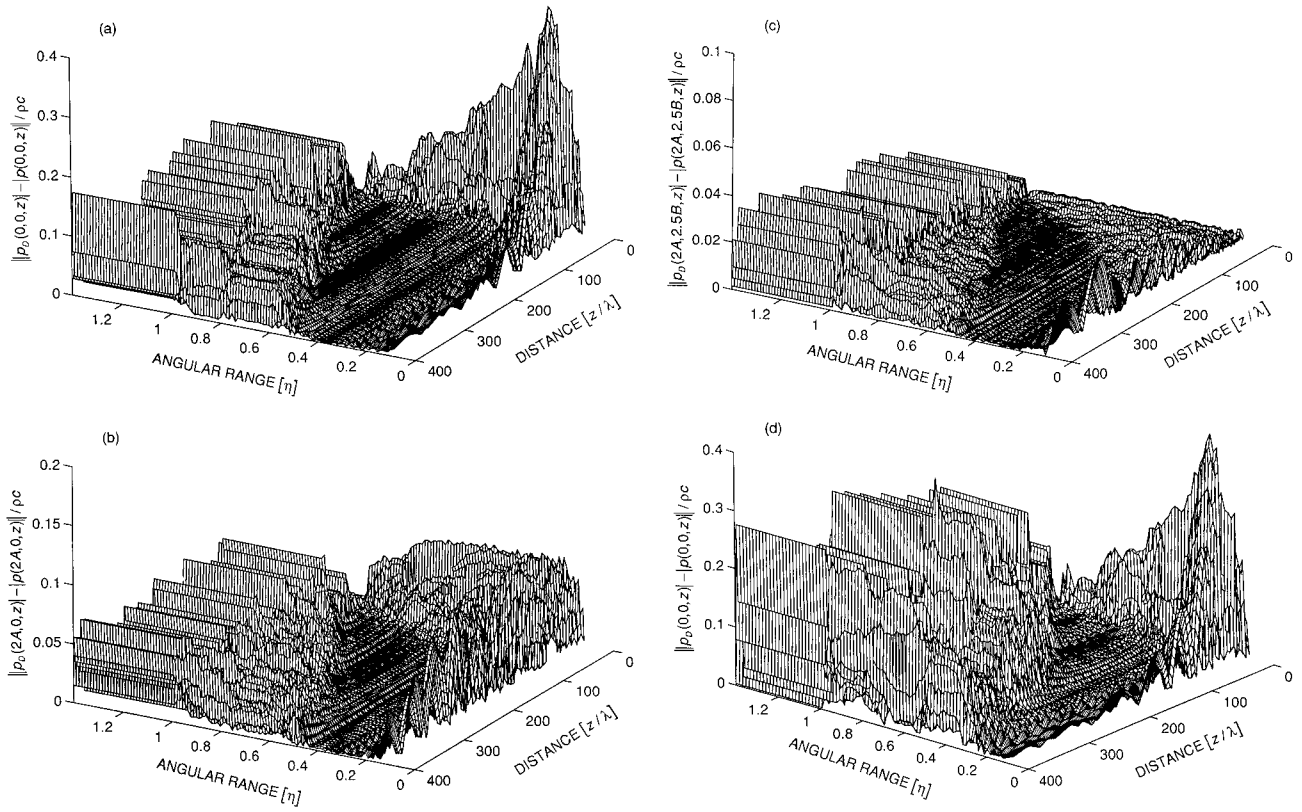


FIG. 3. Difference (absolute) between the ASAs and the analytical solution's pressure magnitudes plotted as a function of angular range. The transducer radiating the fields has the size of $2A=32\lambda$ and $2B=24\lambda$. The sampling frequency $F_s=2/\lambda$ and two different numbers of samples $N=512$ and $N=256$ are used; (a) $N=512$, $x/A=y/B=0$; (b) $N=512$, $x/A=2$, $y/B=0$; (c) $N=512$, $x/A=2$, $y/B=2.5$; and (d) $N=256$, $x/A=y/B=0$. The magnitudes are normalized by ρc. Note that the full scales of the vertical axes in (a), (b), and (c) are different.

where $M=0,1,2,\ldots,\text{int}[\sqrt{2}N/2]$ (int$[\cdot]$ means taking the integer part of) and $\eta=MF_s\lambda/N$, and $(\cdot)$ represents the components of the sum in Eq. (4). Similarly, when $m_x,m_y\leqslant -N/2$ or $m_x,m_y>N/2$, $V_D(m_x\Delta f,m_y\Delta f)$ are not defined in the $N\times N$ matrix and therefore excluded in the calculation of Eq. (6). When $F_s\lambda$ is such an integer that makes $M=N/F_s\lambda$ be an integer, η becomes unity and hence the components of the sum for $\sqrt{m_x^2+m_y^2}=M$ in Eq. (6) become infinite. To eliminate the problem, we subtract a tiny valued number from the medium's sound speed to make $F_s\lambda$ not be an integer without loss of accuracy.

Now we investigate the characteristics of the ASA as a function of angular range by considering a specific case of a 3 MHz rectangular piston transducer with dimension $2A=32\lambda$ and $2B=24\lambda$ (see Fig. 1). The normal velocity on the source surface is assumed to be unity. The medium in which the waves are radiated is water having sound speed $c=1500$ m/s, density $\rho=10^3$ kg/m^3, and no attenuation to be included. Since $\lambda=0.5$ mm and $F_s\lambda$ can be an integer, the sound speed c used is subtracted by 10^{-4} m/s in computation. All the results concerning the pressure fields are normalized by ρc. The investigation of the characteristics is based on the analysis of the difference between the ASAs and the analytical (exact) solutions. The analytical solution is obtained using the method developed by Ocheltree.[10]

To investigate the characteristics of the ASA as a function of angular range, and also to see how it is affected by the source plane size D and the sampling frequency F_s and varies with the field points, two different conditions are used, one where $F_s=2/\lambda$ and the number of samples $N=512(D=256\lambda)$ and the other where $F_s=2/\lambda$ and $N=256(D=128\lambda)$, and the fields to be calculated are on the z axis and off the z axis at $x/A=2$, $y/B=0$ and $x/A=2$, $y/B=2.5$ with the dimension of $2\lambda\sim400\lambda$ and the calculation interval of 2λ. Choosing $F_s=2/\lambda$ allows to be available all homogeneous plane waves whose angular range is $0\leqslant\eta\leqslant1$, but, meanwhile, yields some inhomogeneous plane waves since η runs till $\sqrt{2}$, which is different from the 2-D case.[1,2] The changing of N is to observe the effect of D on the spatial aliasing. Since the source is planar and no beam steering is involved, the symmetry of the angular spectrum is employed. It should be noted that, due to the use of the symmetry, $V_D(0,0)$ must be divided by four and both $V_D(0,m_y\Delta f)$ and $V_D(m_x\Delta f,0)$ by two, and the indices of the sum in Eq. (6) become $m_x=0,1,\ldots,M$ and $m_y=0,1,\ldots,\text{int}[\sqrt{M^2-m_x^2}]$. The results calculated from the above conditions for four cases are shown in Fig. 3 in which the absolute values of the differences are taken. Comparison of the results at different field points of interest for $N=512$ and $N=256$ is illustrated in Fig. 4.

Figure 3(a)–(d) shows the same phenomenon as in the 2-D case studied in Ref. 2 that there exists a deep, flat ''valley'' between ''two mountain chains'' in each subfigure. A detailed analysis of this phenomenon has been presented in Ref. 2. Here we just introduce some definitions and obtain some conclusions. This flat ''valley'' is termed as good ac-

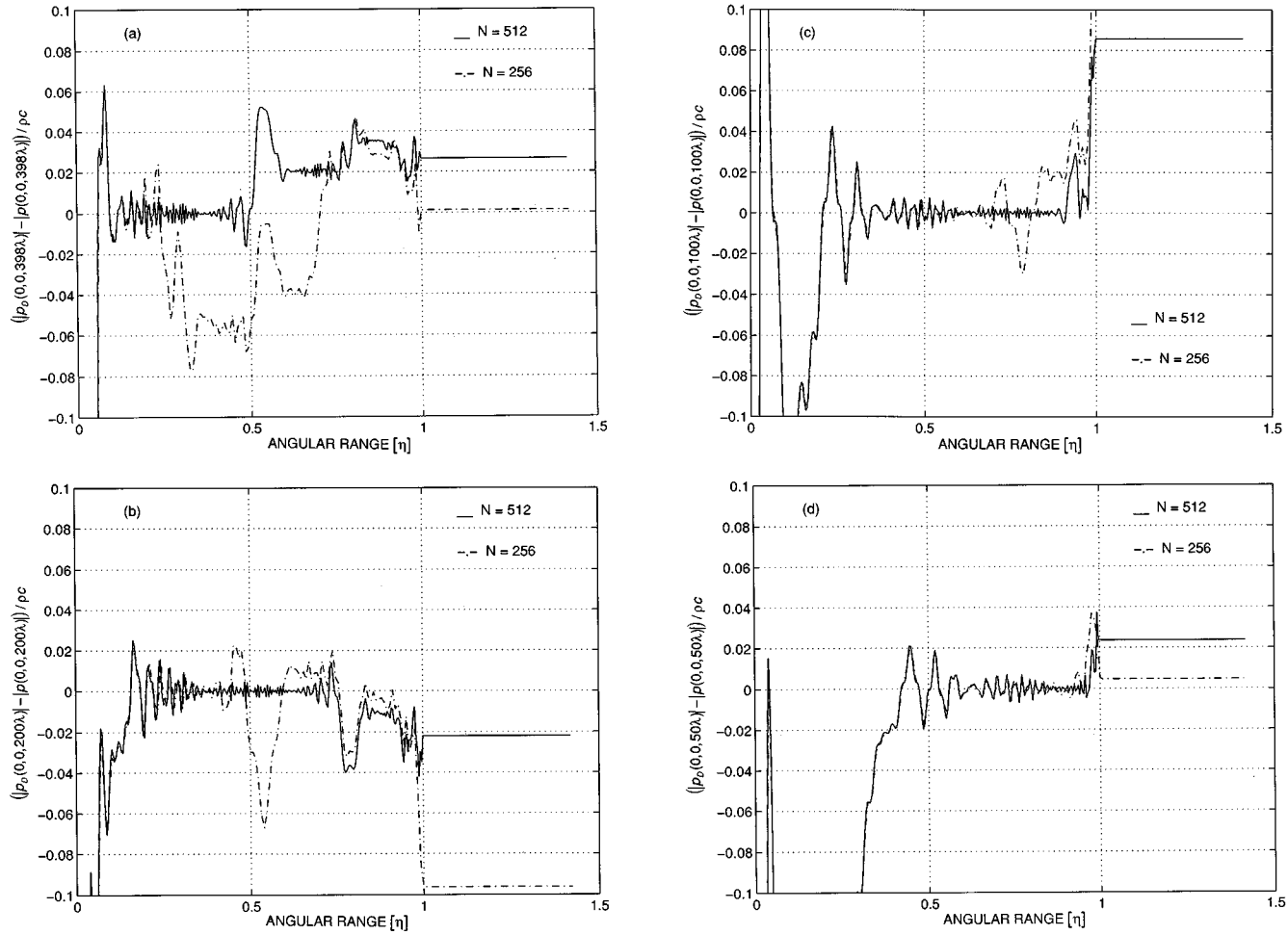


FIG. 4. Difference between the ASAs and the analytical solution's pressure magnitudes plotted as a function of angular range in two cases, corresponding to Fig. 3(a) and (d), where the numbers of samples $N=512$ (solid) and $N=256$ (dashed and dotted), respectively, with the same sampling frequency $F_s=2/\lambda$; (a), (b), (c), and (d) show the difference at different points of interest on the z axis, $z=398\lambda$, 200λ, 100λ, and 50λ, respectively. The magnitudes are normalised by ρc. Note the difference between the ASAs and the analytical solution's pressure magnitudes is not absolute value.

curacy region. The ''mountain'' located in the large angular range is the aliasing region (which will be explained later) and in the small angular range is the pre-well-converging region. The boundary between the aliasing and the good accuracy regions is called upper boundary, and the other between the pre-well-converging and the good accuracy regions is called lower boundary. Comparing Fig. 3(a) with Fig. 3(d), we can see that the size of the good accuracy region is largely decreased when D is decreased from $256\lambda(N=512)$ to $128\lambda(N=256)$ and hence the aliasing region pushed downward. Comparison of Fig. 3(a) with Fig. 3(b) and (c) shows slight influence of the coordinate variation on the size of the good accuracy region and a large influence on the pre-well-converging region (in which the ''mountain'' height largely vary with the coordinates in reference to the ''mountain'' height in the aliasing region). Figure 4 shows a comparison of the results of the difference of the ASAs and the analytical solutions at different positions on the z axis in the two cases where $N=512$ and $N=256$. The comparison reveals that there is an angular range [e.g., $\eta=\{0\sim0.2\}$ in Fig. 4(a), $\eta=\{0\sim0.3\}$ in Fig. 4(b) etc.] over which the results for $N=512$ and $N=256$ are in a good agreement and tend to converge to zero as η increases, and beyond which the results for $N=256$ first go into the aliasing region and start to diverge and later those for $N=512$ follow. This is because the smaller D for $N=256$ makes the spatial aliasing come up earlier than in the case of $N=512$. In Figs. 3 and 4, it can also be discovered that, in the range $1<\eta\leqslant1.414$ over which evanescent waves appear, the difference of the ASAs and the analytical solutions keeps constant. This indicates that the evanescent waves do not contribute to the constructed fields except in the vicinity of the source.

According to the above comparison and analysis, it is possible to conclude that the sampling frequency F_s dominates the convergence of the ASA, and choosing it to be $2/\lambda$ can ensure the convergence as good as required for any high precision provided that D is infinite so that no spatial aliasing appears (unless the extremely near field is considered and the radiation is considered from a source with a sharply curved radiation surface). The discretization size D determines the sampling interval ($\Delta f=1/D$) in the spatial frequency domain and dominates the aliasing region, i.e., the larger the discretization size, the smaller the aliasing region and the larger the good accuracy region.

It is known that the spatial aliasing results from the un-

dersampling of the Fourier transformed Green's function.[2] The transformed Green's function in the cylindrical coordinates is $\exp(jkz\sqrt{1-n_\rho^2})/k\sqrt{1-n_\rho^2}$, which is exactly the same in form as in the 2-D case except that $n_\rho=\sqrt{n_x^2+n_y^2}$. The transformed Green's function can be considered as a ''signal'' of n_ρ in the spatial frequency domain, and its ''instantaneous frequency'' can be found to be[2]

$$f_I(n_\rho)=\frac{1}{2\pi}\frac{d}{dn_\rho}(kz\sqrt{1-n_\rho^2})=\frac{-kzn_\rho}{2\pi\sqrt{1-n_\rho^2}}. \tag{7}$$

From Eq. (7), $|f_I(n_\rho)|\to\infty$ when $n_\rho\to 1$ and it is proportional to z. This indicates that the ''instantaneous frequency'' of the transformed Green's function can be infinite and hence the function cannot be sampled enough for any finite D as $n_\rho\to 1$ and as a result, the spatial aliasing is unavoidable. Fortunately, before the aliasing comes up, as has been seen in Figs. 3 and 4, the integral in Eq. (5) can be well convergent when the angular range η is large enough but less than unity. For this reason, we can choose the wave components within an adequate angular range to get a good accuracy rather than use all decomposed components.

To acquire the adequate angular range, we shall gain more insight into the numerical situation for Eq. (7). In the rectangular coordinates the sampling interval in the angular spectrum is identical to $\Delta f=1/D$ all over the $N\times N$ matrix, but looking at the interval Δf in the cylindrical coordinate system, the interval is not identical, e.g., on the x or y axes $\Delta f=1/D$ and diagonally Δf is maximal and equal to $\sqrt{2}/D\approx 1.414/D$ (refer to Fig. 2), and thus the sampling frequency in the spectral domain, F_f, is not radially constant and the minimal sampling frequency, $F_{f\min}=D/\sqrt{2}$ corresponding to the maximal interval. In order to eliminate the spatial aliasing, letting Eq. (7) meet the Nyquist criterion for all different intervals (in the sense of the cylindrical coordinates), i.e., $zn_\rho/\sqrt{1-n_\rho^2}\leqslant F_{f\min}/2$. Thus, we obtain a transit boundary below which the angular range chosen does not cause the aliasing error:

$$n_{na}=\frac{D}{2\sqrt{2}z}\Bigg/\sqrt{1+\left(\frac{D}{2\sqrt{2}z}\right)^2}=\sin\left(\arctan\frac{D}{2\sqrt{2}z}\right). \tag{8}$$

When η becomes larger than n_{na}, the sampling of the transformed Green's function does not meet the Nyquist criterion and the undersampling happens to the function. Thus, Eq. (6) is mingled with the aliasing. However, the aliasing error is not serious when η increases just over n_{na}. Letting η continue to increase until $zn_\rho/\sqrt{1-n_\rho^2}=F_{f\min}$, we get

$$n_{\rm up}=\frac{D}{\sqrt{2}z}\Bigg/\sqrt{1+\left(\frac{D}{\sqrt{2}z}\right)^2}=\sin\left(\arctan\frac{D}{\sqrt{2}z}\right), \tag{9}$$

which will be applied later to some specific examples and shown to be able to determine the above defined upper boundary of the good accuracy region.

Dealing with the lower boundary, here we adopt the empirical method used in Ref. 2. The existence of the good accuracy regions in Figs. 3 and 4 indicates that for a certain discretization size used, $p_D(i_x\Delta d,i_y\Delta d,z;\eta)$ at some point of interest can be convergent to a good enough precision for a certain range of η. Comparing the results in Figs. 3 and 4 with those in the 2-D case reported in Ref. 2, we may find quite similar convergence in both cases. Hence, the formula presented in Ref. 2 determining the lower boundary is directly applied here, written as

$$n_{\rm low}=\begin{cases}(16\lambda/z)^{1/2}, & \text{for } z\geqslant 16\lambda,\\ 1, & \text{for } z<16\lambda.\end{cases} \tag{10}$$

Applying Eqs. (8)–(10) to the case presented in Fig. 4(a)–(d), we obtain the nonaliasing angular range and the locations of the upper and lower boundaries for the case, which are listed in Table I. Applying the calculated $n_{\rm low}$ and $n_{\rm up}$ in Table I to Fig. 4(a)–(d), we can see that the errors in the regions between $n_{\rm low}$ and $n_{\rm up}$ are much smaller than those outside the regions. In the case of $N=256$ and $z=398\lambda$, i.e., the dashed and dotted curve in Fig. 4(a), the upper boundary $n_{\rm up}=0.22$ is very close to the lower boundary $n_{\rm low}=0.20$ and $n_{na}=0.11<n_{\rm low}=0.20$ which does not satisfy the inequality $n_{na}>n_{\rm low}$, and the error between $n_{\rm low}$ and $n_{\rm up}$ is large compared to the other cases in the figure. This means that $D=128\lambda$ is not large enough to make the inequality $n_{na}>n_{\rm low}$ hold. Whereas in the case of $D=256\lambda$ ($N=512$) and $z=398\lambda$, i.e., the solid curve in Fig. 4(a), $n_{\rm up}=0.41$ and $n_{na}=0.22>n_{\rm low}=0.20$, and the error is much smaller. This demonstrates that the larger D pushes the spatial aliasing toward the larger angular range region and hence the accuracy region increases in size. From Table I it is not hard to find that the position of the nonaliasing angular range is in between the two boundaries and hence the angular range determined by Eq. (8) in general may result in a better accuracy than those on the two boundaries. Up to now Eqs. (8)–(10) have been confirmed to well fit with the numerical results.

TABLE I. Locations of the lower and upper boundaries and the nonaliasing angular range calculated from Eqs. (8)–(10) for Fig. 4(a)–(d).

Fig. 4		(a)	(b)	(c)	(d)
$n_{\rm low}$	$N=512$	0.20	0.28	0.40	0.56
	$N=256$	0.20	0.28	0.40	0.56
$n_{\rm up}$	$N=512$	0.41	0.67	0.87	0.96
	$N=256$	0.22	0.41	0.67	0.87
n_{n_a}	$N=512$	0.22	0.41	0.67	0.87
	$N=256$	0.11	0.22	0.41	0.67

III. APPLICATION OF OPTIMAL SELECTION TO THE CALCULATION OF ACOUSTIC FIELDS

In the preceding sections we have dealt with the optimal selection of the numerical parameters for the ASA. The optimal selection may be applied to the implementation of the ASA in two schemes.

The first scheme is that the sampling frequency F_s and the discretization size D are chosen once and fixed in the entire computation, and thus the DFT is performed only once

with the F_s and D and used to construct the whole acoustic field of interest. This scheme may be most suitable for a computer which has a large size memory.

Since the spatial frequency aliasing can be eliminated by use of the DSFT proposed in Ref. 1, selection of F_s should be based on the maximal angular range needed for a certain computation accuracy and the dimension of the field to be calculated. According to the analysis in Sec. II, choosing F_s to be $2/\lambda$, one can numerically get all homogeneous plane waves (whose angular range in the angular spectrum is $0 \leqslant \eta \leqslant 1$) but meanwhile yields some inhomogeneous plane waves since $1 < \eta \leqslant \sqrt{2}$. The inhomogeneous plane waves contribute little to the constructed fields and shall be discarded in the calculation. Therefore, $F_s = 2/\lambda$ can ensure that the ASA's result can converge to the analytical solution as well as required provided that D is large enough to push the aliasing region far enough towards the large angular range.

When D is so small that $n_{na} < n_{\text{low}}$ for a given point of interest in a field, the error will be large because the aliasing appears before the convergence of Eq. (6) is sufficiently good. Therefore, when the dimension of the acoustic field to be calculated is given, Eqs. (8) and (10) should be employed simultaneously to ensure $n_{na} > n_{\text{low}}$. Should a better computation accuracy be needed, D should be further enlarged to make n_{na} sufficiently larger than n_{low}. When the field to be calculated starts at a nonzero distance in the z direction, F_s can be chosen to be less than $2/\lambda$ according to Eqs. (8) and (10) if D or N could be given in advance. When the selected F_s and D are applied to the calculation, the angular range should be chosen according to Eq. (8) and the plane wave components within this range should be used to construct the field.

If the memory of a computer is not so large in size that the 2-D DFT can only be performed with a small size matrix, an alternative scheme for applying the optimal selection can be employed; that is, fixing the size of the 2-D matrix but changing F_s and D according to the location of the calculated field in the z direction and Eqs. (8) and (10). Each time when F_s and D are changed, the DFT must be carried out once again.

Here the first scheme is used to the calculation.

A. Acoustic fields in a homogeneous medium

Acoustic fields in water radiated by the rectangular planer transducer specified in Sec. II are evaluated with dimension $x = \{0 \sim 2A\}, z = \{2\lambda \sim 398\lambda\}$, and the intervals of λ and 2λ in the x and z directions, respectively. Since the field to be calculated in the z direction starts from $2\lambda < 16\lambda$, and from Eq. (10), $n_{\text{low}} = 1$ and thus F_s is chosen to be $2/\lambda$; and since the field ends at $z = 398\lambda$, and from the results in Table I obtained from Eq. (10), $n_{\text{low}} = 0.20 > n_{na} = 0.11$ for $N = 256 (D = 128\lambda)$, whereas $n_{\text{low}} = 0.20 < n_{na} = 0.22$ for $N = 512 (D = 256\lambda)$. Obviously, $N = 256$ does not satisfy the inequality $n_{\text{low}} < n_{na}$ but $N = 512$ does. Thus, N is chosen to be 512. Consequently, the fields calculated have the same physical and numerical parameters as those in Fig. 3(a), (b), and (c) except that Eq. (8) is used here to determine the angular range for the decomposed plane waves used in the construction of the field. Fig. 5(a) and (b) show a comparison of the analytical (solid curve) and the ASA's (dotted curve) solutions of the fields on the z axis ($x = y = 0$) and off the z axis ($x = 2A, y = 0$), respectively. By comparison, the ASA's solutions are in a very good agreement with the analytical ones [Fig. 5(a) and (b)] only except that very small ripples appear in the far field and they are mainly due to the

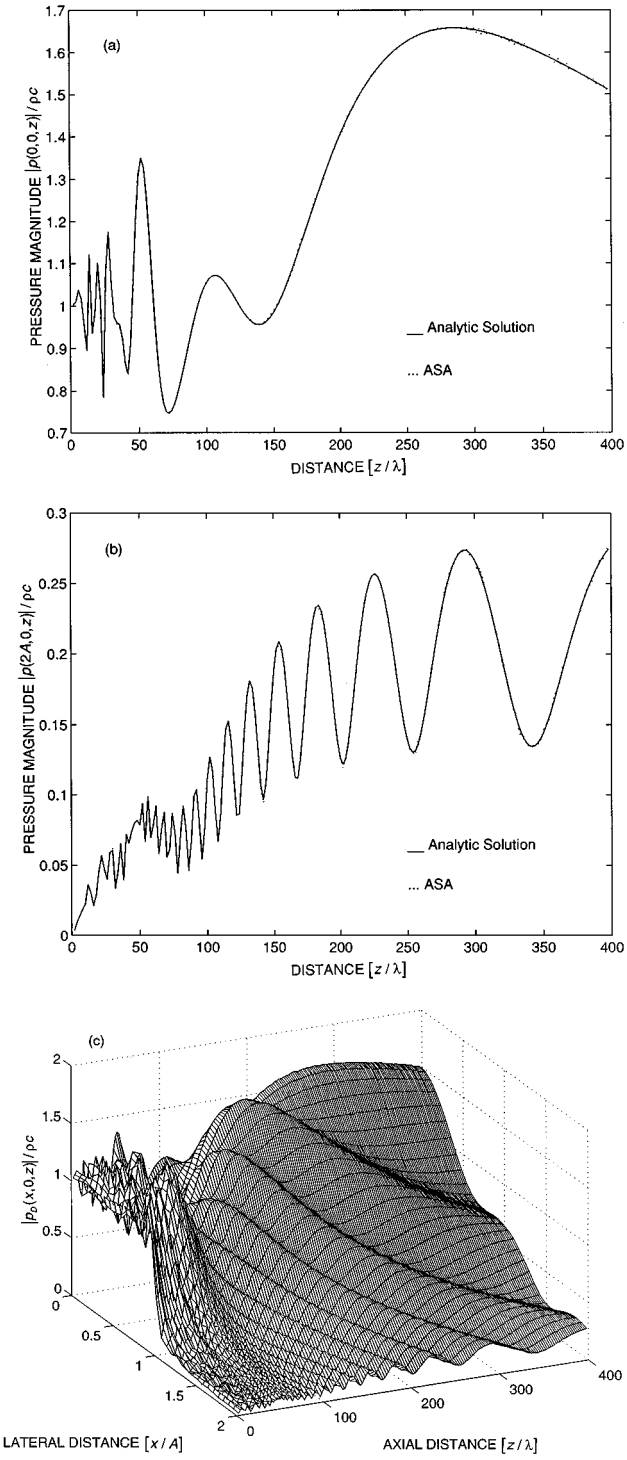


FIG. 5. Acoustic fields from the transducer with the same physical and numerical parameters as in Fig. 3(a) calculated by means of the analytical method (solid line) and the ASA (dotted curve) with the optimal selection of the angular range from Eq. (8). (a) The field on the z axis ($x/A = y/B = 0$) and (b) the field off the z axis ($x/A = 2, y/B = 0$), and (c) 3-D plot of the field on the plane $y/B = 0$ calculated by means of the ASA with the optimal selection.

not-good-enough convergence. Figure. 5(c) shows the 3-D plot of the field in the $x-z$ plane ($y=0$) obtained from the ASA.

B. Acoustic fields in a layered medium

One of the advantages that the ASA possesses is the ability to solve the problems involved with the boundaries. The ASA has therefore, been applied to the computation of fields reflected by a boundary[4,7] and in layered tissues.[8] The effectiveness of the ASA lies in that the decomposition of a finite source into a set of plane waves leads to the easy use of the Snell law and the easy finding of the reflection and transmission coefficients at the boundaries.

To illustrate this, we apply the ASA and the above studied optimal selection to calculating acoustic fields in a layered medium from a 3 MHz linear array. The medium consists of an oil (diesel) layer with thickness $z_{ow}=10$ mm, sound speed $c'=1200$ m/s, and density $\rho'=800$ kg/m^3, and water beneath the oil layer, which has c and ρ specified as in Sec. II. The wavelength in the oil is denoted by λ' and in water still by λ. The linear array used has the same aperture size as the transducer used in Sec. II, but along the side of $2A$, consists of 40 elements which have a width of λ'. No gap between the adjacent elements are considered. The array is intended to generate a focused acoustic field. When it is not focused, the array is equivalent to a rectangular planar transducer. The focusing is carried out only electronically by appropriately delaying the excitation of each element. The delay for each element is calculated by using the geometrical acoustics in such a way that all acoustic pulses coming from the centers of the array elements simultaneously arrive at a focal point expected. Here the expected focal point is located at $z=50$ mm (100λ) on the array axis in water. The field in the layered medium being calculated also has a dimension $x=\{0-16$ mm$\}$, $z=\{1-150$ mm$\}$ or $x=\{0-2A\}$, $z=\{2\lambda-300\lambda\}$ with the intervals of λ and 2λ, respectively, in the x and z directions. The multiple reflections between the transducer and the oil/water interface are not considered here. The pressure field in the layered medium should be calculated separately, that is, the field in the oil layer and the field in the water, which are written as

$$p(x,y,z)=\frac{k'\rho'c'}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{V(k_x',k_y')}{k_z'}\exp[j(xk_x'+yk_y'+zk_z')]dk_x'\,dk_y' \quad (z<z_{ow}), \tag{11a}$$

$$p(x,y,z)=\frac{k'\rho'c'}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{V(k_x',k_y')T(k_x',k_y')}{k_z'}\exp[j(xk_x'+yk_y'+zk_z+z_{ow}(k_z'-k_z))]dk_x'\,dk_y' \quad (z\geqslant z_{ow}), \tag{11b}$$

where k_x',k_y',k_z' and k_x,k_y,k_z are the spatial frequencies in the x,y, and z directions in the oil layer and in water, respectively. If k_x',k_y',k_z' and k_x,k_y,k_z are expressed in terms of spherical angles as

$$(k_x',k_y',k_z')=(k'\sin\theta'\cos\phi,k'\sin\theta'\sin\phi,k'\cos\theta'),$$

and

$$(k_x,k_y,k_z)=(k\sin\theta\cos\phi,k\sin\theta\sin\phi,k\cos\theta),$$

where θ' and θ are the polar angles and also the incident and the refracted angles of plane waves in the oil layer and in water, respectively, and ϕ is the azimuth angle. By means of the Snell's law stating $k'\sin\theta'=k\sin\theta$, we can find the following relation,

$$k_x'=k'\sin\theta'\cos\phi=k\sin\theta\cos\phi=k_x,$$

$$k_y'=k'\sin\theta'\sin\phi=k\sin\theta\sin\phi=k_y,$$

which has been used in deriving Eq. (11b). $T(k_x',k_y')$ in Eq. (11b) is the transmission coefficient of pressure at the liquid/liquid interface and it can be easily found elsewhere (e.g., see Ref. 8). Equation (11b) is not difficult to numerically implement in the way similar to that in Sec. II. And the way of optimally selecting parameters presented in Sec. III can be easily applied to the case of layered medium, just noting that the wavelength used in Eq. (10) must be the one in the medium in which the field points calculated are located, e.g., in the present case, λ' must be used over the range $0\leqslant z<10$ mm and λ over 10 mm$\leqslant z<150$ mm.

F_s is chosen to be $2/\lambda'$ in order to numerically get all homogeneous plane waves. N is chosen to be 512 and hence D is $256\lambda'$ or 205λ. From Eqs. (8) and (10), we obtain that for the farthest field point in water $z=150$ mm, $n_{na}=0.235>n_{\text{low}}=0.231$. By numerical implementation of Eq. (11), the fields radiated by the array without and with focusing into the oil/water layered medium are calculated and presented in Figs. 6 and 7, respectively, in which the pressure magnitude is normalized by $\rho'c'$, the acoustical impedance in oil but the distance z still by λ. Hence, the oil/water interface is located at $z=20\lambda$, where the reflection and refraction happen and discontinuities appear in the fields. From these two figures, we can see that, since the aliasing is eliminated, all the plots are quite smooth only with an exception that some very small ripples remain in the far field zone.

IV. CONCLUSIONS

Optimal selection of parameters for the DFT-based angular spectrum approach (ASA) to the calculation of 3-D acoustic fields has been presented. With the extension of the methods used and the results obtained in the 2-D case in Refs. 1 and 2, the discrete sinc-Fourier transform (DSFT) proposed in Ref. 1 has been applied and the characteristics of the ASA as a function of the angular range in the 3-D case have been investigated. Based on the minimization of spatial aliasing errors in the constructed field, optimal selection of the parameters such as the spatial sampling frequency, the discretization size, and the angular range has been developed. Since the DSFT enables us to remove the spatial frequency aliasing error in the angular spectrum, the spatial sampling frequency may be chosen, according to the dimension of a calculated field, to be less than or equal to $2/\lambda$, where λ is the wavelength in the medium in which the calculated wave fields propagate, in order to obtain a sufficient convergence and to reduce the spatial aliasing error as well. A formula to determine the sufficiently good convergence or

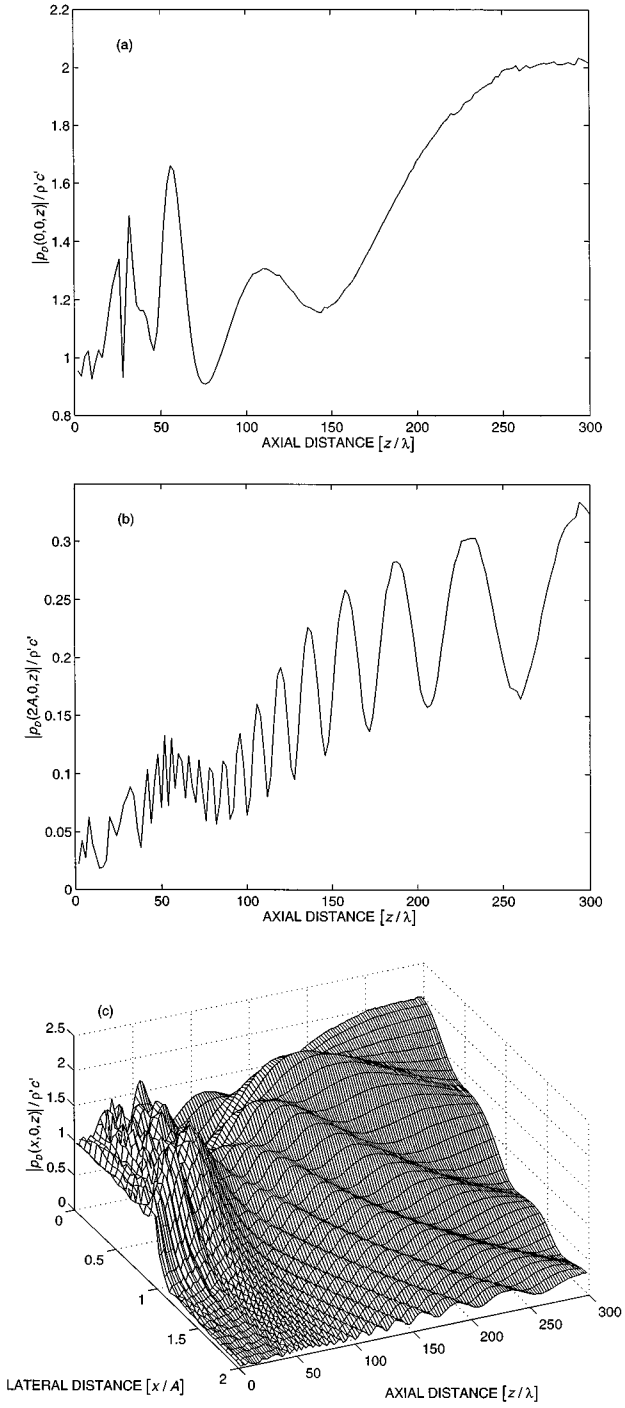


FIG. 6. Acoustic fields from the nonfocused array in the oil/water layered medium obtained using the ASA with optimally selected parameters. The oil layer is 20λ (10 mm thick and starts at $z=0$. The array consists of 40 elements which have a width of λ' along the 2A side and a length of 24λ along the 2B side (see Fig. 1). The numerical conditions used are $F_s=2/\lambda'$ and $N=512$ but the angular range is chosen by Eq. (8). (a) The field on the z axis ($x/A=y/B=0$) and (b) the field off the z axis ($x/A=2, y/B=0$), and (c) 3-D plot of the field on the plane $y/B=0$. Note that the vertical axis is normalized by $\rho' c'$ and the horizontal axis by λ, respectively.

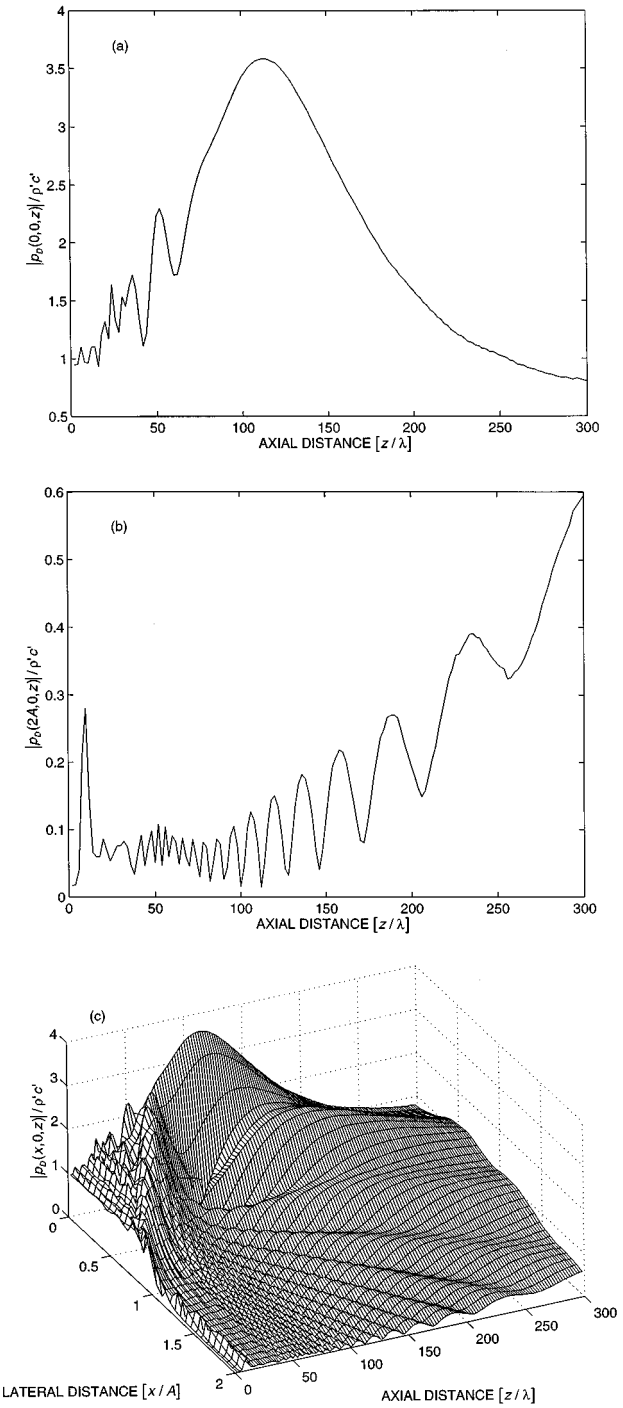


FIG. 7. Acoustic fields from the focused array in the oil/water layered medium obtained using the ASA with optimally selected parameters. The medium, the array and the normalization are the same as in Fig. 6 except that the array is focused. The focal point is expected to be located at $z=100\lambda$ (50 mm) in water.

the lower boundary of a good accuracy region in the 2-D case is confirmed to be valid and effective in the 3-D case. Applying the concept of instantaneous frequency to the Fourier transformed Green's function, which is viewed as a ''signal'' in the transformed domain, the analytical formula has been obtained which, for a given discretization size, allows us to determine the angular range over which the decomposed plane waves should be used for constructing the field with minimal spatial aliasing error. More important is that this formula can be used to predict such a discretization size that the spatial aliasing can be eliminated completely. In combination of the above two formulas, two schemes of optimal selection of parameters for the ASA to calculating 3-D fields have been proposed according to the size of a comput-

er's memory. The ASA with optimal selection of parameters has been applied to the calculations of the fields radiated by rectangular planar sources and linear arrays into homogeneous and layered media. The results obtained demonstrate that the way proposed to select the sampling frequency, the discretization size, and the angular range allows the two types of aliasing errors to be eliminated and hence enables the ASA to obtain a best computation accuracy for the parameters selected.

ACKNOWLEDGMENTS

This work was sponsored by the National Board for Industrial and Technical Development (NUTEK) under Contract No. 8420-94-03570 and by the Swedish Nuclear Fuels under Contract No. 040528.

[1] P. Wu, R. Kazys, and T. Stepinski, ''Analysis of the numerically implemented angular spectrum approach based on the evaluation of two-dimensional acoustic fields. Part I. Errors due to the discrete Fourier transform and discretization,'' J. Acoust. Soc. Am. **99**, 1339–1348 (1996).

[2] P. Wu, R. Kazys, and T. Stepinski, ''Analysis of the numerically implemented angular spectrum approach based on the evaluation of two-dimensional acoustic fields. Part II. Characteristics as a function of angular range,'' J. Acoust. Soc. Am. **99**, 1349–1359 (1996).

[3] D. P. Orofino and P. C. Pedersen, ''Efficient angular spectrum decomposition of acoustic sources—Part I. Theory,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **40**, 238–249 (1993).

[4] D. P. Orofino and P. C. Pedersen, ''Efficient angular spectrum decomposition of acoustic sources—Part II. Results,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **40**, 250–257 (1993).

[5] 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).

[6] E. G. Williams, ''Numerical evaluation of the radiation from unbaffled, finite plates using the FFT,'' J. Acoust. Soc. Am. **74**, 343–347 (1983).

[7] D. P. Orofino and P. C. Pedersen, ''Evaluation of angle-dependent spectral distortion for infinite, planar elastic media via angular spectrum decomposition,'' J. Acoust. Soc. Am. **93**, 1235–1248 (1993).

[8] P. T. Christopher and K. J. Parker, ''New approaches to the linear propagation of acoustic fields,'' J. Acoust. Soc. Am. **90**, 507–521 (1991).

[9] P. T. Christopher and K. J. Parker, ''New approaches to nonlinear diffractive field propagation,'' J. Acoust. Soc. Am. **90**, 488–499 (1991).

[10] K. B. Ocheltree and L. A. Frizzell, ''Sound field calculation for rectangular sources,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **36**, 242–248 (1989).

[11] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill New York, 1968), Chap. 7, pp. 306–399.

[12] G. C. Sherman, ''Application of the convolution theorem to Rayleigh's integral formulas,'' J. Opt. Soc. Am. **57**, 546–547 (1967).

Two Scholte–Stoneley waves on doubly fluid-loaded plates and shells

J.-P. Sessarego, J. Sagéloli, and C. Gazanhes
CNRS, Laboratoire de Mécanique et d'Acoustique, 31 Chemin Joseph Aiguier, Marseilles 13402 Cedex 20, France

H. Überall
CNRS, Laboratoire de Mécanique et d'Acoustique, 31 Chemin Joseph Aiguier, Marseilles 13402 Cedex 20, France, Ecole Supérieure de Mécanique de Marseille, Université Aix-Marseille II, Marseilles 13451 Cedex 20, France and Department of Physics, Catholic University of America, Washington, DC 20064

(Received 14 October 1995; accepted for publication 29 August 1996)

Previous theoretical and experimental studies of sound scattering from plates and from evacuated cylindrical or spherical shells with one-sided water loading have demonstrated the existence of a water-borne Scholte–Stoneley wave, and its acoustic excitation, in addition to that of the Lamb-type plate or shell wave modes. For two-sided water loading two Scholte–Stoneley waves, of symmetric (S) and antisymmetric (A) nature, were predicted on plates, with only the A wave surviving for the case of one-sided loading, while for loading with two different fluids, again two such waves have been demonstrated theoretically. In the present investigation, these two Scholte–Stoneley waves are studied experimentally via short-pulse scattering from water-immersed, thin-walled cylindrical shells filled alternatingly with air, water, and alcohol, and a theoretical analysis of their dispersion curves is presented. © *1997 Acoustical Society of America.* [S0001-4966(97)01201-0]

PACS numbers: 43.20.Tb, 43.30.Jx, 43.40.Ey [ANN]

INTRODUCTION

Scholte–Stoneley-type interface waves have first been introduced in the geophysical context,[1,2] but their name has more recently been applied[3,4] to surface waves on immersed elastic objects, excited by incident sound waves. Test objects in corresponding experiments have often been thin-walled cylindrical or spherical metal shells, and ever since the pioneering experiments on cylindrical shells by Bunney *et al.*,[5] the complete analogy and great quantitative similarity of the elastic-wave modes on thin metal shells to those on free plates, symmetric or antisymmetric ''Lamb waves,'' has repeatedly been emphasized.[6–15] If such plates (or shells) are water-loaded on both sides, an additional pair of waves has been predicted[16] consisting of a symmetric (S) and antisymmetric (A), largely fluid-borne wave, while the corresponding pseudo Lamb-wave modes S_i and A_i ($i=1,2,3,...$) on water-loaded plates or shells remain largely shell-borne and, as stated are very similar to the free-plate Lamb waves. If the water loading is removed from one side of the plate (or, for the case of submerged shells, the shell is evacuated or air-filled), the S wave disappears and only the A wave remains;[6,7] this A wave (no longer antisymmetric) is the one that has been referred to as the ''Scholte–Stoneley wave.'' Figure 1 shows dispersion curves of the S, A and the lowest-order flexural Lamb wave A_0 for a duraluminum plate, plotted versus the frequency-thickness product fd (in MHz×mm); the lower scale refers to the value of the reduced wave number ka (k=acoustic wave number in water, a=radius) for a cylindrical duraluminum shell of inner-to-outer ratio $b/a=0.96$ that has been used in our experiments (see below).

The connection between the scales is given by

$$fd=\frac{C_w}{2\pi}\left(1-\frac{b}{a}\right)ka, \tag{1}$$

the sound velocity in water being taken as $C_W=1480$ m/s.

The figure shows the phase velocity dispersion curve of the flexural, lowest-order Lamb wave A_0 on a free plate[13] as a dashed line. One-sided water-loading splits (or ''bifurcates'')[11,12] this curve into the two solid-line phase velocity ($C^{\rm ph}$) curves[12,13] labeled A_0 and A, i.e., the fluid-borne Scholte–Stoneley wave A has been introduced by the water loading. Also shown is the solid-line group velocity curve (C^g) of the A wave[17] obtained from the formula

$$C^g=\left\{\frac{d}{d(fd)}\frac{fd}{C^{\rm ph}}\right\}^{-1}. \tag{2}$$

The nature of the A and A_0 waves has been discussed recently.[13] Proceeding downward in frequency from the right-hand side of Fig. 1, their dispersion curves start out to converge, A_0 being the plate-borne and A the fluid-borne (Scholte–Stoneley) wave. However, the convergence fails to materialize near the critical (''coincidence'') value $C^{\rm ph}(A_0)\cong C_W$ and an apparent repulsion of the phase velocity curves takes place (although eventually as $fd\to 0$, both curves do end up together at $C^{\rm ph}=0$). In the repulsion region, the waves exchange their nature,[13] so that in the left-hand region of Fig. 1, A_0 is the fluid-borne, and A is the plate-borne wave.

For double-sided water loading of the plate, the curves of $C^{\rm ph}$ and C^g for the A wave get replaced by the slightly different[17] dash-dot curves shown in Fig. 1, and the symmetric S-wave curve S of the second Scholte–Stoneley wave

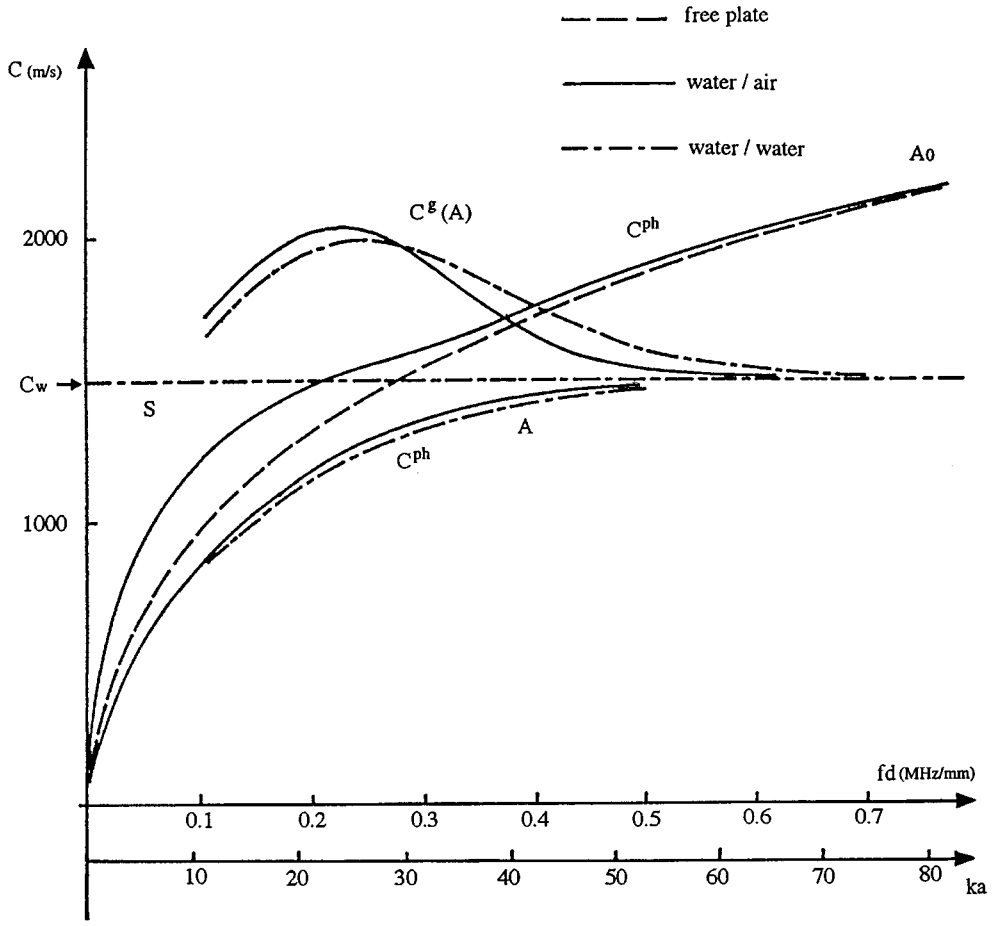


FIG. 1. Phase and group velocity dispersion curves of the S, A Scholte–Stoneley waves and lowest-order (A_0) Lamb wave on a duraluminum plate: dashed, free-plate; solid: one-sided water loading; dash-dot: two-sided water loading.

also appears,[16] which over its entire length is practically equal to the constant water sound speed C_W. It is thus nondispersive, so that its phase and group velocities[7] coincide.

It should be stated here the a different naming convention of the A wave exists in some parts of the literature,[11,18] where it is referred to as the a_{0-} wave.

I. ONE-SIDED WATER LOADING

For the case of thin air-filled cylindrical shells immersed in water, short-pulse experiments[6,7,18] have shown that in ka regions centered around the coincidence frequency (20–50) and even higher up, two series of backscattered pulses appear that have sequentially circumnavigated the shell (here and in all the following, incidence normal to the cylinder axis is maintained): one high-speed pulse that can be interpreted as being due to the S_0 pseudo-Lamb wave, and one of lower speed that agreed well with the A wave (i.e., regarding its group velocity C_g as determined by the time interval between successive pulses). In these experiments, and in related calculations,[19] no evidence for pulses corresponding to the excitation of the flexural A_0 wave has been seen, which is also the case for all our experiments mentioned below. This can be attributed[19] to the large width of the A_0 resonances, which at coincidence becomes exceedingly large.[6,10] Although the A_0 resonances get narrower and increase in magnitude with increasing frequency, the flexural wave becomes significant[19] only beyond $ka \approx 100$.

We have carried out experiments with short pulses of center frequency 1 MHz. The experimental arrangement is described in Fig. 2. A very short electrical pulse was applied to a transducer Panametrix (1 in. in diameter and 1-MHz bandwidth at -6 dB) which was used for both signal transmission and reception. The emitted signal was normally incident on an air-filled cylindrical duraluminum shell whose characteristics were: $a=8.25$ mm, $b/a=0.96$, length $L=150$ mm, density $\rho=2.7$ g/cm^3, compressional speed $C_L=6300$ m/s, and shear speed $C_T=3080$ m/s. (The same shell will be used below with water or alcohol fillings, and with similar incident pulses.) Regarding the narrow beamwidth of the transducer compared to the length of the cylinder, it can be concluded that the geometry was two dimensional. Figure 3 shows the time series of backscattered echo pulses (completely analogous to those of Refs. 6 and 7) where the first pulse on the left represents the specular reflection, the series of fast pulses indicated by arrows represents the S_0 pseudo-Lamb wave[6,7] (which will not be considered here in any detail), and the slow pulses labeled $A^{(1)}$,$A^{(2)}$,... constitute the first, second,... circuits of the A wave. In the resonance analysis method for short pulses,[20–22] the resonance spectrum of these pulses has been obtained by gating out the

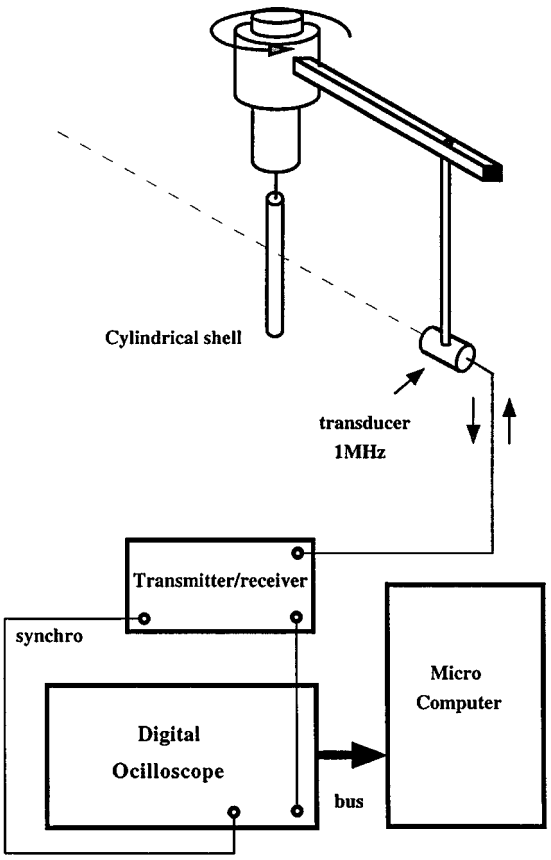


FIG. 2. Experimental arrangement.

specular pulse, and Fourier-analyzing the remainder, with the results (shown in Fig. 4) exhibiting the resonances of both the S_0 and the A wave.

We shall here, in contrast to previous literature, go one step further and gate out in Fig. 3 *all* pulses except those of the A wave, as illustrated in Fig. 5. Fourier analysis of this time series will then provide the pure resonances of the A wave alone, uncluttered by those of the S_0 (or any other)

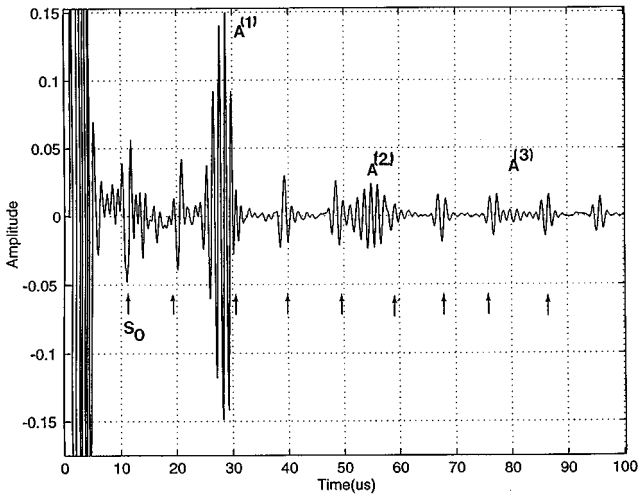


FIG. 3. Experimental backscattering echo response of air-filled, water-immersed cylindrical duraluminum shell (outer radius $a=8.25$ mm, ratio of radii $b/a=0.96$).

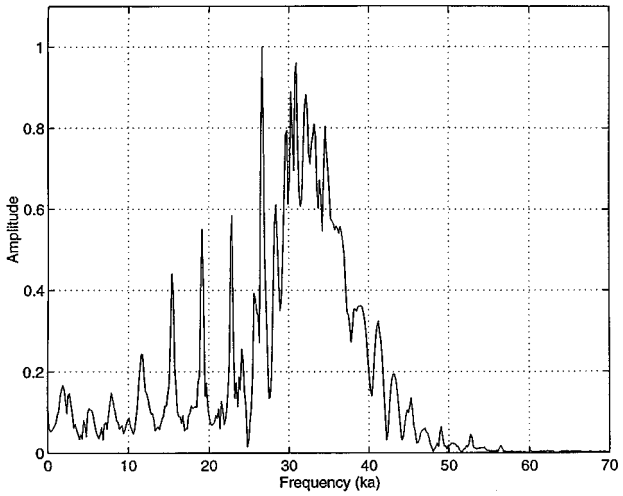


FIG. 4. Resonance spectrum of Fig. 3 after gating out the first (specular) pulse.

wave, as demonstrated in Fig. 6. This approach will be used below to great advantage, since for fluid-filled shells the time series are further cluttered up by transmitted wave pulses in addition to those of the circumferential waves.

From the information presented in Figs. 3 (or 5) and 6, we should be able to obtain agreement with the solid-line dispersion curves of Fig. 1, at our carrier frequency of 1 MHz and with a plate thickness of $d=0.33$ mm (i.e., $ka=35.02$). The group velocity C^g is obtained most simply from the successive arrival times (the repeat times of the circumferential pulses $A^{(2)},A^{(3)},...$ starting from the first pulse $A^{(1)}$). The *entire* phase-velocity dispersion curve C^{ph} could be obtained from the values $(ka)^*$ of the resonance frequencies read off from Fig. 6 via[3,23]

$$C^{\mathrm{ph}}/C_w=(ka)^*/n, \tag{3}$$

where n is the resonance order, if the latter were known. (This is not *a priori* the case here.) The group velocity, besides being obtainable from the phase-velocity curve via Eq. (2), can also be found from the *spacing* Δka^* of the resonances via[7,24]

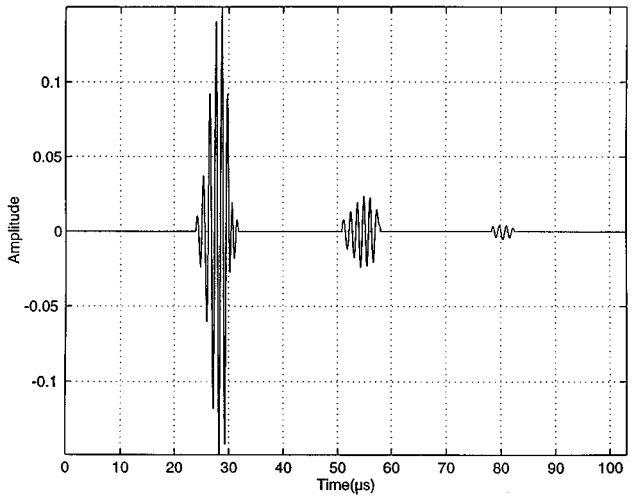


FIG. 5. Echo response as in Fig. 2, but for A wave only.

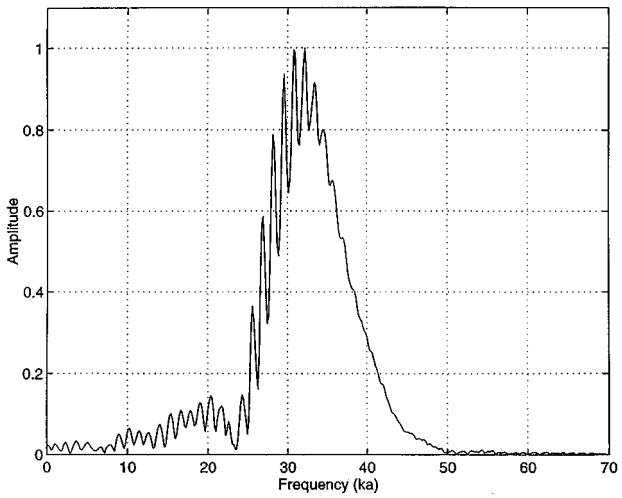


FIG. 6. Resonance spectrum of Fig. 4.

$$C^g/C_w = \Delta ka^*. \tag{4}$$

These various approaches for obtaining values of the phase and group velocities may serve to corroborate their values obtained via one method, by comparing with those obtained from another. In particular, values of n may be assigned to the resonances of Fig. 6 with a reasonable precision, using this approach. If, for example, we ascribe to the resonance at $(ka)^*=30.80$ the resonance order $n=34$, then the values for C^{ph} and C^g obtained from the resonance formulas Eqs. (3) and (4), respectively, agree within 5% with the corresponding solid-curve values in Fig. 1 at $fd=0.33\text{MHz}\times\text{mm}$. In addition, the entire curve of C^{ph}, found from Eq. (3) using the successive values of $(ka)^*$ and n of Fig. 6, is in very good agreement with that of Fig. 1.

The same approach can be used for S_0 waves. Figure 7(a) gives a time representation of all pulses except those of the A wave. This time series has been truncated in order to avoid the overlapping between A and S waves, at the beginning of the response. Fourier analysis of this truncated time series will then provide the pure resonances of the S_0 wave alone, uncluttered by those of the A wave as demonstrated in Fig. 7(b).

II. TWO-SIDED WATER LOADING

The experiment to be described in this section was carried out with a water-filled cylindrical shell. Prevous studies for water-filled cylinders[3,25] have only considered the region $ka \leqslant 10$, while our region of interest is $20 \leqslant ka \leqslant 50$. The echo time series obtained in our experiment (or rather, its computer simulation, see below) is shown in Fig. 8; it is considerably richer in echo pulses than that of the air-filled cylinder, Fig. 3. The reason for this is the presence of transmitted pulses which can penetrate through the water filling of the shell, being backreflected by internal reflections. Strictly speaking, Fig. 8 represents a computer simulation of our experimental results since the experimental pulses showed excessive overlap (although a 1:1 correspondence of pulses existed).

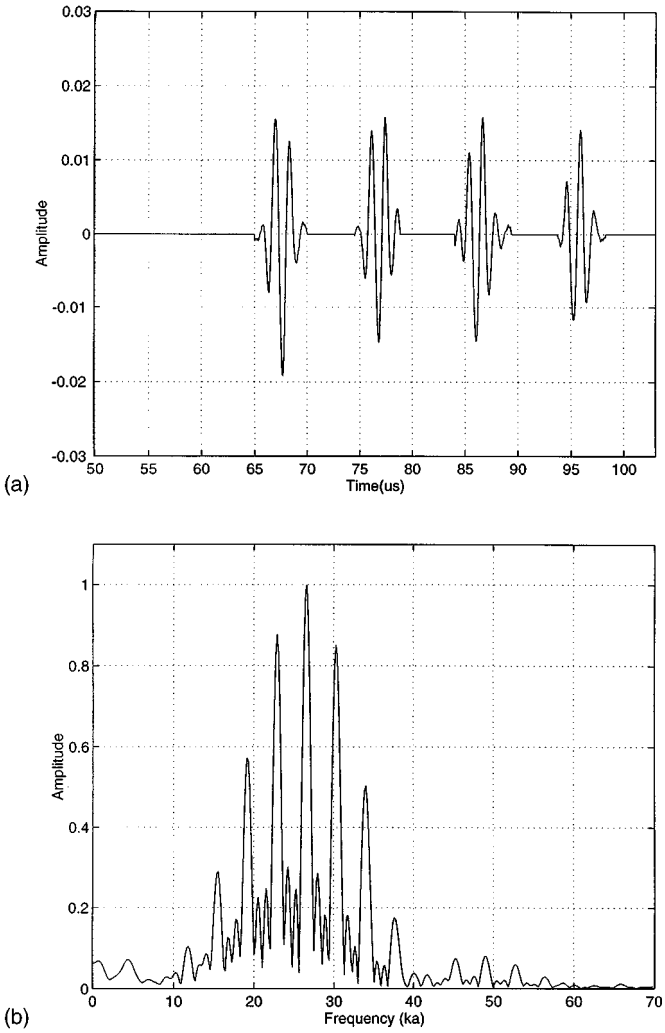


FIG. 7. (a) Time series representing S_0 waves alone. (b) Corresponding resonance spectrum.

Figure 9 illustrates the first three modes of transmission of sound pulses into a water-filled cylinder, and their eventual backreflection. In addition to the first return, there will be subsequent returns due to the additional multiple internal reflections as indicated in each figure. These multiple arrivals

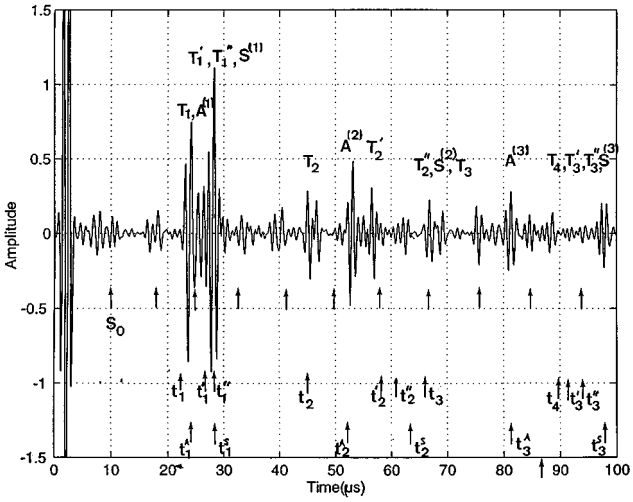


FIG. 8. Calculated backscattering echo response of water-filled, water-immersed cylindrical duraluminum shell (as in Fig. 3).

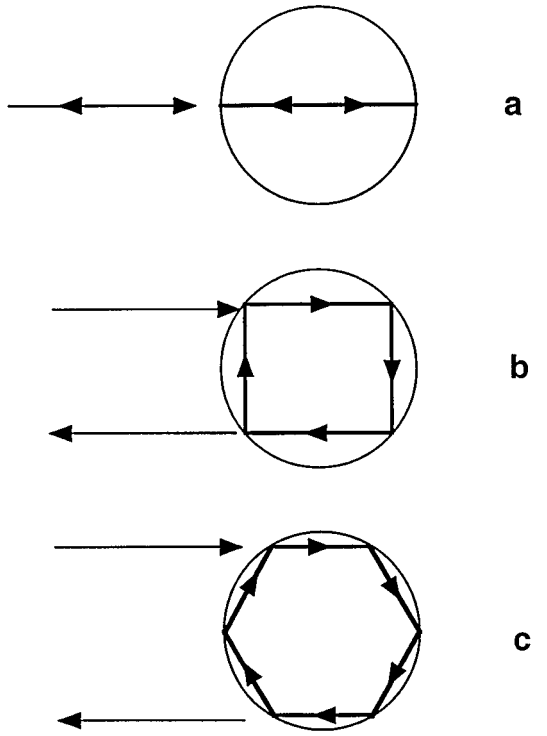


FIG. 9. First three transmission/backreflection modes of water-filled cylinder.

will be labeled t_i for Fig. 9(a), t_i' for Fig. 9(b), and t_i'' for Fig. 9(c) ($i=1,2,\ldots$). Since these higher-order returns get successively smaller (see below), further path forms such as octagonal etc.... are not considered; we also neglect the effect of the thin shell walls and only take the group propagation times in the (external or internal) water media into account. The corresponding arrival times of the transmitted echo pulses have been indicated by arrows labeled by t_i, t_i', and t_i'' in Fig. 8, as were those of the fast S_0-wave pulses which are very similar to those in Fig. 3, due to the fact that all Lamb or pseudo-Lamb waves on thin shells are affected very little by water loading[6,7] (except, of course, the A_0 wave which is coupled to the A wave). From these arrival times, the assignments $T_1, T_1', \ldots, T_3''$ of transmitted pulses as indicated in the figure will be adopted and it is seen that these pulses get successively smaller.

The remaining pulses appearing in Fig. 8 can be interpreted as those of the S- and the A-type Scholte–Stoneley waves whose dispersion curves are shown in Fig. 1 as dash-dot lines. These (subsonic) waves are excited on the shell by the incident wave at tangential incidence,[7] and use of their dispersion curves $C^g(A)$ and $C^g(S)=C^{\mathrm{ph}}(S)$ from Fig. 1 leads to their group arrival times marked by arrows labeled t_i^A and t_i^S ($i=1,2,...$) in Fig. 8, respectively. From these arrival times, the assignments $A^{(1)}, S^{(1)}, A^{(2)}, \ldots, S^{(3)}$ of the multiply-circulating A and S waves can be made in the figure, and it is seen that as in Fig. 1, these initially give rise to quite prominent pulses although they overlap with the initial transmitted arrivals.

In spite of such overlaps, we have again used a gating of the pulse series $A^{(i)}$ alone, and of $S^{(i)}$ alone, in order to obtain a spectrum of resonances of these waves individually,

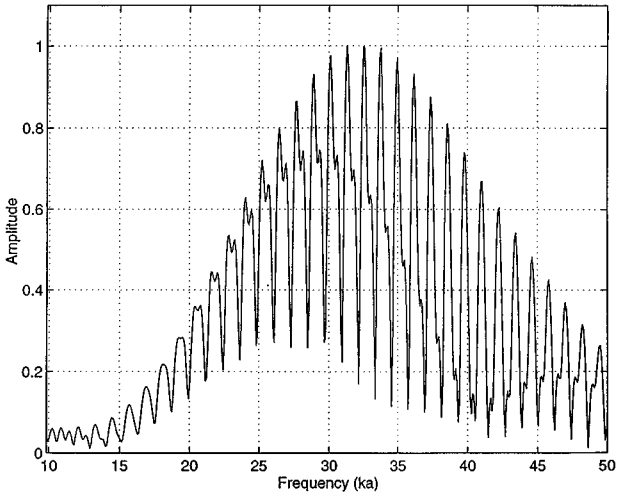


FIG. 10. Resonance spectrum of A wave from Fig. 8.

as we did for the A wave in Fig. 6. The mentioned overlaps, e.g., of $A^{(1)}$ and T_1 nevertheless did not significantly impair the acquisition of a useful A-wave resonance spectrum as shown in Fig. 10; assigning the value $n=52$ to the resonance at $ka=49.30$ led to a perfect reproduction of the phase velocity curve of the A wave in Fig. 1 between $ka=15$ and 50 (the region covered by Fig. 10). Our analysis has thus provided evidence for the existence, and the acoustic excitation, of *two* different Scholte–Stoneley waves A and S on a doubly water-loaded shell which for the related case of a doubly water-loaded plate correspond to an antisymmetric or symmetric wave. In the following, we shall study the case of fluid loading by two different fluids on opposite sides of the shell (or plate) where this symmetry property is perturbed, but the existence of two different Scholte–Stoneley wave modes is retained.

II. TWO-SIDED LOADING BY DIFFERENT FLUIDS

The case of a plate with two different fluids on its two sides has been briefly studied analytically recently.[26] It was found that in addition to the A_0 wave whose dispersion curve of C^{ph} approaches from above the larger one of the two fluid sound speeds when descending from the high-frequency region, there are two (subsonic) phase velocity curves that individually approach the two different fluid sound speed values from below when the frequency increases. These can obviously be interpreted as the generalizations of the two dash-dot (S and A) phase velocity curves of Fig. 1 where the loading fluid was the same (water) on both sides of the plate, and the S and A curves both approached the common value of C_w at high frequencies. In the present case, there are two different fluid limits, with the new ''S'' wave approaching one and the new ''A'' wave the other. This picture can be used in the interpretation of our experimental results for this case. It should be said, however, that our interpretations are here more tentative than for the case of two-sided water loading.

Previous experiments on water-immersed thin cylindrical shells filled with different fluids (chroroform or glycerol)

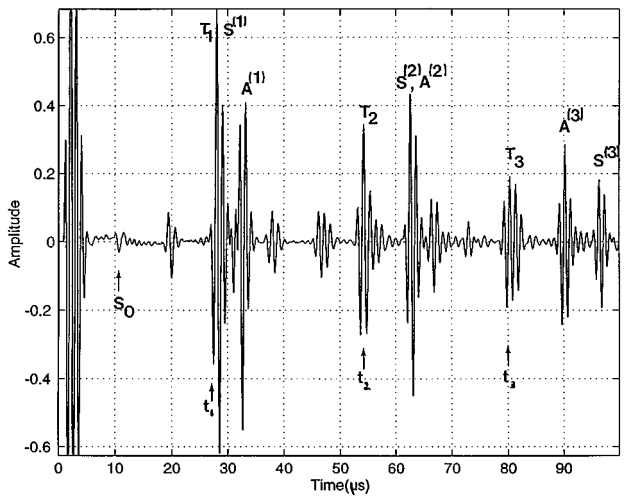


FIG. 11. Experimental backscattering echo response of alcohol-filled, water-immersed cylindrical duraluminum shell (as in Fig. 3).

are discussed in the literature,[3,4] but again as for the case of a water-filled shell, they refer to a ka region ($\leqslant 10$) much lower than the one considered by us here.

Our experiment was carried out with the same cylindrical shell as before, immersed in water ($\rho_w = 1$ g/cm^3, $C_w = 1480$ m/s) and filled with alcohol ($\rho_a = 0.8$ g/cm^3, $C_a = 1200$ m/s). Figure 11 shows the pulse returns for this case, these being the actual experimental results rather than a computer simulation. The S_0 pulse series is still present as in Fig. 8 or Fig. 3, but is not indicated by arrows. When comparing amplitudes, the different scales of Figs. 11 and 8 should be noted, indicating that the pulses of Fig. 11 are not as big as they might appear.

Of the transmitted pulses, we have only indicated by arrows the arrival times t_i of the transmissions corresponding to Fig. 9(a); the higher-order transmissions are altered compared to Fig. 9(b) and (c) by refraction, and are shifted to larger times due to the lower sound speed in alcohol. This leads to the identification of the pulses T_1, T_2 and T_3 as indicated in Fig. 11.

The remaining large pulses in Fig. 11 are candidates for the ''A'' and ''S'' wave, and a study of Figs. 1 and 8 will provide some hints for their identification. Although $C^{\mathrm{ph}}(A) < C^{\mathrm{ph}}(S)$ there, the opposite holds for the group velocities which apply to the pulses. Also, an overlap of one of these first arrivals with T_1 is quite likely, as in Fig. 8. We thus propose the assignments of $S^{(i)}$ and $A^{(i)}$ ($i = 1,2,3$) as indicated in Fig. 11, which ascribes a larger group velocity to A, and maintains equal spacing for repeatedly circumnavigating pulses for both S and A waves. This implies an overlap between T_1 an $S^{(1)}$, as well as between $S^{(2)}$ and $A^{(2)}$. Our following arguments will corroborate these assignments in several different ways.

First, we subject the pulse series $S^{(i)}$ of Fig. 11, as well as the pulse series $A^{(i)}$ ($i = 1,2,3$) to the individual gated-

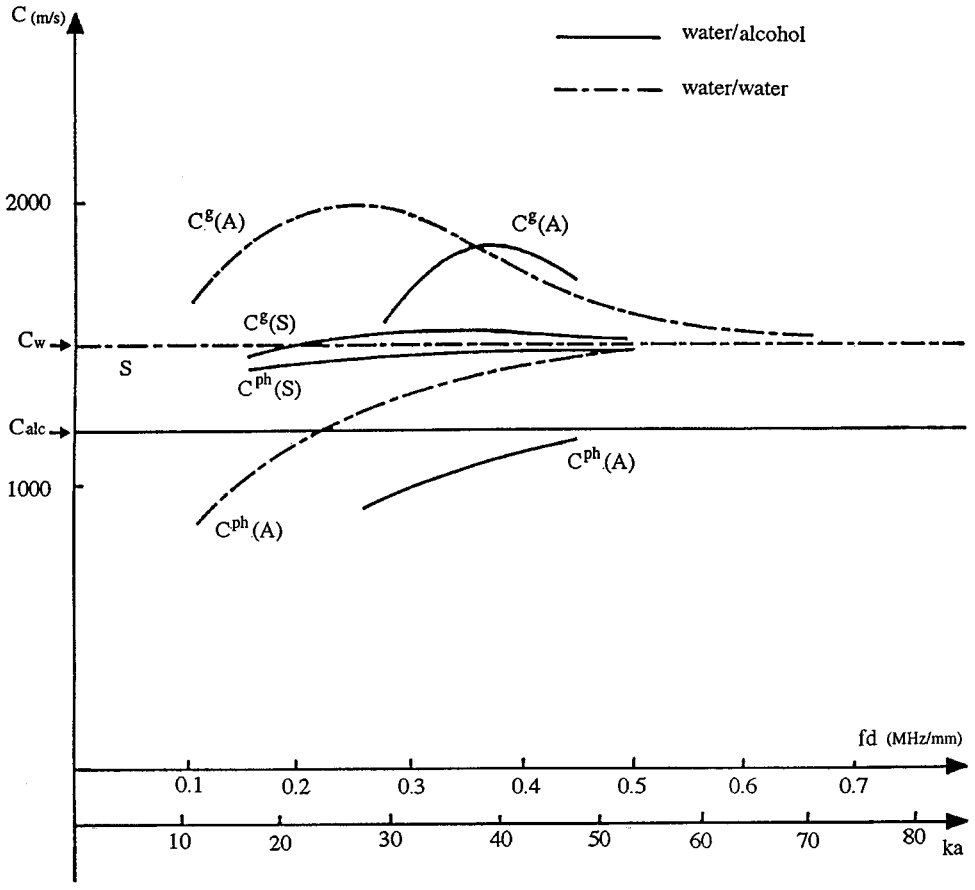


FIG. 12. Phase and group velocity dispersion curves of the S and A Scholte–Stoneley waves on a duraluminum plate or cylinder; solid: water loaded from one side, alcohol loaded from the other; dash-dot: two-sided water loading.

Fourier transform procedure as outlined above, see, e.g., Fig. 5. This leads to a resonance spectrum for both the ''*S*'' and the ''*A*'' waves similar to Fig. 10, and with an appropriate assignment of n values to the sets of resonances, phase velocity dispersion curves may be obtained for these from Eq. (3). The assignment of $n=46$ to the $ka=45.78$ resonance of the *S* wave leads to its limit $C^{\mathrm{ph}}\rightarrow C_w$ for $ka\rightarrow\infty$, and that of $n=60$ to the $ka=45.46$ resonance of the *A* wave leads to its limit $C^{\mathrm{ph}}\rightarrow C_a$ for $ka\rightarrow\infty$, as suggested by the study of Ref. 26. From these choices, the phase and group velocity dispersion curves shown in Fig. 12 are obtained for the *S* and *A* waves of the water/alcohol-loaded shells; here, the group velocity curves were obtained from the phase velocity curves using Eq. (2).

For a reliable identification of *S* and *A* wave pulses, the above procedure *per se* may not be sufficient since it depends on an assumption for the n values. However, two different checks on the group velocities provide convincing corroborration since these do *not* depend on the knowledge of n. Trivially, the arrival time of the pulses provide the group velocity, especially those for the repeated circulations $A^{(2)},A^{(3)},\ldots$ after the first one, $A^{(1)}$. Second, the group velocities obtained from Eq. (4) are based on the intervals Δka^* between resonances, independently of the n value assigned to each individual resonance. Using these two methods for a test on the group velocities $C^g(S)$ and $C^g(A)$ of Fig. 12 we obtain excellent agreement, thus confirming our group velocity results over our ka range of interest and hence, by implication, also our phase velocity results. As an example, at the experimental carrier frequency 1 MHz, the curves of Fig. 12 yield $C^g(A)=1761$ m/s, while the repeat pulse arrivals of the *A* wave in Fig. 11 yield $C^g(A)=1809$ m/s; the use of Eq. (4) and the resonance spacing of the *A* wave leads to $C^g(A)=1753$ m/s, showing a maximum uncertainty of only 3%.

One feature of Fig. 11 may be noted here which deserves further study. Although the group velocity of the pulses at 1 MHz is larger for the *A* wave than for the *S* wave (leading to more rapid repeat arrivals), the *first* arrival of the *S* wave occurs before that of the *A* wave. The latter can be assumed to be excited tangentially[7] (i.e., at the shadow boundary of the cylinder), but using the angular analysis of Ref. 7, one finds that the early arrival of the *S* wave is consistent with its excitation at the apex line of the cylinder.

III. CONCLUSIONS

The present study, based on experimental sound scattering from immersed thin cylindrical duraluminum shells and its theoretical analysis, has provided evidence for the existence and excitation of two Scholte–Stoneley waves on doubly fluid-loaded shells or plates: of *S* and *A* type for double water loading, and of related ''*S*'' and ''*A*'' types for loading by two different fluids. While predictions for the first case go back to the classic Osborne–Hart study on plates,[16] for the second case they are of more recent nature.[26] Our study used short incident sound pulses and observed their backscattering indicating transmitted or repeatedly circumnavigating waves on the cylinder. The latter were analyzed by an extension of the Numrich–de Billy short-pulse spectral analysis method,[20,21] where we gated out *all* the pulses except the candidate pulses, and subsequently Fourier-analyzing these. Their spectra provided phase and group velocity information, in agreement with the predictions, sometimes dependent on assumptions on the mode order value of the resonances and sometimes independent thereof. In addition, the pulse arrival times provided group velocity data directly. These various methods represent a redundancy that serves to corroborate our assignments and pulse identifications which, based on a single approach, would have been more uncertain.

[1] J. G. J. Scholte, ''On seismic waves in a spherical earth,'' Koninklijk Nederlands Meteorologisch Instituut, Commun. No. 65, The Hague, 1956, pp. 1–55.

[2] R. Stoneley, ''Elastic waves at the surface of separation of two solids,''Proc. R. Soc. London, Ser. A **106**, 416–420 (1924).

[3] See, e.g., M. Talmant, J.-L. Izbicki, G. Maze, G. Quentin, and J. Ripoche, ''External wave resonances on thin cylindrical shells,'' J. Acoust. **4**, 509–523 (1991).

[4] J.-L. Izbicki, J.-L. Rousselot, A. Gérard, G. Maze, and J. Ripoche, ''Analysis of resonances related to Scholte–Stoneley waves around circular cylindrical shells,'' J. Acoust. Soc. Am. **90**, 2602–2608 (1991).

[5] R. E. Bunney, R. R. Goodman, and S. W. Marshall, ''Rayleigh and Lamb waves on cylinders,'' J. Acoust. Soc. Am. **46**, 1223–1233 (1969).

[6] M. Talmant, ''Rétrodiffusion d'une impulsion ultrasonore brève par une coque cylindrique à paroi mince,'' Ph.D. thesis, University of Paris VII, 1987.

[7] M. Talmant and G. Quentin, ''Backscattering of a short ultrasonic pulse from thin cylindrical shells,'' J. Appl. Phys. **63**, 1857–1863 (1988).

[8] M. Zakharia, P. Flandrin, J. Sageloli, and J.-P. Sessarego, ''Analyse temps-fréquence appliquée à la caractérisation acoustique de cibles,'' J. Acoust. **1**, 185–188 (1988).

[9] M. Talmant, G. Quentin, J.-L. Rousselot, J. V. Subrahmanyam, and H. Überall, ''Acoustic resonances of thin cylindrical shells and the resonance scattering theory,'' J. Acoust. Soc. Am. **84**, 681–688 (1988).

[10] G. Quentin and M. Talmant, ''The plane plate model applied to the scattering of the ultrasonic waves from cylindrical shells,'' in *Elastic Wave Propagation*, Proceedings of the 2nd IUTAM-IUPAP Symposium on Elastic Wave Propagation, Galway, Ireland, 20–25 March 1988, edited by M. F. McCarthy and M. A. Hayes (Elsevier Science, North-Holland, Amsterdam, 1989).

[11] G. S. Sammelmann, D. H. Trivett, and R. H. Hackman, ''The acoustic scattering by a submerged, spherical shell, I: The bifurcation of the dispersion curve for the spherical antisymmetric Lamb wave,'' J. Acoust. Soc. Am. **85**, 114–124 (1989).

[12] M. Talmant, H. Überall, R. D. Miller, M. F. Werby, and J. W. Dickey, ''Lamb waves and fluid-borne waves on water-loaded, air-filled thin spherical shells,'' J. Acoust. Soc. Am. **86**, 278–289 (1989).

[13] 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).

[14] D. Décultot, F. Lecroq, and G. Maze, ''Diffusion acoustique par des coques cylindriques. Etude des ondes axiales,'' J. Phys. IV, Colloq. C5, Suppl. J. Phys. III, **4** (1995), Proceedings of the 3rd French Conference on Acoustics, Toulouse, France, 2–6 May 1994, pp. C5-841–C5-844.

[15] G. Maze, F. Léon, F. Lecrocq, D. Décultot, and H. Überall, ''Nature de l'onde d'interface de Scholte sur une coque cylindrique,'' J. Phys. IV, Colloq. C5, Suppl. J. Phys. III, **4** (1995), Proceedings of the 3rd French Conference on Acoustics, Toulouse, France, 2–6 May 1994, pp. C5-849–C5-852.

[16] M. F. M. Osborne and S. D. Hart, ''Transmission, reflexion, and guiding of an exponential pulse by a steel plate in water. I. Theory,'' J. Acoust. Soc. Am. **17**, 1–18 (1945); ''II. Experiment,'' **18**, 170–184 (1946).

[17] G. Maze, F. Luppé, J. Duclos, P. Pareige, F. Lecroq, A. Tinel, D. Décultot, M. Ech-Cherif, H. Duflo, and J. Ripoche, ''Diffusion acoustique par des cibles élastiques immergées de forme géométrique simple ou complexe,'' Rapport final, Contrat DRET No. 90/182, Tome II, LAUE, Université du Havre, 76610 Le Havre, France.

[18] G. Kaduchak and P. L. Marston, ''Observation of the mid-frequency enhancement of tone bursts backscattered by a thin spherical shell in water

near the coincidence frequency,'' J. Acoust. Soc. Am. **93**, 224–230 (1993).
[19] M. F. Werby and H. Überall, ''The excitation of water-borne waves at the interface of evacuated elastic spherical shells and pseudo-Stoneley resonances,'' J. Phys. IV, Colloque C1, Suppl. J. Phys. III, Vol. 2 (1992), Proceedings of the 2nd French Conference on Acoustics, Arcachon, France, 14–17 April 1992, pp. C1-1087–C1-1090.
[20] S. K. Numrich, N. Dale, and L. R. Dragonette, ''Generation and exploitation of plate waves in submerged, air-filled shells,'' in *Advances in Fluid-Structure Interactions* (ASME, New York, 1984), PVP Vol. 78/AMD Vol. 64, pp. 59–74.
[21] M. de Billy, ''Determination of the resonance spectrum of elastic bodies via the use of short pulses and Fourier transform theory,'' J. Acoust. Soc. Am. **79**, 219–221 (1986).
[22] S. K. Numrich and H. Überall, ''Scattering of sound pulses and the ringing of target resonances,'' in *Physical Acoustics, High Frequency and Pulse Scattering*, edited by A. D. Pierce and R. N. Thurston (Academic, New York, 1992), Vol. 21, pp. 235–318.
[23] L. Flax, L. R. Dragonette, and H. Überall, ''Theory of elastic resonance excitation by sound scattering,'' J. Acoust. Soc. Am. **63**, 723–731 (1978).
[24] S. K. Numrich, W. E. Howell, J. V. Subrahmanyam, and H. Überall, ''Acoustic ringing response of the individual resonances of an elastic cylinder,'' J. Acoust. Soc. Am. **80**, 1161–1169 (1986).
[25] J.-L. Izbicki, G. Maze, and J. Ripoche, ''Influence of the free modes of vibration on the acoustic scattering of a circular cylindrical shell,'' J. Acoust. Soc. Am. **80**, 1215–1219 (1986).
[26] J. Dickey, G. Maidanik, and H. Überall, ''The splitting of dispersion curves for the fluid loaded plate,'' J. Acoust. Soc. Am. **98**, 2365–2367 (1995).

Nonlinear propagation applied to the improvement of resolution in diagnostic medical ultrasound

B. Ward, A. C. Baker, and V. F. Humphrey
School of Physics, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom

(Received 16 October 1995; revised 20 July 1996; accepted 1 August 1996)

Medical B-mode scanners operating under conditions typically encountered during clinical work produce ultrasonic wave fields that undergo nonlinear distortion. In general, the resulting harmonic beams are narrower and have lower sidelobe levels than the fundamental beam, making them ideal for imaging purposes. This work demonstrates the feasibility of nonlinear harmonic imaging in medical scanners using a simple broadband imaging arrangement in water. The ultrasonic system comprises a 2.25-MHz circular transducer with a diameter of 38 mm, a membrane hydrophone, also with a diameter of 38 mm, and a polymer lens with a focal length of 262 mm. These components are arranged coaxially giving an imaging geometry similar to that used in many commercial B-scanners, but with a receiver bandwidth sufficient to record the first four harmonics. A series of continuous wave and pulse-echo measurements are performed on a wire phantom to give 1-D transverse pressure profiles and 2-D B-mode images, respectively. The reflected beamwidths w_n decrease as $w_n/w_1 = 1/n^{0.78}$, where n is the harmonic number, and the reflected sidelobe levels fall off quickly with increasing n. In imaging terms, these effects correspond to a large improvement in lateral resolution and signal-to-clutter ratio for the higher harmonics. [S0001-4966(97)03801-0]

PACS numbers: 43.25.Ts, 43.25.Cb, 43.80.Vj [MAB]

INTRODUCTION

With the increasingly widespread use of diagnostic ultrasound scanners in a wide range of medical environments, e.g., obstetric and vascular imaging, the need to improve the quality of the displayed visual output is of very real interest, both to manufacturers and to clinicians. Since the spatial resolution attainable with such equipment is dependent ultimately upon the wavelength of the ultrasound used, the chosen frequency of operation should be high if the size of the object under study is small. However, higher frequencies are attenuated more rapidly in biological and other media meaning that the depth of penetration inside the object is reduced as the wavelength of the acoustic field decreases. In practice, a compromise must be established between these two conflicting effects in each new imaging application by the choice of a scan head of suitable center frequency, typically in the range 1 to 10 MHz. An example of this is the now routine examination of the developing human fetus. Earlier on during pregnancy a 5-MHz head would often be used to view the fetus through the filled bladder, the attenuation in urine being relatively low. At later stages, on the other hand, the frequency would be reduced to 3.5 MHz as the fetus grew in size and the depth of penetration became the dominant factor.

One possible method of enhancing the quality of B-mode ultrasound images might be to exploit the effect of nonlinear propagation on the acoustic signal as it travels through the human body. All finite-amplitude ultrasonic waves undergo a degree of nonlinear distortion when traveling through real media, and this manifests itself in the frequency domain as the appearance of additional harmonic signals at integer multiples of the original frequency. The idea that commercial diagnostic B-mode scanners might produce signals of large enough amplitude to cause appreciable nonlinear effects in normal clinical use was first proposed in 1980 by Muir and Carstensen,[1] who also worked on a series of experiments in water with medical transducers to support their predictions.[2] The first observations of nonlinear distortion in tissue, however, were not recorded until some 5 years later by Starritt *et al.*,[3,4] who looked at (i) human calf muscle *in vivo* using a clinical pulse-echo scanner, and (ii) excised bovine liver *in vitro* using a type of focused transducer commonly used in commercial ultrasound systems. In both cases the source pressures and frequencies were typical of those used during medical examinations at the time, i.e., about 600 kPa and 2.5 MHz, respectively. Although the amount of nonlinear distortion occurring was significantly less in tissue than in water, it was clearly measurable, the second harmonic being on average 17 and 12 dB below the fundamental for muscle and liver, respectively.

Wherever appreciable nonlinear distortion does occur during clinical work with B-mode scanners, then it should in principle be possible to use the higher harmonics to improve the spatial resolution of the resulting images. Current medical systems do not make significant use of these higher harmonics because of the limited response of the receiving transducer at the harmonic frequencies. Only by analyzing the newly created harmonic beams may some of this information be recovered in a useful form. Indeed, since the widths of the beams decrease with increasing harmonic number, an actual improvement in quality of the visual information available to the clinician should be expected.

This kind of approach has already been shown to work in the field of acoustic microscopy, where the frequency of operation is typically 100 MHz, i.e., at least an order of magnitude higher than in medical B-mode scanning. Ko-

mpfner and Lemons[5] in 1976 were the first to demonstrate second harmonic imaging in water using a specially tuned microscope at a fundamental frequency of 400 MHz. They increased the power applied to the transmitter by 40 dB above its normal linear level and noted improvements in both the resolution and the contrast of the second harmonic image compared to the fundamental. Rugar[6] in 1984 looked at the problem from a slightly different standpoint. He transmitted and received at a single frequency in the low gigahertz range and found the spatial frequency response of images formed at nonlinear drive levels to be a factor of 1.4 better than those formed during conventional linear imaging. This was attributed to a conversion of the energy in the higher harmonics down to the fundamental frequency in the region behind the object. The work was carried further by Germain and Cheeke[7] using a device operating at 15 MHz with ethanol as the coupling liquid to maximize harmonic generation. They separated the harmonic components out from one another and observed an improvement in resolution proportional to $n^{1/2}$ for $n \leqslant 10$, n being the harmonic number. More recently, work has begun using acoustic microscopes with pressurized cryogenic liquids.[8] Here, the improvement in resolution for the second harmonic is better than a factor of $2^{1/2}$, possibly due to the special nonlinear properties of the coupling liquid.

The use of harmonic imaging in the context of side-scan sonar has been considered in some detail by Muir.[9] He looked specifically at a system based on a 1.0-m$\times$0.1-m rectangular transducer using tone bursts with a center frequency of 100 kHz. For targets between 50 and 100 m from the source, considerable improvements in angular resolution and sidelobe suppression were observed as the harmonic number was increased from 1 to 5. In going from the fundamental to the second harmonic, for example, the half-power resolution was reduced from 0.8° to 0.5° and the suppression went up from 13 to 20 dB. As a consequence, the quality of image obtained using the higher harmonics was significantly better.

Due to the relatively high attenuation of ultrasound in human tissue and the complexity of today's commercial diagnostic B-mode scanners, a good starting point for studying nonlinear propagation at medical frequencies and source pressures is to consider a simple focused circular source in water. This approach has the added advantage that a large body of work has already been built up on the circular arrangement, enabling comparisons between experiment and theory to be more easily made. Although the levels of the harmonics generated via nonlinear processes are bound to be higher in water than in tissue, they should display the same kinds of features. Since the aim of this study is to demonstrate the feasibility of harmonic imaging in medical systems, it should act as a basis for further work with more realistic media. Previous work in this field[10] demonstrated the enhanced beamwidths available for clinical ultrasound but did not produce any images.

The numerical model that has since become a standard method for calculating the acoustic field ahead of plane and focused circular sources was implemented by Aanonsen *et al.* in 1984, and is now widely referred to as the Bergen code as a result.[11,12] It is based on the KZK equation, which is a parabolic approximation to the nonlinear wave equation for the system, and it provides both amplitude and phase information along the acoustic axis of the transducer and radially. The model has been extensively tested by comparison with experimental data from plane[13,14] and focused[15] transducers, and found to give excellent results for continuous-wave (cw) sources with diameters large compared to the acoustic wavelength. Since commercial diagnostic B-scanners tend to produce pulsed fields with very complicated structures, it has proved somewhat harder to model their outputs effectively. Baker[16] has adapted Aanonsen's code for use with pulsed sources and obtained reasonable agreement between experiment and theory for a ''real'' focused transducer working at 3.5 MHz in water, while Watson *et al.*[17] have looked at cw propagation in tissue-like media. Yet to be attempted, however, is a combination of these two approaches applied to commercial scanners in a clinical environment.

The work described in this paper considers the relatively simple case of nonlinear imaging in water using a focused circular transducer with a center frequency of 2.25 MHz, a focal gain of around 6, and source pressures of up to 400 kPa. These values are comparable with diagnostic B-mode scanners in medicine. Measurements are made on a wire phantom using a focused broadband receiver with a diameter equal to that of the transmitter, thus simulating the imaging geometry in many real medical ultrasound systems. This is similar to the arrangement studied by Gavrilov *et al.*[18] and Levin *et al.*,[19] who were concerned with remote linear measurements of acoustic pressure in biological media. Whereas their experiments were performed in transmission mode, however, the work described in this paper uses a reflection mode setup. It has recently been suggested by Wu and Shung[20] that it should be possible to ''prebias'' the signal applied to the transducer for imaging purposes, i.e., to mix higher frequency components in at the source, thus enhancing the amplitudes of the harmonics around the focal region, but this has not been attempted in this work. It should be emphasized that an increase in drive is not required or proposed for this technique, it merely makes better use of the existing ultrasonic dose.

I. MAIN EXPERIMENTAL ARRANGEMENT

Measurements of the broadband acoustic signal reflected from the test object were performed using the experimental arrangement shown in Fig. 1(a). The system was based on a circular single-element immersion transducer with a resonant frequency f_0 of 2.25 MHz and a nominal radius a of 19 mm. The transducer was driven at its center frequency via an adjustable attenuator and a 50-dB broadband rf power amplifier using a pulsed function generator. For pseudo cw measurements a sinusoidal tone burst of 8 cycles was applied to the device. For pulse-echo B-mode measurements, in contrast, it was necessary to use a short single-cycle sinusoidal pulse to achieve good axial resolution. In both types of measurement the maximum voltage applied to the transducer was limited to about 300 V pk-pk corresponding to an average pressure of 400 kPa across the face of the source. In addition, the pulse repetition frequency was set relatively low at 250 Hz in order to limit the power dissipated in the transducer.

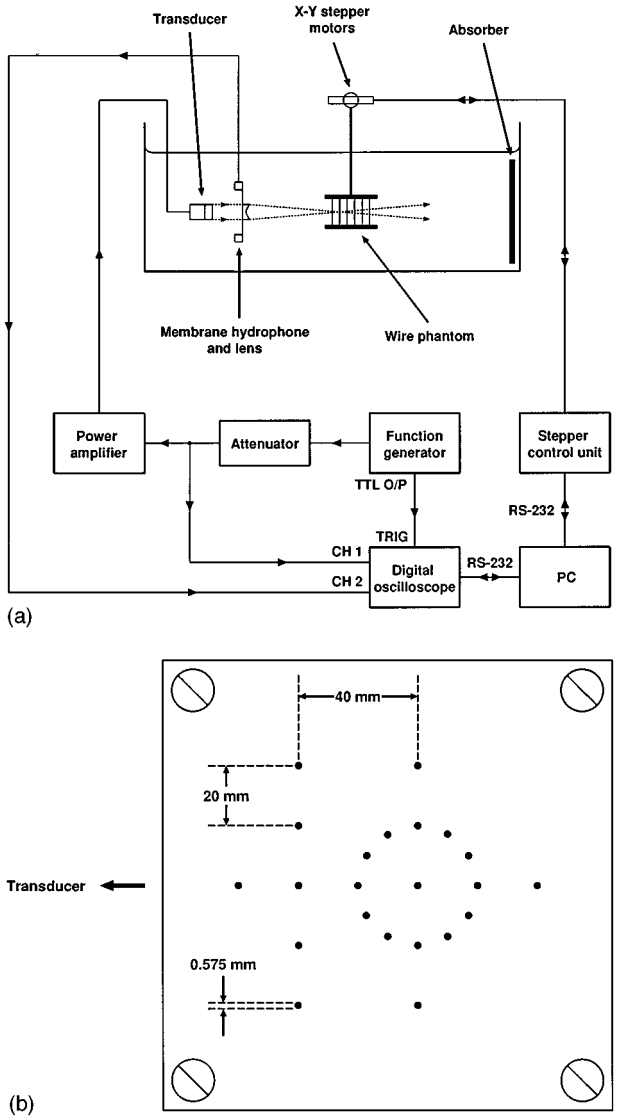


FIG. 1. (a) Diagram of the main experimental arrangement. (b) Schematic plan of the ''wire'' phantom.

The backscattered signal was measured using a polyvinylidene-fluoride (PVdF) bilaminar shielded membrane hydrophone manufactured by GEC-Marconi Ltd. of Chelmsford, England. This device was constructed with a circular active area at the center of the membrane with a nominal radius b of 19 mm, i.e., the same as that of the transducer. Due to its wide bandwidth and relatively flat frequency response the hydrophone was suitable for measurements at frequencies up to at least 20 MHz. In the work described here only the first four harmonics of the transmitted signal were considered, corresponding to a maximum frequency of 9 MHz. This meant that the hydrophone was only used well within its useful range of operation. Since the active area of the device was much larger than usual for hydrophones of this type, it was first necessary to characterize its output by comparison with a smaller-area precalibrated device. This second hydrophone was also produced by GEC-Marconi and was similar in construction to the large-area device, but had an active area with a diameter of only 1 mm. The calibration of the larger hydrophone over the frequency range of interest was performed by placing it in the nonlinear acoustic field formed by the transducer for a source pressure p_0 of about 350 kPa. The source pressure was determined using the small hydrophone at short axial ranges. The electrical signal produced in this case was then compared with the transverse profile obtained at the same point in the field by scanning the smaller hydrophone over the width of the larger device. By taking into account the amplitude and phase variations over this region and the different areas of the two devices, the relative sensitivities of the hydrophones were determined. Using these values it was then possible to calculate the sensitivity of the uncalibrated large-area hydrophone.

The transducer and hydrophone were positioned along a common horizontal axis in a tank of distilled water during the measurements of backscatter from the object [see Fig. 1(a)]. A small amount of chlorine was added to the water after the tank had been filled in order to prevent the growth of algae. The tank had dimensions of 1.0 m$\times$0.5 m$\times$0.4 m and all measurements were performed at ambient room temperature. The output of the transducer was allowed to pass through the hydrophone at an axial range z_0 of 50 mm and then on through a plano-concave PMMA lens with a focal length D in water of 262 mm. The lens had a maximum thickness of 6 mm, a diameter of 44 mm, and a radius of curvature of 117 mm. The resulting focal gain G in water was 5.7 and the linear beam waist (full-width half-maximum, FWHM) w_0 at the focal plane was about 6 mm. The lens was acoustically coupled with a thin layer of commercial coupling gel and held in place with self-adhesive tape. After reflection from the test object, most of the acoustic signal passed back through the lens and onto the hydrophone for measurement. Using this arrangement the transmitted and received signals were focused by roughly equal amounts and the source and receiver were nearly conjugate.

The electrical signal produced by the hydrophone was passed by way of a coaxial cable to a 300-MHz digital storage oscilloscope (D.S.O.). This instrument (LeCroy 9310) had two input channels, each comprising an 8-bit A-to-D converter with a maximum sampling rate of 10^8 s^{-1}, and a 10 000-point acquisition memory. During each measurement sequence 1000 consecutive waveforms were averaged within the D.S.O. resulting in an effective 13-bit signal with a 30-dB reduction in the digitization error. In addition, the acquired data were passed through the oscilloscope's internal 30-MHz bandwidth limiter leading to a further reduction in the background noise. The D.S.O. was triggered direct from the TTL ''synchronization'' output of the function generator (Wavetek 191), and the D.S.O. trigger delay was set manually so as to compensate for the water path between the points of transmission and reception.

The D.S.O. was controlled remotely via an RS-232-C serial connection from a PC with a 486DX2 processor running at 66 MHz. Waveforms were transferred from the instrument to the PC as a series of coded ASCII characters at a rate of 9600 baud. Software for the control of the D.S.O. was written in the form of a series of Microsoft QuickBASIC routines and the stored waveforms were subsequently processed using the MathWorks MATLAB software package.

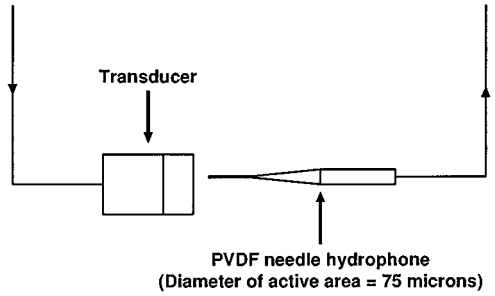


FIG. 2. Diagram of the arrangement used to measure the amplitude and phase profiles across the source transducer.

The PC was also used to control a pair of stepping motors on two translation stages via a second RS-232 interface. The stages were used to move the test object along two mutually orthogonal horizontal axes, enabling long series of measurements to be carried out automatically.

The test object itself is shown in Fig. 1(b). The supporting frame was made from two parallel 5-mm sheets of PMMA held apart using spacers with a length of 50 mm. Through each sheet was drilled an array of small holes at intervals of roughly 20 mm. These holes were used to hold 22 stainless-steel ''wires'' made from hollow hypodermic needles, each with a diameter of about 0.6 mm. The object was placed in the tank in front of the transducer and hydrophone with the polymer sheets oriented horizontally and the first row of five wires at the focal plane of the lens (i.e., at $z=D=262$ mm). For the pseudo cw measurements only this row of wires was considered, so the length of the recorded D.S.O. trace could be limited to 5 μs (compared to a tone burst duration of 3.6 μs), thus allowing the maximum sampling frequency of 100 MHz to be used. For the ''A-line'' measurements, on the other hand, it was necessary to use a trace length of 200 μs corresponding to a depth in water of about 150 mm. This meant that a sampling frequency of only 50 MHz could be achieved giving a total of 10^4 data points.

II. FIELD CHARACTERIZATION MEASUREMENTS

Before carrying out measurements on the wire phantom using the system described above, it was first necessary to characterize the acoustic field around in the focal region of the lens. This was achieved by making two preliminary sets of measurements: one close to the face of the transducer, and one within the region of interest itself. These were carried out in the same tank of water with slight modifications to the basic experimental arrangement. The first set of measurements was used in a numerical model to predict the harmonic pressure levels around the focus; the second set was then compared with the numerical results to check the validity of the model.

Measurements of the amplitude and phase of the acoustic wave field emitted from the source transducer were made using the arrangement shown in Fig. 2. A PVdF needle hydrophone (Precision Acoustics) with a nominal diameter of 75 μm was scanned across the face of the transducer at a range z of 5 mm in the x and y directions using manual and motorized translation stages, respectively. The signal from the hydrophone was passed by way of an external buffer amplifier to the D.S.O., the maximum recorded voltage being of the order of a millivolt (using 50-Ω coupling on the D.S.O.). Once again, the measurements were performed using a toneburst of eight cycles from which a single cycle was extracted and Fourier analyzed. The total duration of the tone burst was sufficiently short to avoid the direct electrical pick-up from the power amplifier that appeared on the signal for $t<4$ μs, but long enough to provide several cycles of effectively cw oscillation.

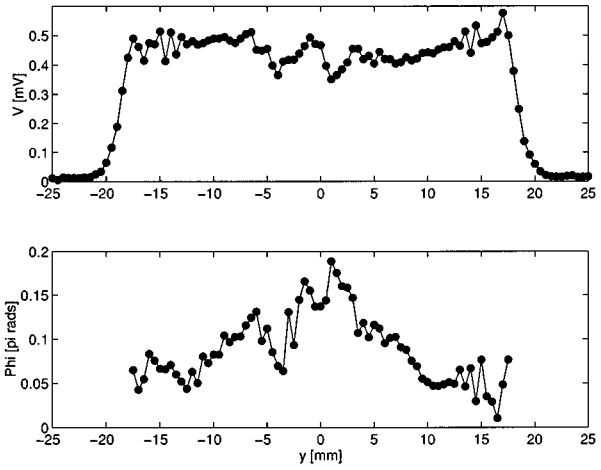


FIG. 3. Horizontal (a) voltage and (b) phase profiles, $V(y)$ and $\phi(y)$, across the center of the transducer face for a nominal source pressure p_0 of 100 kPa. The measurements were made at an axial range z of 5 mm using a PVdF needle hydrophone with a diameter of 75 μm.

Shown in Fig. 3 are the results of a series of measurements for a nominal source pressure p_0 of 100 kPa. The amplitude $V(y)$ and phase $\phi(y)$ of the hydrophone voltage were recorded in a horizontal plane through the center of the transducer face at an interval Δy of 0.5 mm. The phases were determined using the relationship $\phi(y)=\phi_0-2\pi f_0\Delta t(y)$, where f_0 is the drive frequency (2.25 MHz) and $\Delta t(y)$ is the time delay to the beginning of the extracted cycle at the fundamental frequency. The value of ϕ_0 (26.9π rads) was chosen so as to compensate roughly for the time required for the extracted cycle (i.e., the sixth of eight) to propagate out from the transducer to $z=5$ mm.

Looking at Fig. 3(a), it is clear that the amplitude of the emitted acoustic field varies by at least $\pm 5\%$ over the width of the transducer face. Moreover, although not so clear, there is a ''hot spot,'' centered on the acoustic axis, that extends out to about $|y|=3$ mm. The existence of this feature is probably linked to the presence of the electrical connection on the rear side of the transducer element. There is also a noticeable rippling effect close to the edge of the transducer face. This is due to interference between the edge and direct acoustic waves and is less prominent nearer the center of the transducer since the finite size of the hydrophone leads to spatial averaging. The effective radius a of the active element, as measured from the profile of Fig. 3(a), is slightly less than the nominal value of 19 mm. In fact, from comparisons of the positions of axial maxima and minima predicted by linear theory with measurements made at low drive levels, the value of a was found to be 18.62 mm, around 2% less.

The phase profile of Fig. 3(b) also differs significantly from what would be expected for an ideal plane-piston radiator which would have constant phase across its face. First there appears to be a linear phase ramp of about 0.05π radians across the transducer diameter. Such a linear ramp can be ignored since it corresponds to a slight misalignment between the acoustic axis of the transducer and the translation stage. Second, the measured phase peaks at the transducer center ($y=0$). This phase advance (about 0.15π radians) is equivalent to a slight defocusing of an ideal transducer, the effective focal length of the system being -3.5 m and the gain being -0.5.

Having characterized the source transducer in this way, it was then possible to calculate the pressure distribution around the focus of the lens in the main experimental arrangement of Fig. 1 using a numerical model. The model was based around a finite-difference solution of the KZK nonlinear parabolic wave equation for a circularly symmetrical source, and took into account the effects of (i) diffraction, (ii) nonlinearity, and (iii) absorption.[11,12] Solutions for the amplitude and phase of the first four harmonics were obtained in the frequency domain along the acoustic axis of the source and transversely at three axial distances. The parabolic approximation provided meaningful solutions for values of ka much greater than unity, k being the wave number of the source, as long as the axial range z was sufficiently large. More precisely the parabolic approximation becomes increasingly inaccurate as z falls below a critical range given by $a(ka)^{1/3}$ (about 100 mm in this case).

The calculations were performed in two main stages. First, the acoustic wave emitted from the transducer was allowed to propagate linearly to the polymer lens at an axial range z_0 of 50 mm. Second, the extra phase curvature due to its passage through the lens was added, and the wave was allowed to propagate nonlinearly toward the focal region at $z=262$ mm measured relative to the lens. This procedure was used because it was not at the time possible to include the amplitude and phase variations across the higher harmonic beams during the initialization of the numerical model. Although an error was introduced by the assumption of linear propagation ahead of the lens, it was fairly small for a source pressure p_0 of 100 kPa, the second harmonic being almost 30 dB below the fundamental at $z=50$ mm.

The model was initialized at the transducer face using a pair of idealized amplitude and phase profiles, $p(r/a)$ and $\phi(r/a)$, based on the measurements made using the needle hydrophone. These are shown in Fig. 4 as a function of the normalized horizontal distance y/a across the surface of the device. The amplitude ''hot spot'' for $r<3$ mm was modeled by adding a Gaussian to the standard top-hat distribution with a height equal to one-eighth that of the top hat itself and a standard deviation σ of 1 mm. The phase variation across the surface was much slower, and was modeled using a Gaussian with a value for σ of 8 mm. A constant value was subtracted from the resulting profile to give $\phi=0.15\pi$ radians at $r=0$ and $\phi=0$ at $r=a$. The nonspherical shape of the initial wavefront meant that the effect of the phase variation was not equivalent to a simple (spherical) defocusing, as implied previously. As a consequence it was found that the axial minima had finite values instead of falling to zero.

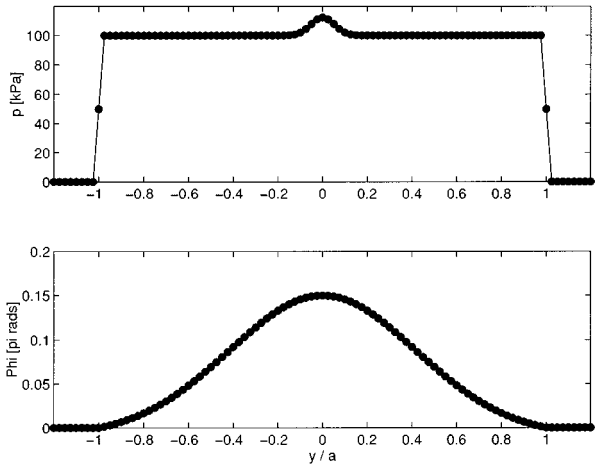


FIG. 4. Idealized (a) amplitude and (b) phase profiles, $p(y/a)$ and $\phi(y/a)$, used to initialize the program, where a is the effective source radius.

For the propagation from the transducer to the lens the source pressure p_0 was set at 1 kPa, while the value of p_0 in the nonlinear region was increased to the ''correct'' value of 100 kPa. Furthermore, the aperture size at the source was the effective radius a of the transducer (i.e., 18.62 mm), whereas the value at the start of the nonlinear stage was the physical radius ''a'' of the lens (i.e., 22 mm). Since the radial step size was held at $a/40$ during both stages of the calculation (e.g., as in Fig. 4), it was necessary to rescale the linear radial profile at the lens before it could be applied to the nonlinear stage. This was achieved by interpolating the linear profile at $z_0=50$ mm for $r\leqslant 22$ mm using a cubic spline. The axial step size was increased from 0.25 mm in the linear region to 0.30 mm in the nonlinear region in order to maintain a suitable ratio of radial to axial step size in the numerical model.

The Rayleigh distances R_0 in the first and second parts of the calculation were 1.65 and 2.30 m, respectively, and the focal gain G in the region after the lens was 8.79. All other parameters required in the model related to the acoustical and other properties of the host medium and were assigned values suitable for water at room temperature.[13,15]

The results obtained using the numerical model were compared with measurements made using a second PVdF membrane hydrophone, this time with a radius b of only 0.5 mm. The experimental arrangement used in this case is shown in Fig. 5. The small-area hydrophone was attached to the motorized translation stages in place of the test object [in Fig. 1(a)] enabling it to be scanned in a horizontal plane through the acoustic axis of the transducer. As during the measurements on the phantom, the large-area hydrophone and lens were positioned coaxially with the transducer at $z_0=50$ mm. Again, measurements were made under pseudo cw conditions, a tone burst of 16 cycles with $f_0=2.25$ MHz being applied to the transducer. Three transverse profiles were obtained: at $z=D-50$ mm$=212$ mm, $z=D=262$ mm, and $z=D+50$ mm$=312$ mm. The spatial resolution of these measurements was $\Delta y=0.5$ mm. The z value for the first transverse profile coincided roughly with the measured posi-

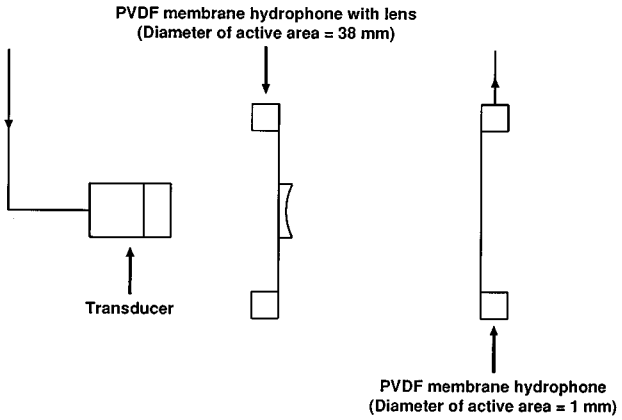


FIG. 5. Diagram of the arrangement used to measure the distributions of pressure, $p_n(y)$, for the first four harmonics, $n=1$ to 4, around the focal plane of the lens.

tion of the final axial maximum for $p_0=100$ kPa, which occurred at $z=218$ mm.

Presented in Fig. 6 are comparisons of theory and experiment at the three axial ranges described above. Within the mainlobe ($|y|<6$ mm) the agreement between the two sets of data is generally good. Outside this region, however, there are significant differences, particularly for the higher harmonics at the larger values of z. These discrepancies are probably due to edge effects at the polymer lens, partly due to the use of an oversimplified source distribution in the numerical model. Moreover, the low values of calculated axial harmonic pressure in Fig. 6(a) may be attributed (i) to the assumption of linear propagation ahead of the lens and (ii) to a slightly inappropriate choice of effective aperture for the lens (giving a final axial maximum at $z=226$ mm). The experimental results show a slight left–right asymmetry in the field, especially around the first sidelobe of the fundamental. This is not surprising considering the differences observed between the left- and right-hand portions of the transducer face in Fig. 3.

As expected, the FWHM beamwidths w_n in Fig. 6 for $n=1$ to 4 (n being the harmonic number) are less for $z=212$ mm than for $z=262$ or 312 mm, and they increase with increasing distance from the final axial maximum. These observations are summarized in Table I, for which measurements were made direct from the three sets of experimental and theoretical profiles. In going from $z=212$ to 312 mm, for example, the measured fundamental beamwidth w_1 increases from 5.40 to 7.57 mm, or by about 40%. For $n=4$ the focusing is tighter, and the increase is consequently much greater: over 90% in fact. In the focal plane, $z=D$, on the other hand, w_n decreases from 6.35 mm down to 2.74 mm between $n=1$ and $n=4$. Indeed, assuming an inverse power law of the form $w_n/w_1=1/n^m$, we find that $m\simeq 0.61$. Thus the drop-off in beamwidth at the ''-6 dB'' points is slower than might be expected from a simple consideration of the linear far-field beam pattern[21] of a plane-piston source for which $m=1$.

The maximum sound-pressure levels, p_{mn} and p_{sn}, of the mainlobe and the first sidelobe for $n=1$ to 4 and $z=D$ are given in Table II. These relate to a source pressure p_0 of 100 kPa. Here, the term ''sidelobe'' describes the first off-axis maximum for a particular harmonic. (For $n\geqslant 2$ the extra sidelobes generated in the near-field as a result of nonlinear propagation are often referred to as ''fingers.'' [12,14]) The value of p_{m1} in Table II is 432 kPa: this compares to a value for Gp_0 of 879 kPa, G being the geometrical focal gain. The largest part of the difference is due to the formation of the

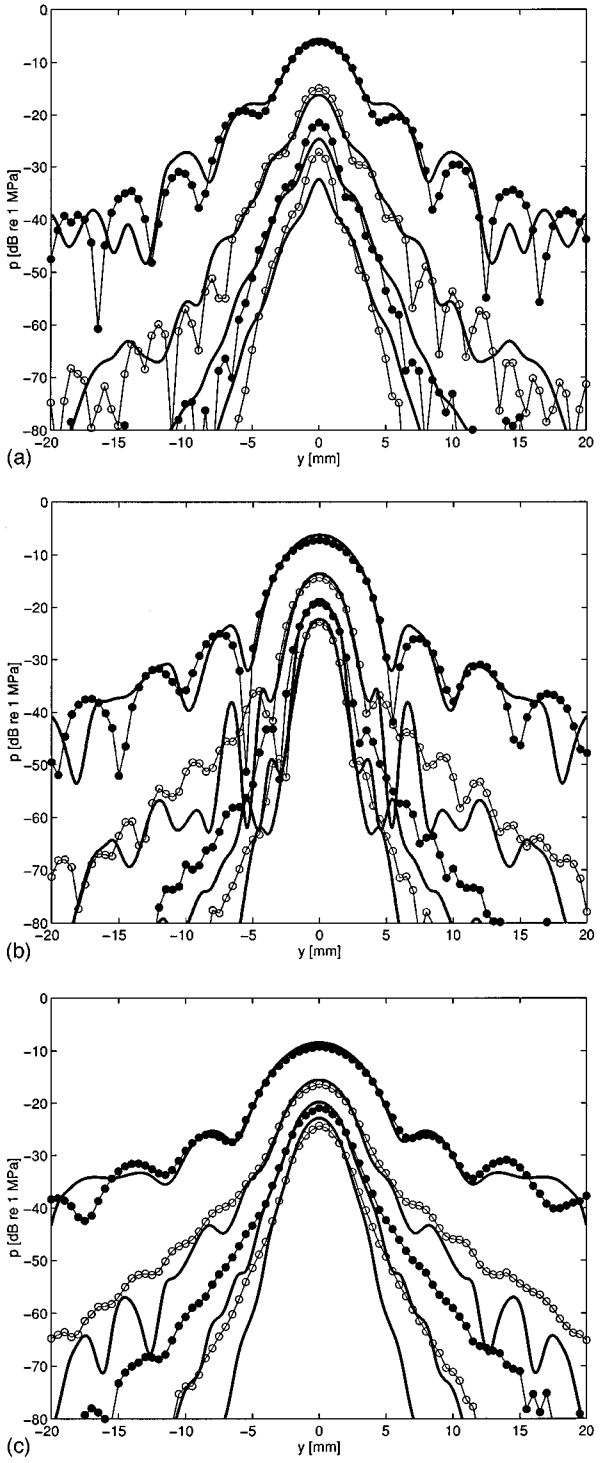


FIG. 6. Comparison of measured and calculated transverse pressure profiles, $p_n(y)$, ahead of the lens at axial ranges z of (a) 212 mm, (b) 262 mm, and (c) 312 mm for a source pressure p_0 of 100 kPa. Thick solid lines: calculated results for $n=1$ to 4 (top to bottom curves, respectively). Filled circles: experimental results for $n=1$ and 3. Empty circles: experimental results for $n=2$ and 4.

TABLE I. Widths (FWHM) w_n of the fundamental ($n=1$) and higher harmonic ($n=2$ to 4) probe beams around the focus of the polymer lens for $p_0=100$ kPa. Values measured from Fig. 6.

		Theory	Experiment			
	z [mm]	$n=1$	$n=1$	$n=2$	$n=3$	$n=4$
w_n [mm]	212	5.67	5.40	3.35	2.51	2.05
	262	6.62	6.35	4.11	3.23	2.74
	312	8.10	7.57	5.21	4.41	3.95
w_n/w_1	212	1	1	0.62	0.46	0.38
	262	1	1	0.65	0.51	0.43
	312	1	1	0.69	0.58	0.52
$w_n(z)/w_n(D)$	212	0.86	0.85	0.82	0.78	0.75
	312	1.22	1.19	1.27	1.37	1.44

higher harmonics, although attenuation at the fundamental frequency also plays a small role. (The attenuation coefficient α of water at room temperature for $f=2.25$ MHz is 1.1 dB m^{-1}.) The high levels of the harmonic axial pressures compared to the fundamental indicate a fairly well developed shock in this region. In fact, a difference of 6 dB between the values of p_{mn} for $n=1$, 2, and 4 would correspond to a sawtooth-type waveform if all the harmonics had been in phase with one another.

An indication of the useful signal level available at each of the four measured harmonic components (for an object at an axial range $z=D$) may be obtained by looking at the simple parameter $p_{dn}=p_{sn}-p_{mn}$, i.e., the difference in pressure between the center of the mainlobe and the peak of the first sidelobe. For $1\leqslant n\leqslant 4$, the measured values of p_{dn} increase steadily with n, although the change in going from $n=1$ to $n=2$ is greatest. (In fact, $p_{d2}-p_{d1}=3.8$ dB, whereas $p_{d4}-p_{d1}=8.2$ dB.) Furthermore, the amplitudes (and widths) of the second and subsequent sidelobes decrease more quickly with increasing distance from the acoustic axis for the higher harmonic beams than they do for $n=1$. This means that images produced with $n\geqslant 2$ will have significantly lower sidelobe levels than those constructed with the fundamental component alone. In addition, what sidelobe images are generated with $n\geqslant 2$ will be confined to a narrower region around the main reflection. On both accounts the clarity of the resulting B-scans should be improved.

III. NONLINEAR IMAGING MEASUREMENTS

As mentioned in Sec. I, the first set of measurements made on the phantom was performed using a similar technique to that employed to characterize the field ahead of the source. A sinusoidal tone burst was applied to the transducer—this time of eight cycles—and the harmonic content of the signal reflected off the object was determined by analyzing a single cycle from the recorded acoustic signal. Three source pressures p_0 were considered: 10 kPa (almost linear), 100 kPa (for comparison with the results of Sec. II), and 400 kPa (the maximum signal it was possible to generate using the apparatus). To obtain the transverse harmonic profiles, $p_n(y)$, 200 waveforms were recorded at intervals Δy of 0.5 mm, allowing all five wires of the front row of the phantom to be included.

TABLE II. Sound-pressure levels, p_{mn} and p_{sn}, of the mainlobe and the first sidelobe for $n=1$ to 4 at the focus of the polymer lens ($z=D$) with $p_0=100$ kPa. The difference between these two values, $p_{dn}=p_{sn}-p_{mn}$, gives an indication of the useful probe signal level. Values measured from Fig. 6.

	Theory	Experiment			
	$n=1$	$n=1$	$n=2$	$n=3$	$n=4$
p_{mn} [dB *re*: 1 MPa]	−6.3	−7.3	−14.3	−19.1	−22.7
p_{sn} [dB *re*: 1 MPa]	−24.0	−25.6	−34.6	−43.4	−49.2
p_{dn} [dB]	17.7	18.3	22.1	24.3	26.5

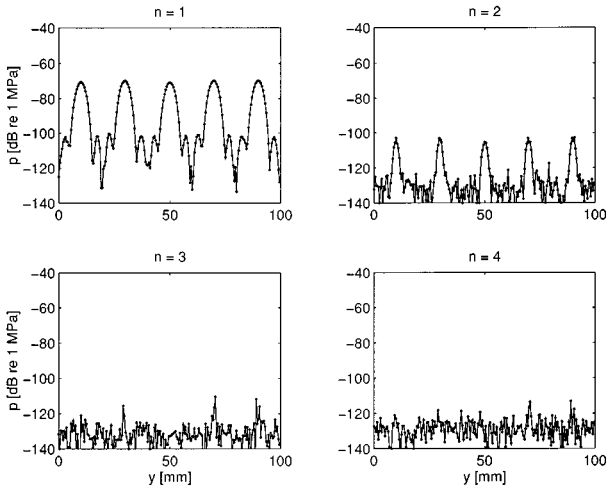


FIG. 7. Reflected signal profiles $p_n(y)$ from the first row of wires for the (a) first, (b) second, (c) third, and (d) fourth harmonics with $z=D=262$ mm and $p_0=10$ kPa.

Results for $p_0=10$ kPa and $n=1$ to 4 are presented in Fig. 7(a) to (d), respectively. The reflections from the five wires, centered on $y=10$, 30, 50, 70, and 90 mm, are clearly visible for $n=1$ and $n=2$. However, for $n\geqslant 3$ the equivalent acoustical noise level ($\sim$0.5 Pa) is of the same order as the peak reflected signals from the phantom. Measurements of beam width w_n and sound pressure level taken from these curves are shown in Tables III and IV, respectively. The width w_1 of the reflected fundamental is 4.6 mm, while that of the second harmonic (w_2) is around half this value. Furthermore, the maximum reflected signal p_{m1} is about 30 dB down on the source pressure p_0, and p_{m2} is a further 33 dB down on p_{m1}. Interpreting the fundamental profile is made much more difficult by the presence of the sidelobes, which lie only $p_{d1}=31$ dB below the peak reflected signal. Despite the relatively low S/N ratio for $n=2$, the positions of the wires are much more clearly defined in the second harmonic profile.

The effect on the reflected signal profiles $p_n(y)$ of increasing p_0 by 20 dB to 100 kPa is shown in Fig. 8. In addition to the second harmonic, the third and fourth harmonics have now increased well above the noise floor of the system, whereas the shape of $p_1(y)$ has remained more or less the same, but has also been shifted up by almost 20 dB. The average widths w_n are less than those for the probe beam, decreasing from 4.6 mm for $n=1$ down to $\sim w_1/3$ for $n=4$. Again, if we assume that $w_n/w_1=1/n^m$, then $m=0.78$

TABLE III. Average widths (FWHM) w_n of the beams reflected from the front row of wires for $n=1$ to 4 and $p_0=10$, 100, and 400 kPa. Values measured from Figs. 7–9.

	p_0 [kPa]	$n=1$	$n=2$	$n=3$	$n=4$
w_n [mm]	10	4.6	2.2	···	···
	100	4.5	2.6	1.9	1.5
	400	4.6	2.7	2.0	1.6
w_n/w_1	10	1	0.48	···	···
	100	1	0.58	0.42	0.33
	400	1	0.59	0.43	0.35

compared to the previously calculated value for the probe beam of 0.61. Hence, not only are the reflected harmonic beams narrower than the equivalent probe beams, they also decrease more rapidly in width with increasing n. The difference p_{dn} between the peak and background signals increases from 31 dB at $n=1$ up to 49 dB at $n=3$ and then decreases back down to 42 dB at $n=4$ as the signal begins to fall back toward the noise floor. For $n=3$ the ''sidelobe'' levels are roughly equal to the random noise level and the system is at its most sensitive.

Raising the source pressure p_0 further to 400 kPa produces the changes shown in Fig. 9. Here, an increase in p_0 of 12 dB has only produced an equivalent increase in the peak reflected pressure p_{m1} of about 8 dB, the remainder of the energy having either gone into the formation of the higher harmonics or been lost through attenuation. The production of the higher frequency components causes a drop in the relative height of the mainlobe compared to the sidelobes for $n=1$, an effect related to the saturation of the fundamental beam at high drive levels. In fact, $p_{d1}\simeq 27$ dB for $p_0=400$ kPa compared to 31 dB for $p_0=10$ and 100 kPa. Thus the B-mode artifacts arising from the fundamental sidelobes become worse as the propagation becomes more and more nonlinear. The FWHM widths of all the harmonic beams for $p_0=400$ kPa are marginally greater than those for $p_0=100$ kPa, the difference presumably being due to the gradual changes in beam shape that occur as the amount of nonlinear generation increases. Again, the maximum useful signal for imaging purposes occurs at $n=3$, but the values of p_{dn} for $n=1$ to 4 are now all less than they were for $p_0=100$ kPa.

TABLE IV. Average sound-pressure levels, p_{mn} and p_{bn}, of the maximum and background signals reflected from the front row of wires for $n=1$ to 4 with $p_0=10$, 100, and 400 kPa. The difference between these two values, $p_{dn}=p_{bn}-p_{mn}$, gives an indication of the signal-to-clutter ratio of the imaging system. Values measured from Figs. 7–9.

	p_0 [kPa]	$n=1$	$n=2$	$n=3$	$n=4$
p_{mn} [dB *re*: 1 MPa]	10	−69.6	−102.7	···	···
	100	−50.0	−64.9	−71.2	−80.1
	400	−42.3	−53.2	−59.1	−66.2
p_{bn} [dB *re*: 1 MPa]	10	−100.5	−117.1	···	···
	100	−80.5	−104.9	−120.4	−122.1
	400	−68.8	−89.3	−105.5	−107.7
p_{dn} [dB]	10	30.9	14.4	···	···
	100	30.5	40.0	49.2	42.0
	400	26.5	36.1	46.4	41.5

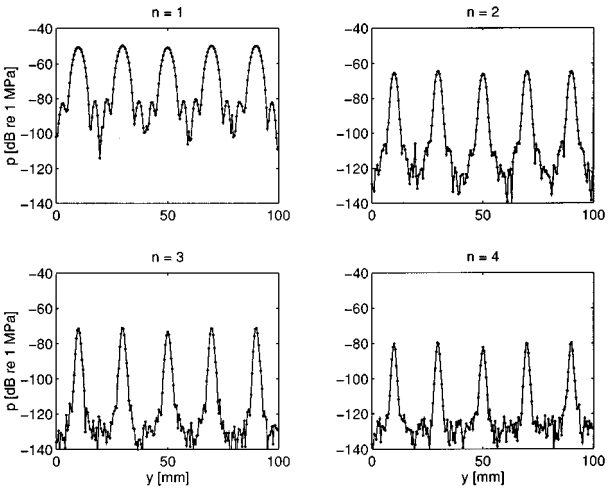


FIG. 8. Reflected signal profiles $p_n(y)$ from the first row of wires for the (a) first, (b) second, (c) third, and (d) fourth harmonics with $z=D=262$ mm and $p_0=100$ kPa.

This reduction in effective harmonic amplitude at high drive levels is linked partly to the amount of nonlinear generation occurring in the water, and partly to its frequency-squared dependence of attenuation (i.e., $\alpha_n/\alpha_1=n^2$, α_n being the attenuation coefficient in water at the nth harmonic).

The second set of measurements on the phantom took the form of a series of rf pulse-echo ''A-lines,'' which were processed digitally and placed side by side to form four B-mode images of the whole object with $n=1$ to 4. A single-cycle tone burst was applied to the transducer giving an average source pressure p_0 of 400 kPa, and waveforms were recorded at intervals Δy of 1 mm. The total length of each trace was 200 μs and the sampling frequency was 50 MHz (see Sec. I). The harmonic components of the signal were obtained by passing the raw data through four digital Chebyshev filters with passbands centered on $f=f_0$, $2f_0$, $3f_0$, and $4f_0$. The magnitude and phase characteristics of these filters, $A_n(f)$ and $\theta_n(f)$, are plotted in Fig. 10(a) and (b), respectively. The width of each passband was set to 1.5 MHz so as

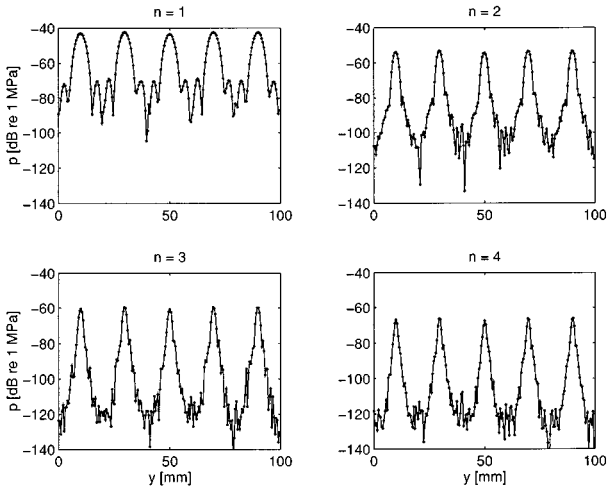


FIG. 9. Reflected signal profiles $p_n(y)$ from the first row of wires for the (a) first, (b) second, (c) third, and (d) fourth harmonics with $z=D=262$ mm and $p_0=400$ kPa.

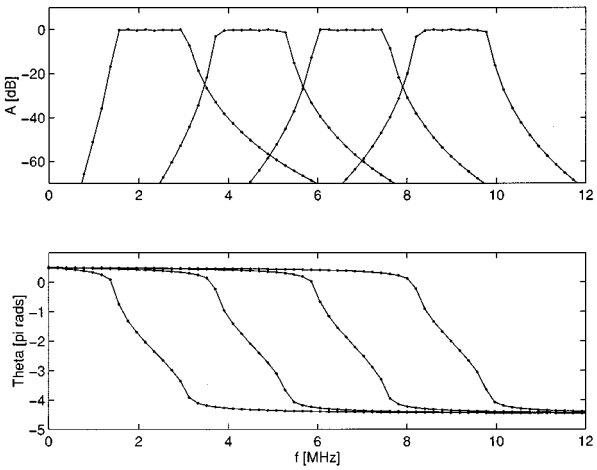


FIG. 10. (a) Magnitude and (b) phase responses, $A_n(f)$ and $\theta_n(f)$, of the four Chebyshev digital bandpass filters used to extract the first four harmonics from the captured waveforms.

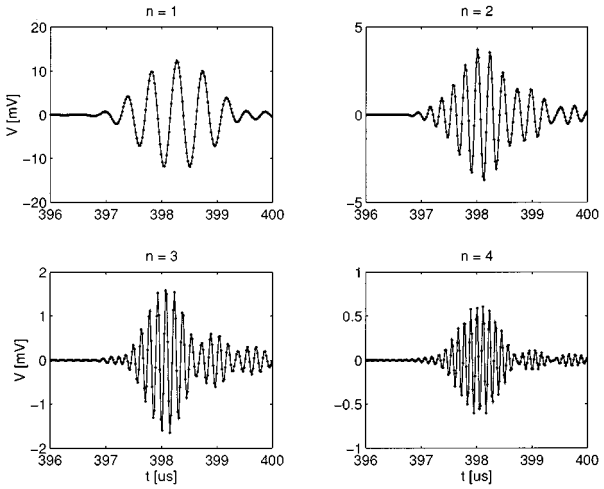


FIG. 12. Harmonic traces of the signal reflected from the unobstructed front-row wire [i.e., filtered versions of Fig. 11(a)] for (a) $n=1$, (b) $n=2$, (c) $n=3$, and (d) $n=4$, where n is the harmonic number.

to include most of the frequency components arising from the single-cycle acoustic pulse, and the ripple on top of each passband was limited to 0.5 dB. The calculated phase responses $\theta_n(f)$ of the digital filters are almost identical with respect to the four center frequencies meaning that the time delays $\Delta t_n=(1/2\pi)(d\theta_n/df)$ caused by their introduction were approximately equal. In fact, $\Delta t_n=(0.84\pm0.05)$ μs for $n=1$ to 4. Therefore, in order to obtain correct axial ranges z for the B-scan, where $z=c_0t$ and $c_0=1471$ ms^{-1} (for $T=16.5$ °C), the filtered waveforms $p_n(t)$ were shifted backward in time by the values Δt_n.

Plotted in Fig. 11(a) is a 4-μs portion of one of the unprocessed ''A-line'' waveforms described above for $p_0=400$ kPa. This corresponds to the point $y=30$ mm on the cw transverse profile of Fig. 9, i.e., a reflection off one of the unobstructed wires in the front row with $z=D$. Clearly, the acoustic pulse is more than a single cycle in length, partly due to its interaction with the object, but mainly due to the limited bandwidth of the transducer (about 600 kHz). The top–bottom asymmetry of the distorted waveform is the most obvious feature arising from the nonlinear propagation, the ratio of the peak positive and negative voltages being about 7:4. Also evident, however, is the steepening of the wavefronts themselves, particularly around the center of the pulse where the amplitude is greatest. The Fourier transform of the $V(t)$ curve of Fig. 11(a) is given in Fig. 11(b). It is clear from this diagram that the acoustic pulse is rich in harmonic content up to at least $n=8$.

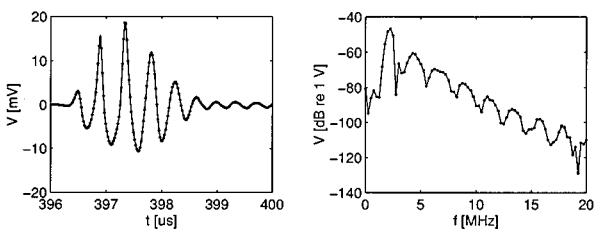


FIG. 11. Maximum signal reflected from one of the unobstructed front-row wires with $z=D=262$ mm and $p_0=400$ kPa. (a) 4-μs portion of the raw unfiltered A-line. (b) Frequency spectrum of the same portion of the A-line.

Filtered versions of the time waveform for $n=1$ to 4 over the same range of t values are shown in Fig. 12(a) to (d), respectively. As expected, each waveform has suffered a delay of approaching a microsecond compared to the raw data, and the amplitude drops off rapidly with increasing n (note the scale changes). Of particular interest is the decrease in length of the pulse envelope for the higher harmonics. This is most likely due to the fact that the greatest harmonic generation occurs where the fundamental pressure is highest. Whereas the duration (FWHM) of the envelope for $n=1$ is about 1.6 μs, it has dropped off to about 0.9 μs by $n=4$. Thus the effective axial resolution of the pulse-echo arrangement has been improved by almost a factor of 2 between $n=1$ and 4. This is in addition and separate to the improvement in lateral resolution obtained from the expected reduction in beamwidth with increasing frequency.

In commercial medical scanning devices, B-mode images are constructed from the raw rf ''A-lines'' by first rectifying the analogue voltage, then performing some kind of envelope detection upon it, e.g., low-pass filtering. In our system, we achieve a similar result by taking the absolute values of all the sampled points within each waveform and then averaging them in blocks of ten, so reducing the length of each trace from 10 000 to a more manageable 1000. The total length of the scan in the y direction was 100 mm and the spatial resolution Δy was 1 mm giving a B-mode grid of 100×1000 points for each value of n. The resulting B-mode voltages, $|V_n(y,z)|$, were displayed on a decibel scale, the reference value for each harmonic being taken as the maximum over the 100 processed A-lines, and any values more than 60 dB below this maximum being ignored.

The results of this procedure for $n=1$ to 4 are given in Fig. 13(a) to (d), respectively. Here, the transducer is off to the left of each image at $z=-50$ mm, and the PMMA lens is between the transducer and the object at $z=0$. The vertical lines extending across the whole ''width'' of each B-scan at $z\simeq250$, 300, and 350 mm are due to multiple reflections between the transducer and hydrophone, and thus decrease exponentially in amplitude with increasing z. The positions

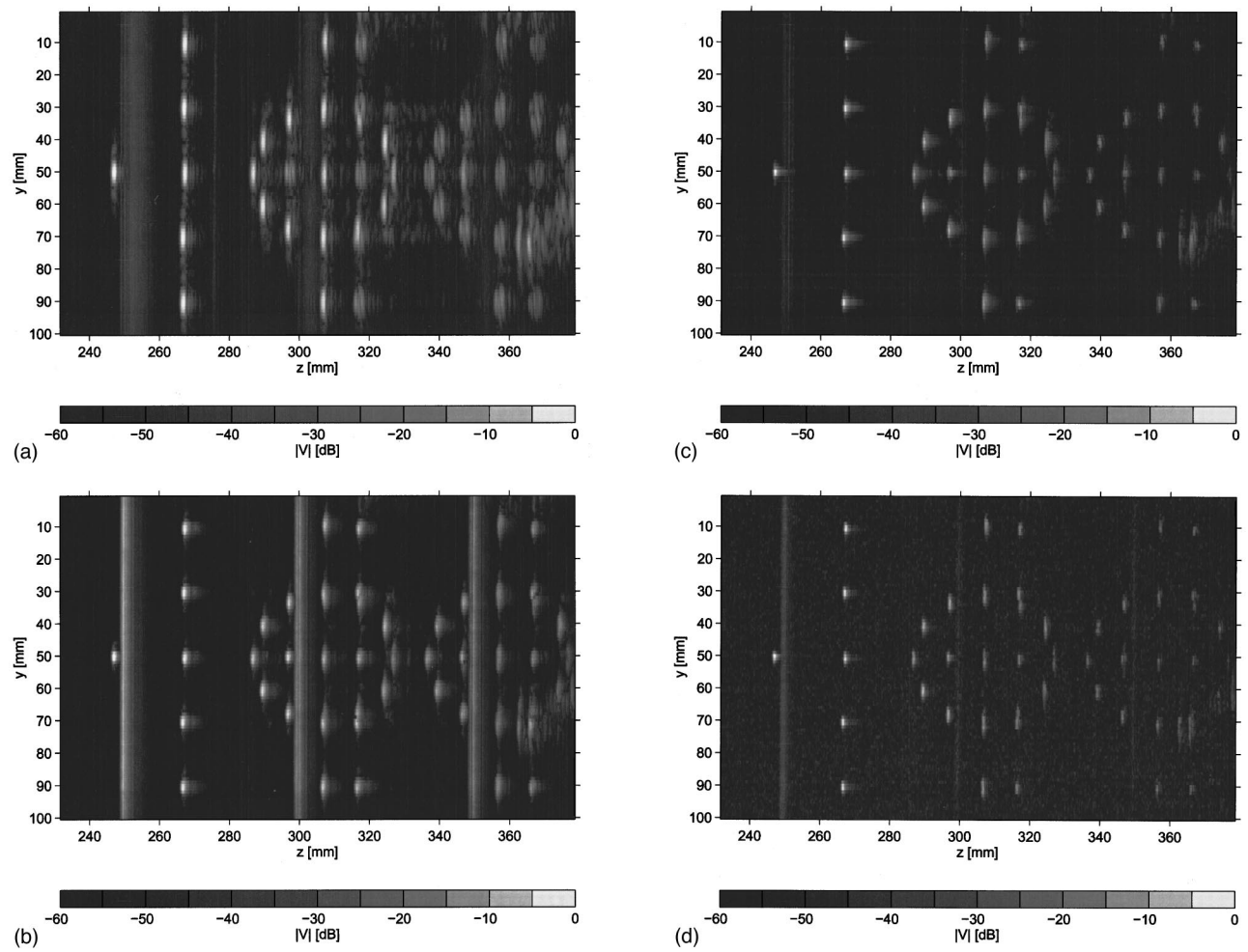


FIG. 13. B-mode images of the wire phantom at the (a) first, (b) second, (c) third, and (d) fourth harmonics. The source pressure p_0 is 400 kPa and the phantom is positioned so that the first row of wires lies at $z=D=262$ mm.

of the wires in the phantom are clearly distinguishable when compared to the schematic of Fig. 1(b), and in particular the front row of wires may be seen near $z=D=262$ mm. The ''ghost'' images of the wires that occur a distance Δz of 50 mm to the right of the ''direct'' images also arise from a single extra reflection off the transducer, but this time after the acoustic signal has been reflected off the object. The bright feature around $(y,z)=(75$ mm, 365 mm) is a clamp used to hold the phantom in place in the tank.

Ignoring the various unwanted reflections in Fig. 13(a), the most striking thing about the fundamental B-scan is the presence of the distinct sidelobes adjacent to each of the ''main'' reflected lobes. Although the S/N ratio between the ''light'' and ''dark'' areas around $z=D$ is at least 60 dB, the difference between the mainlobe and the sidelobes is only 31.6 dB. This compares to a value for p_{d1} from the cw transverse profiles of 26.5 dB (see Table IV). Thus despite the fact that the sidelobes on the pulse-echo B-mode image are about 5 dB lower than they are on the cw profile, they still cause a significant degradation of image quality in the y direction. Another interesting feature of the B-scan is the gradual widening of the ''real'' reflections with increasing distance from the final axial maximum at $z=218$ mm. This is in accordance with the changes in width of the mainlobe of the probe beam for $n=1$ in Fig. 6.

The equivalent B-scan for $n=2$ [Fig. 13(b)] shows significant improvements on that for the fundamental, both in terms of spatial resolution and background clutter reduction. Indeed, the reflected beamwidth w_2 in the focal plane has decreased by an amount similar to that seen for the cw profile of Fig. 9(b), while the FWHM ''duration'' of each of the reflected pulses has also become slightly less. On the negative side, the ''tail'' that appears behind each wire's image in the z direction still has a length roughly equal to that for $n=1$, and the relative levels of the ''ghost'' reflections are somewhat higher. This last effect would not be a problem in a real broadband system, however, since the separation z_0 of the transducer and hydrophone would probably be much less than 50 mm, and possibly even be zero.

Considering next the third and fourth harmonics [Fig. 13(c) and (d), respectively], the lateral beamwidths w_n and relative sidelobe levels p_{dn} continue to decrease in a manner similar to that seen in the pseudo cw measurements. Due to the much diminished peak pressures, however, the random background noise on the B-scans begins to rise well above the -60-dB cutoff level: for $n=3$ and 4 the noise is about 50

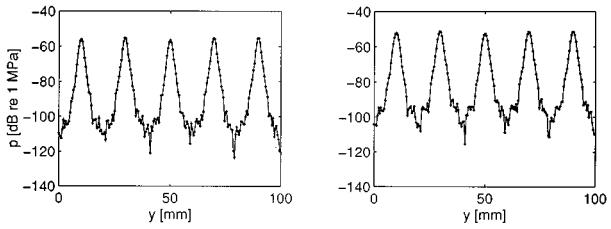


FIG. 14. Geometrical averages of the transverse profiles $p_n(y)$ along the first row of wires for $n=1$ to 4 with $z=D$ and $p_0=400$ kPa. (a) Unweighted average: $\overline{p}=(p_1p_2p_3p_4)^{1/4}$. (b) Weighted average: $\overline{p}=(p_1^4p_2^3p_3^2p_4^1)^{1/10}$.

and 40 dB beneath the maximum values, respectively. However, looking between the wires of the front row for example, it is clear that the useful signal available to the imaging system in this area is still greater than that for $n=1$ since the sidelobes have now fallen back into the noise. Although the background S/N ratio for $n=1$ is much better than that for $n\geqslant 2$, the presence of the large sidelobes serves to reduce the useful dynamic range in the lateral direction by almost 30 dB. Also evident in the B-scans for $n=3$ and 4 is the tighter focusing of the higher harmonic beams compared to that of the fundamental beam. This effect shows how the greatest improvements in lateral beamwidth for the higher frequency components are bound to be confined more closely to the region around the focal plane of the probe beam.

In a practical imaging system the extra information received from the higher harmonics could be either (i) accessed individually or (ii) incorporated into a single compound B-scan using colour to code the information. One of the simplest and quickest ways of combining the data would be to multiply the values of $|V_n(y,z)|$ together for $n=1$ to 4 and then to take the fourth root of the result, so forming a geometrical average. This would have the advantage of bringing some of the high-resolution information from the higher frequency components into the output, while also maintaining the good S/N ratio observed for the fundamental. In Fig. 14(a) this procedure is applied in 1-D to the transverse profiles $p_n(y)$ for $p_0=400$ kPa (Fig. 9). Here, the reflected beamwidth $\overline{w}$ is 2.3 mm, and the useful signal level $\overline{p_d}$ is 41.6 dB, both of these values lying about midway between the equivalent values for $n=2$ and $n=3$. Comparing the combined and original profiles, however, it is clear that the process of averaging has produced a definite improvement in the smoothness of the peaks, as well as causing a reduction in the fluctuations in the background signal. These observations are confirmed in the averaged 2-D B-mode image for $p_0=400$ kPa of Fig. 15(a).

Another slightly more sophisticated method of bringing together information from the four harmonics would be to perform a weighted geometrical average on the data for $n=1$ to 4. In this way it would be possible to place more emphasis on the lower, more reliable values of n, particularly when the levels of the higher harmonics are very low. Presented in Fig. 14(b) is the averaged transverse profile $\overline{p(y)}=(p_1^4(y)p_2^3(y)p_3^2(y)p_4^1(y))^{1/10}$ with $\overline{w}=2.8$ mm and $\overline{p_d}=38.0$ dB. The equivalent averaged B-scan is shown in Fig. 15(b). The values of $\overline{w}$ and $\overline{p_d}$ are now similar to those for $n=2$ in the original profiles since the ''center of gravity''

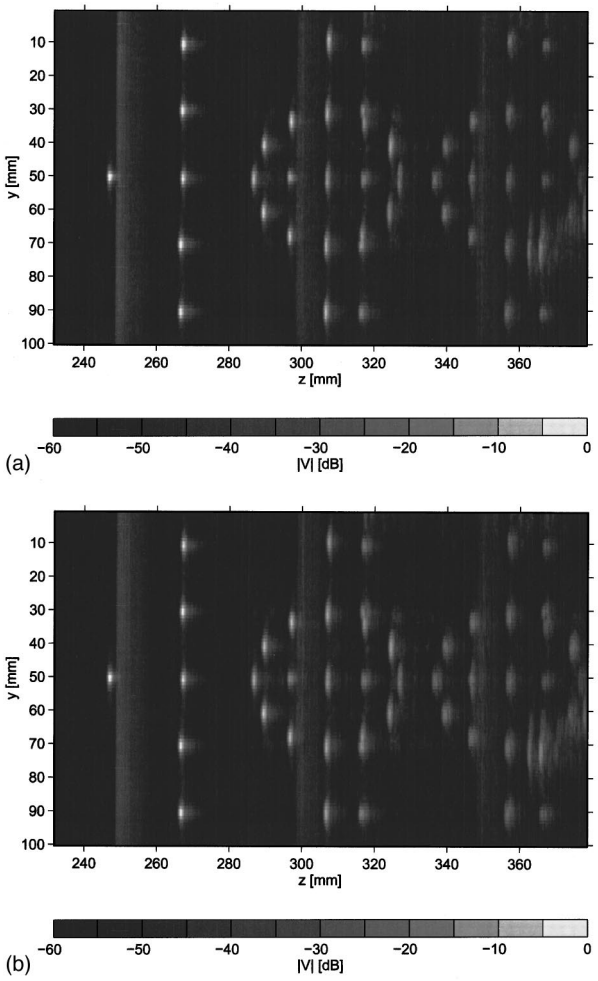


FIG. 15. Geometrical averages of the B-mode images for $n=1$ to 4 with $p_0=400$ kPa. (a) Unweighted average: $|V|=(|V_1\|V_2\|V_3\|V_4|)^{1/4}$. (b) Weighted average: $|V|=(|V_1|^4|V_2|^3|V_3|^2|V_4|^1)^{1/10}$.

of the average has been shifted toward the lower values of n. For this particular imaging arrangement there is very little to choose between the averaged and unaveraged B-mode images, implying perhaps that it would be better to view the higher harmonic B-scans alone. In cases where relatively little nonlinear generation had occurred, however, it might still be better to combine the images in some way.

IV. CONCLUSIONS

Four harmonic B-mode scans of a simple wire phantom in water have been obtained with a broadband imaging arrangement comprising a circular immersion transducer, a membrane hydrophone, and a PMMA lens. The operating frequency of the transducer was at the lower end of the range of values currently used in clinical work, while the source pressure and physical dimensions of the arrangement are comparable with those found in many commercial diagnostic medical ultrasound systems. As such the results of these experiments are relevant to a range of imaging applications, including medical B-mode scanning. Two main types of improvement were noted in going from the fundamental to the higher harmonics: (i) the widths w_n of the images of the wires at the focus decreased according to the relationship

$w_n/w_1 = 1/n^m$, where $m \simeq 0.78$; and (ii) the levels of the ''sidelobes'' on the reflected beams decreased as n went up from 1 to 4, n being the harmonic number. In imaging terminology, these two effects correspond to an enhancement of the lateral resolution of the system and a reduction in the signal-to-clutter ratio. A third but rather less obvious type of improvement was due to the gradual decrease in length of the reflected B-mode pulse echoes with increasing n. This was equivalent to a modest enhancement of the axial resolution of the system.

The greatest improvement in image quality occurred in going from $n=1$ to $n=2$, particularly in terms of lateral resolution. By widening the receive bandwidth of a medical imaging system to include information from a much wider range of frequencies, therefore, it would be expected that significantly more detail would be available to the operator than was previously the case. Although widening the bandwidth would cause the total amount of noise recorded by the system to increase, the S/N ratio of the filtered fundamental signal would be similar to that from a conventional narrowband system. Of course, the relative background noise level for the second harmonic would be higher, but this effect would be countered by the enhanced resolution and sidelobe suppression at this frequency. Further improvements could be achieved by going to the third and higher harmonics, but they would not be as dramatic as those for $n=2$. Information from the recorded harmonics could be presented separately as individual B-scans or combined in some way to produce a single image. Taking a simple geometrical average of the harmonic B-scans would be a convenient and effective means of doing this.

The imaging work described here was limited to water as the host medium, but it should in future be possible to extend it to more realistic tissue-like media. Although the attenuation of ultrasound is at least two orders of magnitude greater in human tissue than in water at 1 MHz, it does increase less rapidly with frequency in tissue. This means that there should still be sufficient high-frequency content in the reflected signal to form B-scans of at least the second harmonic. Experiments are currently underway to confirm this.[22]

The extraction of harmonic information could in the future prove useful in a number of situations in diagnostic medical ultrasound where the quality of the visual information supplied to the clinician is of great concern. At the moment, medical ultrasound systems do not make significant use of any harmonics generated, harnessing only a few of the extra frequency components may allow both the resolution and the clarity of the resulting images to be improved. It should be reiterated that it is not advocated that drive levels are increased to achieve this goal. The technique could also find application in other fields of ultrasonic imaging, e.g., nondestructive testing and side-scan sonar, where significant amounts of nonlinear distortion may also occur.

ACKNOWLEDGMENTS

The authors would like to thank Francis Duck of the Royal United Hospital in Bath for the loan of the wire phantom and for his help and encouragement during this work. The research assistantship for B. Ward was provided by the U.K. Engineering and Physical Sciences Research Council (Grant No. GR/H36368). Funding for the purchase of the transducer and the large-area membrane hydrophone was provided by the Royal Society (Grant No. RSRG 11546). The needle hydrophone was loaned by Precision Acoustics Ltd. of Dorchester in the U.K.

[1] T. G. Muir and E. L. Carstensen, ''Prediction of nonlinear acoustic effects at biomedical frequencies and intensities,'' Ultrasound Med. Biol. **6**, 345–357 (1980).
[2] E. L. Carstensen, W. K. Law, N. D. McKay, and T. G. Muir, ''Demonstration of nonlinear acoustical effects at biomedical frequencies and intensities,'' Ultrasound Med. Biol. **6**, 359–368 (1980).
[3] H. C. Starritt, M. A. Perkins, F. A. Duck, and V. F. Humphrey, ''Evidence for ultrasonic finite-amplitude distortion in muscle using medical equipment,'' J. Acoust. Soc. Am. **77**, 302–306 (1985).
[4] H. C. Starritt, F. A. Duck, A. J. Hawkins, and V. F. Humphrey, ''The development of harmonic distortion in pulsed finite-amplitude ultrasound passing through liver,'' Phys. Med. Biol. **31**, 1401–1409 (1986).
[5] R. Kompfner and R. A. Lemons, ''Nonlinear acoustic microscopy,'' Appl. Phys. Lett. **28**, 295–297 (1976).
[6] D. Rugar, ''Resolution beyond the diffraction limit in the acoustic microscope: a nonlinear effect,'' J. Acoust. Soc. Am. **56**, 1338–1346 (1984).
[7] L. Germain and J. D. N. Cheeke, ''Generation and detection of high-order harmonics in liquids using a scanning acoustic microscope,'' J. Acoust. Soc. Am. **83**, 942–949 (1988).
[8] K. Karaki, T. Saito, K. Matsumoto, and Y. Okuda, ''Nonlinear resolution improvement and second-harmonic generation of a pressurized superfluid ^{4}He acoustic microscope,'' Appl. Phys. Lett. **59**, 908–910 (1991).
[9] T. G. Muir, ''Nonlinear effects in acoustic imaging,'' Acoust. Imag. **9**, 93–109 (1980).
[10] L. Bjørnø and P. A. Lewin, ''Nonlinear focusing effects in ultrasonic imaging,'' in *1982 Ultrasonics Symposium* (IEEE, New York, 1982), pp. 659–662.
[11] S. I. Aanonsen, T. Barkve, J. Naze Tjøtta, and S. Tjøtta, ''Distortion and harmonic generation in the nearfield of a finite amplitude sound beam,'' J. Acoust. Soc. Am. **75**, 749–768 (1984).
[12] M. F. Hamilton, J. Naze Tjøtta, and S. Tjøtta, ''Nonlinear effects in the farfield of a directive sound source,'' J. Acoust. Soc. Am. **78**, 202–216 (1985).
[13] 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).
[14] J. A. TenCate, ''An experimental investigation of the nonlinear pressure field produced by a plane circular piston,'' J. Acoust. Soc. Am. **94**, 1084–1089 (1993).
[15] A. C. Baker, ''Nonlinear pressure fields due to focused circular apertures,'' J. Acoust. Soc. Am. **91**, 713–717 (1992).
[16] A. C. Baker, ''Prediction of non-linear propagation in water due to diagnostic medical ultrasound equipment,'' Phys. Med. Biol. **36**, 1457–1464 (1991).
[17] A. J. Watson, V. F. Humphrey, A. C. Baker, and F. A. Duck, ''Nonlinear propagation of focused ultrasound in tissue-like media,'' in *Frontiers of Nonlinear Acoustics: 12th ISNA*, edited by M. F. Hamilton and D. T. Blackstock (Elsevier Science, London, 1990), pp. 445–450.
[18] L. R. Gavrilov, V. N. Dmitriev, and L. V. Solontsova, ''Use of focused ultrasonic receivers for remote measurements in biological tissues,'' J. Acoust. Soc. Am. **83**, 1167–1179 (1988).
[19] V. M. Levin, O. I. Lobkis, and R. G. Maev, ''Investigation of the spatial structure of acoustic fields by a spherical focusing transducer,'' Sov. Phys. Acoust. **36**, 391–395 (1991).
[20] J. Wu and K. K. Shung, ''Non-linear energy exchange among harmonic modes and its applications to nonlinear imaging,'' J. Acoust. Soc. Am. **88**, 2852–2858 (1990).
[21] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics* (Wiley, New York, 1982), 3rd ed., pp. 178–182.
[22] B. Ward, A. C. Baker, and V. F. Humphrey, ''Non-linear propagation applied to the improvement of lateral resolution in medical ultrasound scanners,'' in *Proceedings of the 1st World Congress on Ultrasonics* (Berlin, 1995), pp. 965–968.

Porous road pavements: Acoustical characterization and propagation effects[a)]

M. C. Bérengier
Laboratoire Central des Ponts et Chaussées, Centre de Nantes, Division Gestion de l'Entretien des Routes, Section Interaction Route/Véhicule et Acoustique Routière, BP n° 19, 44340 Bouguenais, France

M. R. Stinson and G. A. Daigle
Institute for Microstructural Sciences, National Research Council, Ottawa, Ontario K1A 0R6, Canada

J. F. Hamet
Institut National de Recherche sur les Transports et leur Sécurité, Département Modélisation Mécanique et Acoustique, 109 avenue Salvador Allende, Case 24, 69675 Bron Cedex, France

(Received 15 May 1995; revised 10 August 1996; accepted 26 August 1996)

Measurements of the acoustical properties of some porous road pavements are presented here and an acoustical method for monitoring the performance of these surfaces is presented. Porous road pavements have been used previously because of their driving qualities and drainage capacities during rainy days (i.e., the elimination of water splash and spray) but they have also been found to reduce traffic noise substantially. Reductions in A-weighted sound levels of 3–5 dB, compared to a dense pavement structure, have been measured. To study further their acoustical performance, measurements over real road surfaces have been carried out and the results compared to theoretical predictions based upon models describing the surface impedance and sound propagation. For the impedance characterization, both a phenomenological and a microstructural model were used. Both approaches introduce a viscous and a thermal dependence to account for the different phenomena inside the porous structure. By incorporating these models into the theoretical propagation predictions, it is possible to evaluate the impact of porous asphalt on highway noise levels. A nondestructive testing technique has been designed for determining *in situ* the noise reduction performance of porous road pavements. [S0001-4966(97)01101-6]

PACS numbers: 43.28.Fp, 43.50.Gf [LCS]

INTRODUCTION

Porous road pavements, also called drainage asphalts or previous macadams, are new road structures in which discontinuous granular formulation can produce an important void content (porosity) inside the structure. As an example, it is possible to obtain a porosity value of 20% or more using 0- to 10-mm aggregates and a 2- to 6-mm gap grading. Very strong modified binders are used in order to ensure that the aggregates stay together.[1] When compared to the common dense road pavements, these structures have a different acoustical behavior due mainly to their porosity. To account for the changes in noise level, two different mechanisms have to be considered: the modification in the tire/road noise generation[2] and the absorption properties of the porous structure.[3] A new porous pavement can produce a 3- to 5-dB reduction, or more, in A-weighted sound level with respect to nonporous pavements.[4]

This paper is concerned with developing a technique for measuring and for monitoring the acoustical performance of porous asphalt pavements. In Sec. I, the theory for propagation above porous ground is summarized. In Sec. II, two different approaches,[5–8] phenomenological and microstructural, are investigated for the prediction of the characteristic impedance of the porous structures, a quantity required by the propagation model. Measurements of the acoustic propagation above porous structures are compared with predictions, in Sec. III. This work is a first step before making a direct prediction of traffic noise using the new road coating. Recent studies[1] have shown a modification, due to a change of the physical parameters of the porous pavements, in the traffic noise level after a few years of use. A nondestructive testing technique to monitor this deterioration, based on an inverse fitting of transmission spectra values, is discussed in Sec. III.

I. PROPAGATION ABOVE POROUS GROUND

The porous pavement will be treated as a finite layer of extended reaction material. Figure 1 shows the assumed geometry for propagation from a point source. The quantities r_d and r_r represent the source/receiver and image-source/receiver distances, respectively, and ψ is the grazing angle. Assuming an $\exp(-i\omega t)$ time dependence and assuming no refractive profile, the approximate solution[9,10] can be expressed as

$$p = \frac{A_d}{r_d}\exp(ik_d r_d) + \frac{QA_r}{r_r}\exp(ik_r r_r), \tag{1}$$

[a)]Some of the results presented in this paper have already been discussed at the 125th ASA meeting in Ottawa, Canada [J. Acoust. Soc. Am. **93**, 2407 (A) (1993)].

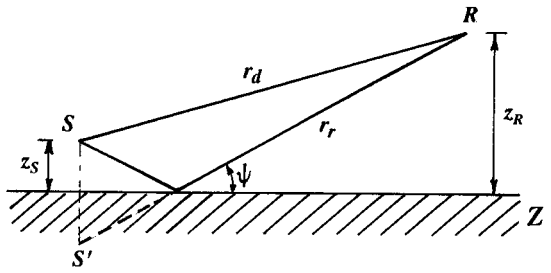


FIG. 1. Sketch showing the geometric definitions for propagation in a homogeneous atmosphere.

where A_d and A_r correspond to the amplitudes of the direct and reflected waves, respectively, and k_d and k_r are the corresponding wave numbers. Here, we are allowing for the effects of turbulence so A_d and A_r are not necessarily equal as they would be for a homogeneous atmosphere nor are k_d and k_r equal. The function Q is the spherical reflection coefficient of the ground,

$$Q=R_p+(1-R_p)F(w). \tag{2}$$

The term R_p is the plane-wave reflection coefficient and the term $(1-R_p)F(w)$ accounts for ground and surface wave components, with $F(w)$ given by

$$F(w)=1+i\pi^{1/2}w\ \exp(-w^2)\mathrm{erfc}(-iw). \tag{3}$$

In Eq. (3), w is a ''numerical distance'' for the extended reaction surface given by

$$w^2=2ik_0r_r\chi^2[Z(1-R_p)]^{-2}, \tag{4}$$

with

$$\chi=[1-(k_0/k)^2\ \cos^2\ \psi]^{1/2}, \tag{5}$$

where Z is the normalized specific surface impedance, k is the complex wave number of the porous pavement, and $k_0=2\pi f/c$, where f is the frequency and c is the homogeneous sound speed in air. Two solutions are possible in taking the complex root for w from Eq. (4). It is necessary[11] to select w so that its complex argument is placed in the range $[-\pi/4,\ 3\pi/4]$. In the case of a nonlocally reacting ground such as porous asphalt, the plane-wave reflection coefficient R_p is given by[12]

$$R_p=(Z\ \sin\ \psi-\chi)/(Z\ \sin\ \psi+\chi). \tag{6}$$

For the typical pavement dealt with in this paper (pervious macadam), the road coating can be considered as a porous structure with a rigid skeleton and a moving fluid (air) inside. For a layer having a thickness l, the surface impedance is given by

$$Z=Z_c\,\frac{Z_T\ \coth(-ikl)+Z_c}{Z_T+Z_c\ \coth(-ikl)}, \tag{7}$$

where Z_c is the characteristic impedance of the road structure and Z_T is the bottom layer impedance. In most cases, Z_T is effectively infinite. With this condition, $Z=Z_c\ \coth(-ikl)$.

A. Homogeneous atmosphere

In the case of a homogeneous atmosphere for which turbulence can be neglected, the direct and reflected wave amplitudes are equal. That is, $A_d=A_r=A_0$, a constant in Eq. (1). Similarly, the wave numbers k_d and k_r are both equal to k_0, for the homogeneous case.

B. Extension to inhomogeneous atmosphere

In a turbulent atmosphere, the amplitudes A_d and A_r and wave numbers k_d and k_r defined in Eq. (1) will fluctuate in time. The fluctuating components of these quantities are defined through

$$A_d=A_0(1+a_d), \tag{8}$$

$$A_r=A_0(1+a_r), \tag{9}$$

$$k_dr_d=k_0r_d+\delta_d, \tag{10}$$

$$k_rr_r=k_0r_r+\delta_r. \tag{11}$$

The components a_d, a_r, δ_d, and δ_r are each assumed to be normally distributed with standard deviations $\langle a_d^2\rangle^{1/2}$, $\langle a_r^2\rangle^{1/2}$, σ_d and σ_r, respectively. The standard deviations can be written[13] as functions of a turbulence strength $\langle\mu^2\rangle$ and a scale of turbulence L.

With the assumption that $\langle a_d^2\rangle=\langle a_r^2\rangle=\langle a^2\rangle$, after some algebra we obtain the following expression for the mean-square pressure:[10]

$$\langle p^2\rangle=A_1p_1^2+A_2p_2^2, \tag{12}$$

with

$$p_1^2=\frac{1}{r_d^2}+\frac{|Q|^2}{r_r^2} \tag{13}$$

and

$$p_2^2=\frac{2|Q|}{r_dr_r}\cos(\phi+\varphi)e^{-\sigma_d^2(1-\rho_d)}, \tag{14}$$

where $\phi=k_0(r_r-r_d)$, $Q=|Q|e^{i\varphi}$, ρ_a and ρ_d are the respective amplitude and phase covariances, $A_1=A_0(1+\langle a^2\rangle)$ and $A_2=A_0(1+\langle a^2\rangle\rho_a)$. Equations (12)–(14) hold for $r_d/r_r\approx1$ and when both r_d and r_r are small enough that the standard deviation of the phase fluctuation along each path is less than about 90°, so that neither amplitude nor phase fluctuations exhibit any limiting phenomena.

For typical atmospheric conditions (wind speed between 0 and 4–5 m/s), $\langle\mu^2\rangle$ can take values between 2×10^{-6} and 20×10^{-6} and L is of the order of 1 m. For a coherent atmosphere, $\langle\mu^2\rangle$ would be identically zero.

II. ACOUSTICAL MODELS FOR POROUS PAVEMENT

In recents years, considerable progress has been made in developing microstructural models to describe the propagation of acoustic waves through porous materials. These models include the cases of propagation through porous materials with an elastic frame[14] and porous materials with a rigid frame.[7,8] The acoustical properties of porous road surfaces is best described by the model of Champoux and Stinson.[8] This is because the pore diameter varies greatly and it can be shown that the viscous and thermal functions must be treated

separately by two pore shape factors. However, the use of this model requires the use of five independent parameters to characterize the material.

Alternatively, the phenomenological approach considers the porous medium to be globally a compressive fluid where dissipations occur. The mathematical expressions are simple enough that an analysis of the influence of the various physical parameters on the impedance value can be performed. The approach introduces a viscous dissipation which is due to the velocity gradient within the porous medium and a thermal dissipation which is due to the thermal gradients within the same medium. The thermal dissipation is particularly important when the airflow resistance is small, i.e., when the structure is highly porous. Recently, an extended phenomenological model has been implemented by Hamet *et al.*[6] The phenomenological model requires only three parameters to characterize the material.

In this section, both microstructural and phenomenological models are summarized. In the following section, it will be shown that in the case of porous pavement the phenomenological model is in close agreement with the microstructural model. Therefore although the microstructural model provides a more complete physical description of porous road pavement, the phenomenological model provides a simpler description which is sufficient for the purpose here.

The acoustical behavior of a rigid-frame porous material is completely characterized by specification of its characteristic impedance Z_c and its complex wave number k. In developing theoretical models to predict these quantities, it is convenient to treat viscous and thermal effects within the pores separately; these effects are described by complex dynamic density ρ_g and bulk modulus K_g functions, respectively. We have, generally,

$$k=\omega[\rho_g(\omega)/K_g(\omega)]^{1/2}, \tag{15}$$

$$Z_c=[\rho_g(\omega)K_g(\omega)]^{1/2}/\Omega, \tag{16}$$

with Ω being the porosity of the material. The roots of Eqs. (15) and (16) are chosen so that the real parts of k and Z_c are positive. In the following, we will be assuming an $\exp(-i\omega t)$ time dependence so appropriate changes have been made to the expressions taken from references which assumed an $\exp(+i\omega t)$ dependence.

A. Phenomenological model

According to the phenomenological model,[6] the dynamic density and bulk modulus functions are given by

$$\rho_g(\omega)=\rho_0 q^2(1+if_\mu/f\ ), \tag{17}$$

$$K_g(\omega)=\gamma P_0[1+(\gamma-1)/(1-if/f_\theta)]^{-1}, \tag{18}$$

where the functions f_μ and f_θ describe the viscous and the thermal dependencies, respectively. They are given by

$$f_\mu=\Omega R_s/(2\pi\rho_0 q^2) \tag{19}$$

and

$$f_\theta=R_s/(2\pi\rho_0 N_{\rm pr}). \tag{20}$$

In the above equations, P_0 is the ambient atmospheric pressure, ρ_0 is the density of air, γ is the specific heat ratio, $N_{\rm pr}$ is the Prandtl number, R_s is the airflow resistivity of the porous structure, Ω is the porosity of air-filled connected pores, and q^2 is the tortuosity[15] (also called the structural factor in Ref. 6). These last three parameters can be independently determined, directly or indirectly, which makes this simple model attractive. The expressions in Eqs. (17) and (18) are similar to those obtained in Wilson's relaxation model.[16]

Introducing the functions

$$F_\mu=1+if_\mu/f, \tag{21}$$

$$F_\theta=1+if_\theta/f, \tag{22}$$

then, from Eqs. (15) and (16), the complex wave number and the characteristic impedance may be written as

$$k=k_0 q F_\mu^{1/2}[\gamma-(\gamma-1)/F_\theta]^{1/2}, \tag{23}$$

$$Z_c=(\rho_0 c q/\Omega)F_\mu^{1/2}[\gamma-(\gamma-1)/F_\theta]^{-1/2}. \tag{24}$$

B. Microstructural model

According to the microstructural approach developed by Champoux and Stinson,[8] the dynamic density ρ_g is

$$\rho_g(\omega)=\rho_0 q^2+iR_s\Omega F(\lambda_\rho)/\omega, \tag{25}$$

with

$$F(\lambda_\rho)=-\frac{1}{4}\frac{\lambda_\rho\sqrt{i}T(\lambda_\rho\sqrt{i})}{1-2T(\lambda_\rho\sqrt{i})/(\lambda_\rho\sqrt{i})}, \tag{26}$$

where

$$T(\xi)=J_1(\xi)/J_0(\xi) \tag{27}$$

is the ratio between Bessel functions of first and zero order. The dimensionless parameter λ_ρ may be obtained using:

$$\lambda_\rho=s_\rho[8q^2\rho_0\omega/(R_s\Omega)]^{1/2}. \tag{28}$$

In Eq. (28), s_ρ is an adjustable parameter which is connected, for real granular structure with a complex geometry, to the viscosity dependence inside the material.

The thermal dependence is accounted for by the bulk modulus

$$K_g(\omega)=\gamma P_0[1+2(\gamma-1)T(\Lambda_k)/\Lambda_k]^{-1}, \tag{29}$$

with

$$\Lambda_k=N_{\rm pr}^{1/2}\lambda_k\sqrt{i} \tag{30}$$

and

$$\lambda_k=s_k[8q^2\rho_0\omega/(R_s\Omega)]^{1/2}. \tag{31}$$

The adjustable parameter s_k is the thermal pore shape factor.

Given the functions ρ_g and K_g, the characteristic impedance and the complex wave number are obtained according to Eqs. (15) and (16). We note that application of the microstructural model requires the knowledge of five parameters, the porosity Ω, the airflow resistivity R_s, the tortuosity q^2, and the viscous and thermal pore shape factors, s_ρ and s_k, respectively.

C. Evaluation of impedance model parameters

In order to compare the two different approaches, the physical parameters, (R_s,Ω,q^2) in the case of the phenomenological model and $(R_s,\Omega,q^2,s_\rho,s_k)$ in the case of the microstructural model, must be specified. Two of the common parameters (R_s and Ω) were measured directly on a sample while the third (q^2) was obtained indirectly from the absorption measurements discussed below. The remaining parameters (s_ρ,s_k) were found not to affect the predictions of the microstructural model too seriously so two sets of values were simply assumed. Once specified, these parameters can be used to calculate the impedance and propagation constant which are needed for calculations of sound-pressure levels.

Core samples of diameter 10 cm and thickness $l=4$ cm were extracted from actual pavement. One of these was used for measurement of the different physical parameters. The aggregates (i.e., the solid component of the sample) were approximately 10 mm in size with gaps of the order of 2–3 mm.

The porosity of the sample was obtained using vertical gamma ray dosimetry.[17] The apparatus consists of a gamma ray source and a detector mounted on the same vertical axis and rigidly connected together. The sample is placed in between source and detector. Gamma rays are absorbed by the solid component of the sample only so knowing the total volume of the sample, the porosity due to the air-filled voids can be obtained. It is noted that this measurement for porosity *includes* the voids that are obstructed and do not communicate with the exterior. In the case of large aggregates such as porous pavement with aggregate sizes greater than 10 mm, the noncommunicating voids represent less than 5% of the total volume. Therefore, the porosity obtained by this method is expected to be nearly the same as the porosity obtained by other techniques[18] which do not include these noncommunicating voids. For the sample under consideration, $\Omega=15\%$ was obtained.

The airflow resistance is measured following the technique described in ISO 9053.[19] The pressure difference across a sample and the flow rate through the sample are measured, with the ratio of these quantities giving the flow resistance R. For a sample cross-sectional area S and thickness l, the flow resistivity is

$$R_s=RS/l. \tag{32}$$

For the sample under consideration, $R_s=5000$ N s m^{-4}. Note that typical nonporous asphalts would have much higher flow resistivities.[20]

The absorption coefficient was obtained using two different approaches. For the first approach, the sample was mounted in an impedance tube and the surface impedance measured. The magnitude r of the reflection coefficient was calculated and used to obtain the absorption coefficient

$$\alpha=1-r^2. \tag{33}$$

The results of this procedure are shown by the solid points in Fig. 2. For the second approach, the absorption coefficient is obtained *in situ* using an impulse technique.[21] The impulse of an 8-mm pistol located 2 m above the surface was measured under normal incidence with a microphone at 0.5 m. This allowed the separation of direct and reflected pulse. The signals were filtered, Fourier transformed, and the ratio of the two Fourier transforms taken to give the reflection coefficient. Since the technique only yields the amplitude of the reflection coefficient at normal incidence accurately, the phase component is not used. For the patch of porous pavement considered, the measurement gave the dashed curve in Fig. 2. The close agreement between the results obtained by the tube technique and the free-field approach suggests relative homogeneity of the surface. This is not generally true in the case of well-used surfaces. The homogeneity found here is due to the virgin state of the surface.

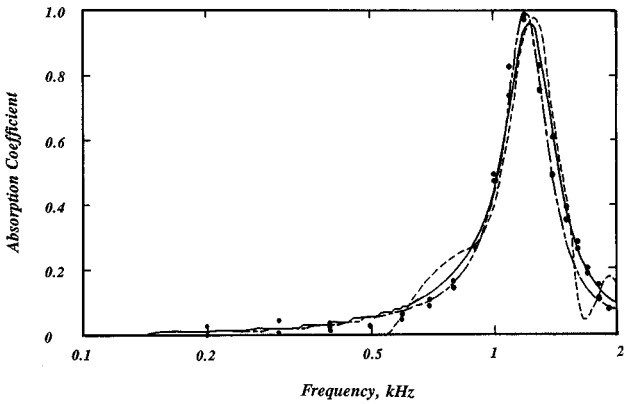


FIG. 2. Absorption coefficient results obtained by different approaches for a 10 cm diameter and 4-cm-thick porous pavement sample. The measured physical parameters are $R_s=5000$ N s m^{-4}, $\Omega=15\%$, $q^2=2.5$, and $s_p=s_k=1$. The solid points are measured results obtained using a standing wave tube, ---- measurement results using an impulse technique,[21] — theoretical prediction using the phenomenological model,[6] and –-– theoretical prediction using the microstructural model.[8]

Tortuosity can be measured directly.[22] However, given the large aggregate sizes of the current pavement, it would not have been possible to cut a sufficiently small sample for mounting in our apparatus. For this reason, the tortuosity was obtained by a fit to the absorption curve. The phenomenological model was assumed and the complex wave number and characteristic impedance calculated from Eqs. (23) and (24). Then with Eqs. (7), (6), and (33), the absorption coefficient was calculated. The position of the resonant peak in Fig. 2 is significantly influenced by the choice of q^2. The solid curve shown in Fig. 2 is obtained from the phenomenological model using the best fit value of $q^2=2.5$.

D. Comparison of impedance models

The absorption coefficient has been calculated using the two impedance models and shown with the experimental data in Fig. 2. The solid curve shows the prediction using the phenomenological model and the dot-dashed curve shows the prediction using the microstructural model. Two additional parameters, the viscous and thermal pore shape factors, s_ρ and s_k, were required for application of the microstructural model. For this initial calculation, s_ρ and s_k have been set equal to unity. A slightly better fit can be obtained

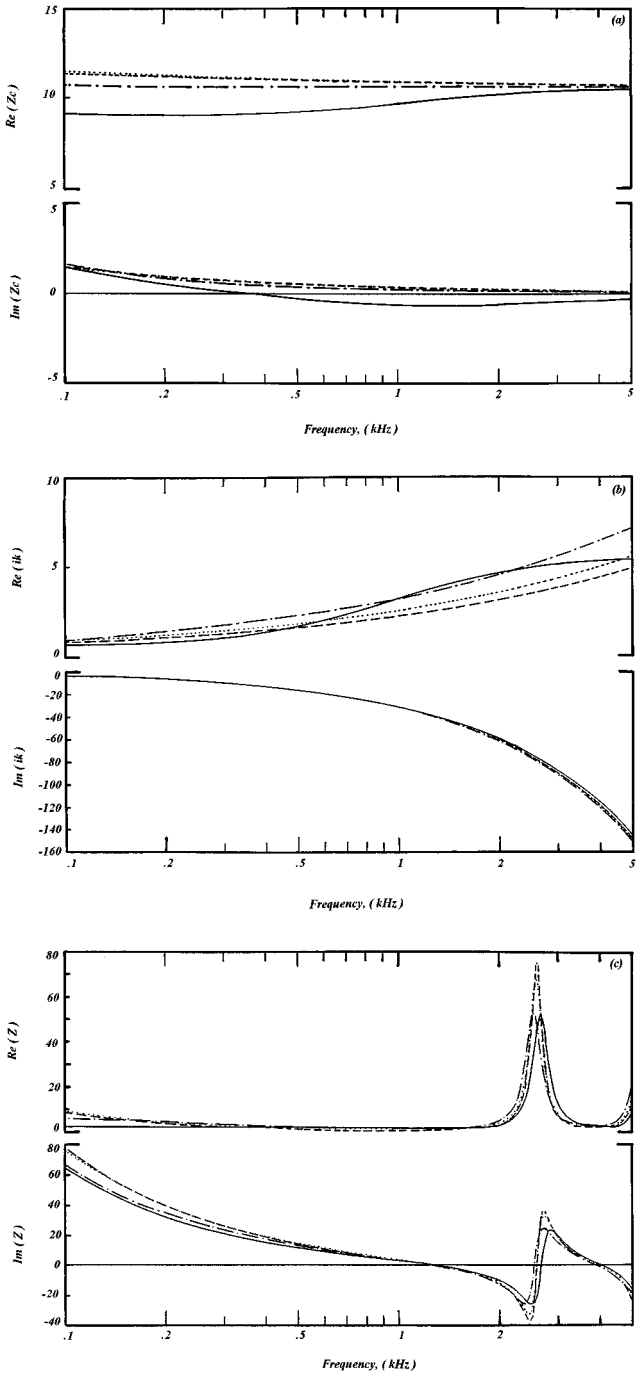


FIG. 3. Theoretical predictions obtained by the different approaches showing (a) characteristic impedance, (b) propagation constant, and (c) surface impedance. Values of $R_s=5000$ N s m^{-4}, $\Omega=15\%$, and $q^2=2.5$ are used by all curves. The different predictions are — phenomenological model; ---- microstructural model, with $s_p=s_k=1$; ···· microstructural model, with $s_p=1.14$ and $s_k=0.88$; and -·-·- microstructural model with $s_p=1.14$ and $s_k=0.44$.

by adjusting the viscous pore shape factor (optimum value of $s_\rho=1.14$). There is close agreement between the two models and with the measured data.

The absorption coefficient does not provide a stringent comparison of the theoretical models. The differences are more apparent in Fig. 3 which shows calculated characteristic impedance, propagation constant (i.e., ik) and surface impedance functions. The results for the phenomenological and

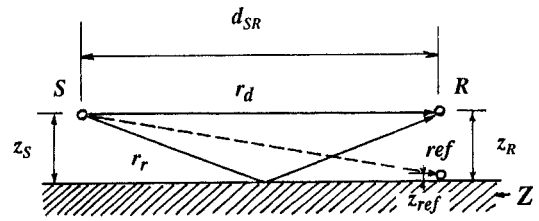


FIG. 4. Sketch showing the measurement setup above a porous pavement to obtain the level difference $L(z_R)-L(z_{\text{ref}})$. The horizontal separation between source and reference d_{Sref} was usually, but not always, the same as that between source and receiver d_{SR}.

microstructural models, shown by the solid and dashed curves, respectively, are based on the same parameter values used in Fig. 2.

In the case of simple pore geometries, with only gradual changes in cross-sectional area and shape along the pore axes, it is possible to obtain simple relationships between s_ρ and s_k. If the cross-sectional area varies but the shape of the pores is constant then it may be shown that $9s_\rho-1/s_k=8s_0$, where s_0 is the shape factor of each section composing the porous material. In the case of circular pores, $s_0=1$ so a value of $s_k=0.44$ would correspond to $s_\rho=1.14$. For this choice of shape factors, the microstructural model yields the dot-dashed curves in Fig. 3. In the case of a constant cross section with a modification of the shape along the pore axis, s_ρ and s_k are related by $s_k=1/s_\rho$. A viscous shape factor of 1.14 gives $s_k=0.88$ and the microstructural model yields the dotted curves in Fig. 3.

The results in Fig. 3 from the microstructural model for these three sets of shape factors are in close agreement with each other and with the phenomenological model. This suggests that the porous pavement sample can be described by shape factors bounded between the two simple hypotheses mentioned above. The relative independence of the acoustical properties of porous pavement is maintained for values of flow resistivity between 2000 and 30 000 N s m^{-4}. Since unity shape factors yield equally good agreement with measured values, the simpler phenomenological model, which uses only three parameters, is retained in the following section to predict measured sound-pressure levels above porous pavements outdoor.

III. ACOUSTIC PROPAGATION ABOVE A POROUS STRUCTURE

To verify that sound fields measured above road surfaces outdoors can be predicted using the theoretical models discussed in the previous section, sets of measurements were made on three different sites. The sound-pressure levels were obtained using the impulse method[21] described earlier. The experimental results are expressed in terms of level differences $L(z_R)-L(z_{\text{ref}})$ between two different receiver locations. The geometry is displayed in Fig. 4.

The solid curve in Fig. 5 shows level differences measured on a test track located at the Laboratoire Central des Ponts et Chaussées (LCPC). The results were obtained at a distance of 2 m from the source for the geometry described in the figure caption. For this site, the measured flow resis-

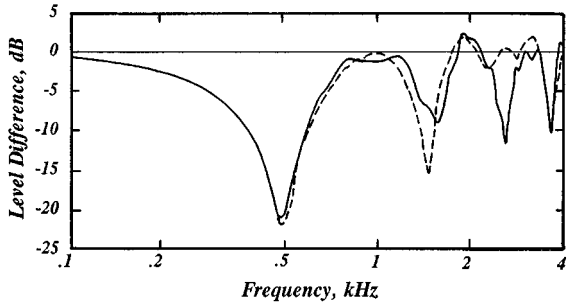


FIG. 5. Level differences versus frequency obtained above the porous asphalt LCPC test track. The geometrical parameters are $z_s=0.6$ m, $z_R=0.6$ m, $z_{\text{ref}}=0.006$ m, $d_{\text{SR}}=d_{\text{Sref}}=2$ m. The acoustic structural parameters are $R_s=15\,000$ N s m^{-4}, $\Omega=15\%$, $q^2=2$, $l=0.025$ m. Curves are:—the measurement results using the impulse technique[19] and ---- the predicted results obtained with Eq. (1).

tivity was $R_s=15\,000$ N s m^{-4} and the measured porosity was $\Omega=15\%$. From the measured absorption coefficient on the site, we obtained $q^2=2$. The thickness of the pavement is $l=0.025$ m. The level difference predicted from Eq. (12) is shown by the dashed curve in Fig. 5. We note that level difference measurements for short distances, as is the case here, are essentially insensitive to atmospheric inhomogeneities. Therefore, the same result would be obtained using Eq. (1) in the absence of atmospheric turbulence.

Figures 6 and 7 show results on a different site, RN6 Dardilly (near Lyon, France). The source–receiver distance is 4 m. Source and receivers heights were $z_S=0.6$ m, $z_R=0.6$ m, and $z_{\text{ref}}=0.006$ m for the measurements shown in Fig. 6 and $z_S=0.025$ m, $z_R=0.6$ m, and $z_{\text{ref}}=0.006$ m, in Fig. 7. The measured level differences are shown by the solid curves. The predicted level difference, using the measured parameters for this site of $R_s=2000$ N s m^{-4}, $\Omega=30\%$, $q^2=3.3$, and $l=0.04$ m, is shown by the dashed curve in Fig. 6. For comparison, the level differences predicted for propagation above a hard surface for this geometry is shown by the dotted curve. Relatively good agreement is obtained between measurements and prediction. There is some mismatch in the regions of maxima and minima for the data of Fig. 7 which used a lower source height. The discrepancy in the position of the extrema resembles the increase in sound-pressure level observed in interference minima due to atmospheric turbulence.[13] However, repeating the calculation using Eq. (12), which would account for turbulent effects, also gives the dashed curve. We can speculate that the discrepancies could result from phase errors inherent in the experimental impulse technique. The increase in the sound pressure level observed in Fig. 7 around 1500 Hz is due to the presence of a surface wave. Such results have been discussed by Stinson *et al.*[23]

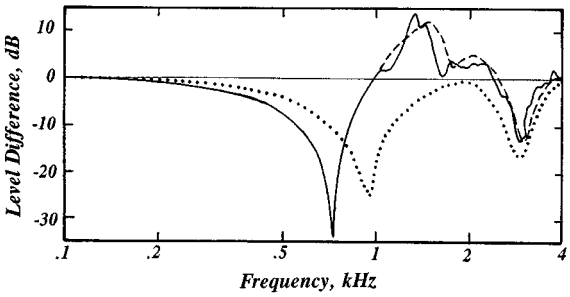


FIG. 6. Level differences versus frequency obtained above a porous asphalt at RN6 Dardilly, France. The geometrical parameters are $z_s=0.6$ m, $z_R=0.6$ m, $z_{\text{ref}}=0.006$ m, $d_{\text{SR}}=d_{\text{Sref}}=4$ m. The acoustic structural parameters are $R_s=2000$ N s m^{-4}, $\Omega=30\%$, $q^2=3.3$, and $l=0.04$ m. The curves are:—the measurement results using the impulse technique,[19] ---- the predicted results obtained with Eq. (1), and $\cdots$ the predicted results obtained with Eq. (1) above a perfectly hard surface.

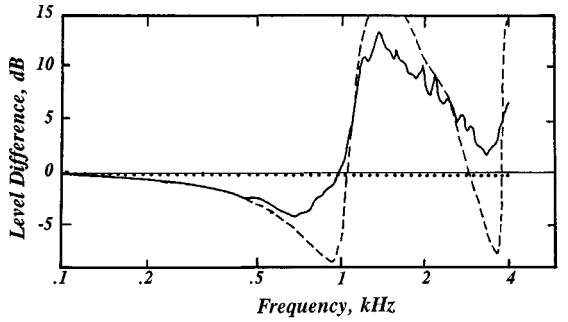


FIG. 7. Level differences versus frequency obtained above a porous asphalt at RN6 Dardilly, France. The geometrical parameters are $z_s=0.025$ m, $z_R=0.6$ m, $z_{\text{ref}}=0.006$ m, $d_{\text{SR}}=d_{\text{Sref}}=4$ m. The acoustic structural parameters are $R_s=2000$ N s m^{-4}, $\Omega=30\%$, $q^2=3.3$, and $l=0.04$ m. The curves are: — the measurement results using the impulse technique,[19] ---- the predicted results obtained with Eq. (1), and $\cdots$ the predicted results obtained with Eq. (1) above a perfectly hard surface.

The above figures show that predictions and measurements are in good agreement for a variety of geometries. We note that the difference between a porous and a hard surface is important in the range of 800 Hz to 1 kHz where traffic noise is the most energetic. For this reason, a measurable reduction of traffic noise is observed above porous pavement. The sensitivity to the surface properties can be used as the basis for an *in situ* technique to monitor the evolution of clogging of the porous pavement.

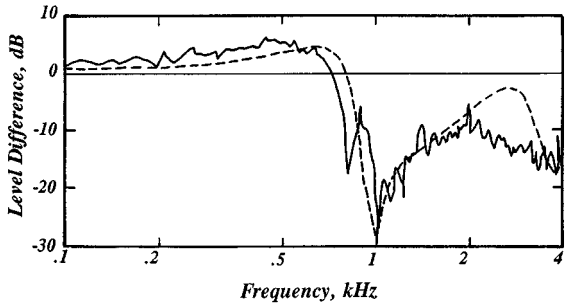


FIG. 8. Level differences versus frequency obtained above a porous asphalt at RN76, France. The geometrical parameters are $z_s=0.025$ m, $z_R=0.006$ m, $z_{\text{ref}}=0.003$ m, $d_{\text{SR}}=32$ m, and $d_{\text{Sref}}=2$ m. The acoustic structural parameters are $R_s=15\,000$ N s m^{-4}, $\Omega=15\%$, $q^2=3.3$, and $l=0.04$ m. The curves are: — the measurement results using the impulse technique,[19] and ---- the predicted results obtained with Eq. (12) using $\langle\mu^2\rangle=12\times10^{-6}$ and $L=1$ m.

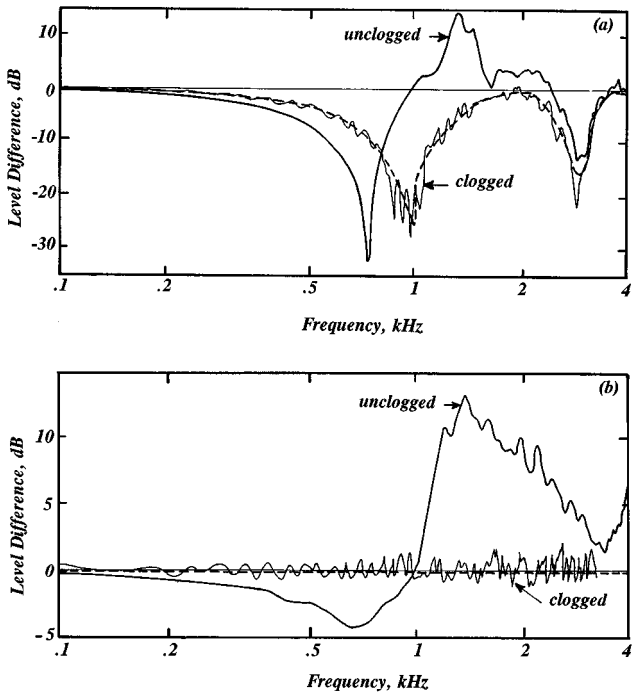


FIG. 9. Measured level differences versus frequency obtained above an unclogged (RN6 Dardilly, France) and a clogged (urban boulevard in Nantes, France) porous asphalt. The geometrical parameters are the same as Fig. 6 for (a) and the same as Fig. 7 for (b). The curves ---- are the predicted results obtained with Eq. (1) above a perfectly hard surface.

During their lifetime, porous structures do not keep their acoustical characteristics. The variation in the acoustical characteristics is due to the partial or total clogging of the structure by various wastes. The clogging principally causes the flow resistivity to increase and the porosity to decrease. This phenomenon can be observed by following the evolution of level difference measurements above porous pavement. In Fig. 9(a) and (b), the solid curves labeled ''unclogged'' reproduce the measured level difference spectra shown previously in Figs. 6 and 7 for the RN6 Dardilly site. Also shown are spectra measured for a totally clogged surface found on the urban Boulevard de Doulon in Nantes, France using the same measurement geometries. These curves are shown in Fig. 9(a) and (b) labeled as ''clogged.'' These measured spectra are similar to those that would be obtained for a hard surface, indicated by the dashed curves in the figure.

IV. CONCLUSION

The microstructural properties of porous pavement samples have been determined. The measured surface impedance and absorption coefficient were compared with two theoretical descriptions of porous materials, a phenomenological and a microstructural model. The microstructural model was shown to provide a good physical description of the acoustical properties of porous pavement. The description relies principally on the flow resistivity and the porosity of the material and, to a lesser extent, on the tortuosity. Further, the properties of the pavement are relatively insensitive to pore shape factor. Therefore, a simpler phenomenological model in terms of air flow resistivity, porosity, and tortuosity was found to provide an adequate description of the surface impedance of the pavement for the purpose of predicting sound-pressure levels. Measured level difference spectra were shown for different sites. Good agreement between measured and predicted spectra was found for a variety of geometries.

When the source and the reference microphone are both close to the surface, measured level difference spectra provide a convenient *in situ* method to characterize the evolution of clogging in the pores of the pavement. Visual inspection alone is inadequate because the problems of clogging often begin close to the bottom of the asphalt layer and would not be immediately visible.

[1] ''Enrobés drainants, le point en 1991,'' Note d'information No 70, Service d'Etudes Techniques des Routes et Autoroutes (SETRA), Centre de la Sécurité et des Techniques Routières (1991).
[2] S. Meiarashi, F. Nakashiba, H. Niimi, M. Hasebe, and T. Nakatsuji, ''Quantitative comparison between noise reduction factors of drainage asphalt pavement,'' Appl. Acoust. **44**, 165–179 (1995).
[3] M. Bérengier, J. F. Hamet, and P. Bar, ''Acoustical properties of porous asphalts: Theoretical and environmental aspects,'' 69th annual meeting of the Transportation Research Board, Washington, DC, Transportation Research Record No. 1265 (1990).
[4] M. Bérengier, ''Traffic noise propagation: Absorbing pavement and screening effect,'' Internoise 92, Toronto, Canada, 143–146 (1992).
[5] M. C. Bérengier, J. F. Hamet, G. A. Daigle, and M. R. Stinson, ''Acoustic impedance of porous pavements: Comparison between a microstructural and a phenomenological model,'' J. Acoust. Soc. Am. **93**, 2407(A) (1993).
[6] J. F. Hamet and M. Bérengier, ''Acoustical characteristics of porous pavements: A new phenomenological model,'' Internoise 93, Leuven, Belgium, 641–646 (1993).
[7] K. Attenborough and C. Howorth, ''Models for the acoustic characteristics of porous road surfaces,'' International Tire/Road Noise Conference, Gothenburg, Sweden, 177–191 (1990).
[8] Y. Champoux and M. R. Stinson, ''On acoustical models for sound propagation in rigid frame porous materials and the influence of shape factors,'' J. Acoust. Soc. Am. **92**, 1120–1131 (1992).
[9] T. F. W. Embleton, J. E. Piercy, and N. Olson, ''Outdoor sound propagation over ground of finite impedance,'' J. Acoust. Soc. Am. **59**, 267–277 (1976).
[10] G. A. Daigle, ''Effects of atmospheric turbulence on the interference of sound waves above a finite impedance boundary,'' J. Acoust. Soc. Am. **65**, 45–49 (1979).
[11] M. R. Stinson, ''A note on the use of an approximate formula to predict sound fields above an impedance plane due to a point source,'' J. Acoust. Soc. Am. **98**, 1810–1812 (1995).
[12] 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).
[13] M. A. Johnson, R. Raspet, and M. T. Bobak, ''A turbulence model for sound propagation from an elevated source above level ground,'' J. Acoust. Soc. Am. **81**, 638–646 (1987).
[14] J. F. Allard, C. Depollier, J. Nicolas, W. Lauriks, and A. Cops, ''Propriétés acoustiques des matériaux saturés d'air et théorie de Biot,'' J. Acoust. **3**, 29–38 (1990).
[15] J. F. Hamet and K. Attenborough, ''Acoustic absorption of porous roads: The phenomenological and microstructural models. Derivation of the acoustical characteristics,'' INRETS report No. MMA 9404 (1994).
[16] D. K. Wilson, ''Relaxation-matched modeling of propagation through porous media, including fractal pore structure,'' J. Acoust. Soc. Am. **94**, 1136–1145 (1993).
[17] ''A 322C vertical gamma densitometer bench,'' publication of Laboratoire Central des Ponts et Chaussées (LCPC), p. 29 (1991).
[18] Y. Champoux, M. R. Stinson, and G. A. Daigle, ''Air-based system for the measurement of porosity,'' J. Acoust. Soc. Am. **89**, 910–916 (1991).
[19] ISO 9053, ''Acoustics: Materials for acoustical applications—determination of airflow resistance'' (1988).
[20] T. F. W. Embleton, J. E. Piercy, and G. A. Daigle, ''Effective flow resistivity of ground surfaces determined by acoustical measurements,'' J. Acoust. Soc. Am. **74**, 1239–1244 (1983).

[21] M. Bérengier and Y. Delanne, ''Caractéristiques acoustiques intrisèques d'une structure par une méthode impulsionnelle,'' Bull. Liaison Lab. de Ponts et Chaussées **139**, 113–118 (1985), and Norme Française S 31-089, ''Acoustique: Code d'essai pour la détermination de caractéristiques acoustiques d'écrans utilisés en champ libre'' (1991).
[22] Y. Champoux and M. R. Stinson, ''Measurement of the tortuosity of porous materials and applications for acoustical modeling,'' J. Acoust. Soc. Am. Suppl. 1 **87**, S139 (1990).
[23] M. R. Stinson, G. A. Daigle, and D. I. Havelock, ''The measurement of air-coupled surface waves over a porous surface,'' Internoise 93, Leuven, Belgium, 1683–1686 (1993).

Fresnel zones for modes

Anatoly L. Virovlyansky
Institute of Applied Physics, Russian Academy of Science, 46 Ul'yanov Street, 603600 Nizhny Novgorod, Russia

Vasily V. Kurin, Nikolay V. Pronchatov-Rubtsov, and Sergey I. Simdyankin
University of Nizhny Novgorod, 23 Gagarin Avenue, 603600 Nizhny Novgorod, Russia

(Received 1 March 1996; accepted for publication 30 July 1996)

The notion of Fresnel zones for modes is introduced which is analogous to the usual Fresnel zones introduced for rays. It is shown that by using Fresnel zones for modes one can simplify the analysis of mode scattering at large-scale inhomogeneities in multimode waveguides. The mode amplitude variations due to scattering by an opaque screen and by refractive index inhomogeneities are considered. In the latter case the formulas obtained in the scope of this approach are statistical averaging suited and similar to those of geometrical optics and the Rytov method widely used to calculate fluctuations of ray complex amplitudes. The results of a tank experiment carried out to test the predictions concerning the mode scattering by an opaque screen are presented. © *1997 Acoustical Society of America.* [S0001-4966(97)02012-2]

PACS numbers: 43.30.Bp, 43.30.Dr, 43.30.Ft, 43.20.Fn [MBP]

INTRODUCTION

High-frequency acoustic propagation in waveguides is usually analyzed in terms of ray or normal mode representations. It is well known that there are aspects of mode theory which have parallels in ray theory, and vice versa.[1–10] The utilization of such parallels grants additional physical insight and therefore clarifies the mechanism of wave propagation. Besides, a hybrid ray-mode description can be introduced[5] which draws upon the advantages of each.

Despite the fact that most results on studying parallels between rays and modes have been obtained for range-independent waveguides or waveguides slowly varying with range, ray-mode analogy goes beyond these simple models. In particular, mode amplitude variations due to scattering by medium inhomogeneities with scales considerably exceeding the wavelength can be described using the relations similar to the well-known formulas widely used for studying of ray amplitude variations.[11–13] Such relations were obtained for mode scattering by an opaque screen,[11,13] by refractive index inhomogeneities,[12,14] and by a rough waveguide boundary.[15,16] It has been shown that to simplify analysis of diffraction effects in the above problems, Fresnel zones for modes can be introduced, which are analog of usual Fresnel zones for rays.[17] Fresnel zones for modes, like their ray prototypes, are introduced in the scope of the Huyghens–Fresnel–Kirchhoff approach: A cylindrical surface surrounding a point source is divided into zones (Fresnel zones for modes) in such a way that each zone contains the virtual (secondary) sources, contributions from which to the field of the given mode add in phase. The influence of a large-scaled inhomogeneity (e.g., an opaque screen) on the mode amplitude depends on what zones and to what extent it blocks. The mathematical formalism here is analogous to the standard one used to analyze ray amplitude variations.

This paper is devoted to a discussion of the concept of the Fresnel zones for modes. In Sec. I we present the necessary propagation equation and the ray and mode representations of the field in a range-independent waveguide. In Sec. II we present some results on studying the relationship between modes and rays in a range-independent waveguide which is used in the following parts of the article. Here we mainly follow Ref. 3. The introduction of the notion of the Fresnel zones for modes and their use when analyzing mode diffraction by an opaque screen is considered in Sec. III. In Sec. IV we deal with applying the notion to problems of wave scattering by large-scaled inhomogeneities of refractive index. The results of a tank experiment carried out to test the predictions made in the framework of the approach under consideration as applied to the problem of sound scattering by an opaque screen are described in Sec. V.

Despite the fact that the subject is treated here mainly in terms of underwater acoustics, the obtained results are of a more general nature and they can be used to analyze diffraction effects in multimode waveguides of different physical origin.

I. THE RAY AND MODE REPRESENTATIONS IN A RANGE-INDEPENDENT WAVEGUIDE

Let us consider a range-independent acoustic waveguide with the sound speed c being the function of the downward increasing vertical coordinate z only. The waveguide is a horizontal water layer $0 \leqslant z \leqslant z_b$ with a pressure release upper boundary $z=0$, lying over a half-space $z > z_b$. The refractive index profile is defined by the relation $n(z)=c_{\min}/c(z)$, where $c_{\min}$ is the minimum sound speed inside the layer. We seek time-harmonic solutions of the wave equation

$$\frac{\partial^2 p}{\partial x^2}+\frac{\partial^2 p}{\partial y^2}+\frac{\partial^2 p}{\partial z^2}+k^2 n^2(z)p=-4\pi\delta(x)\,\delta(y)\,\delta(z-z_0) \tag{1}$$

subject to excitation by a point source located at $(0,0,z_0)$. Here, k is the wave number in a reference medium with $n=1$. A time dependence $e^{-i\omega t}$ is suppressed.

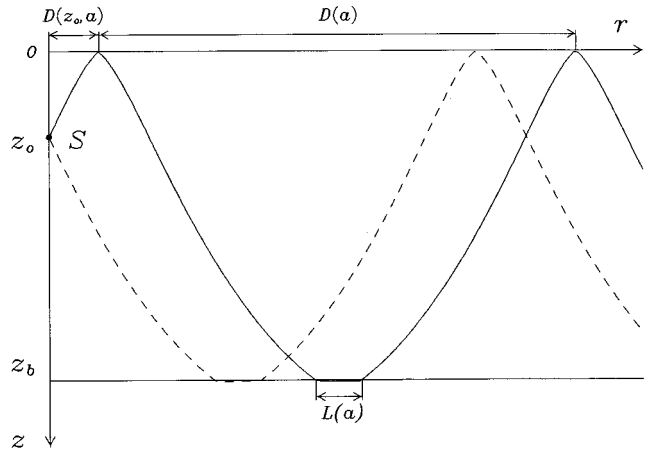


FIG. 1. Ray trajectories with the turning points intercepting the upper and lower boundaries.

We restrict ourselves to analysis of the constituent of the total wave field that is formed by modes with horizontal propagation numbers belonging to the discrete part of the spectrum. In terms of the ray approach, it means that we consider only the rays with the grazing angles at the water–bottom interface $z=z_b$ not exceeding the angle of the total internal reflection.

A. Ray representation

Each ray trajectory is governed by Snell's law $n(z)\cos(\theta)=a$, where θ is the current grazing angle, $a=\cos\chi$, is the ray parameter, and χ is the grazing angle at the axis of the sound channel, i.e., at the horizon where $c(z)=c_{\min}$. Let $z_{\min}$ and $z_{\max}$ be the coordinates of the upper and lower ray turning points, respectively. The part of the ray trajectory connecting the upper turning point and the point at the depth z has the length (in horizontal direction)[2]

$$D(z,a)=a\int_{z_{\min}}^{z}\frac{dz}{\sqrt{n^2(z)-a^2}}.$$

The ray cycle distance[2]

$$D(a)=2a\int_{z_{\min}}^{z_{\max}}\frac{dz}{\sqrt{n^2(z)-a^2}}+L(a)$$

includes a lateral shift on the lower boundary[2,18] $L=-(1/k)(\partial\phi_b/\partial a)$, where ϕ_b is the phase of the reflection coefficient at the interface $z=z_b$. If the lower turning point is inside the water bulk, then $L=0$. The above designations are illustrated in Fig. 1, where two ray trajectories that intercept both boundaries are shown. The equation of the ray path escaping the point source and having the ray parameter a can be written as

$$r=R_{N,j}(z,a)=\alpha_j D(z_0,a)+ND(a)+\beta_j D(z,a), \qquad (2)$$

where $r=\sqrt{x^2+y^2}$ (x and y are the horizontal coordinates), N is the number of lower turning points, and α_j and β_j are constants equal to $+1$ or -1, depending on signs of the launch and arrival angles. The rays are separated into four groups distinguished by the index $j=1,2,3,4$, according to whether the trajectory leaves the source in the upward or downward direction and arrives at the point (r,z) from above or below.

In the scope of the ray representation the solution of the Helmholtz equation (1) is equal to

$$p(r,z)=\sum_{j=1}^{4}\sum_{N}A_{N,j}e^{i\Psi_{N,j}}\Big|_{a=a_{N,j}}, \qquad (3)$$

with the ray amplitudes

$$A_{N,j}=\frac{a^{1/2}e^{i\pi/4}}{\sqrt{r|\partial R_{N,j}/\partial a|}}\frac{1}{\sqrt[4]{(n^2(z_0)-a^2)(n^2(z)-a^2)}},$$

and phases

$$\Psi_{N,j}=k[\alpha_j J(z_0,a)+NI(a)+\beta_j J(z,a)+ar]+N(\phi_s+\phi_b)+\frac{\alpha_j+\beta_j}{2}\phi_s-\frac{\pi}{4}\operatorname{sgn}\left(\frac{\partial R_{N,j}}{\partial a}\right), \qquad (4)$$

where

$$J(z,a)=\int_{z_{\min}}^{z}\sqrt{n^2(z)-a^2}\,dz,$$

$$I(a)=2\int_{z_{\min}}^{z_{\max}}\sqrt{n^2(z)-a^2}\,dz,$$

the function sgn shows a sign of its argument, and ϕ_s and ϕ_b are the reflection phase shifts at the upper and lower turning points, respectively.

The parameters $a_{N,j}$, corresponding to rays connecting the source and the observation point (r,z), satisfy the equation

$$\frac{\partial\Psi_{N,j}}{\partial a}=k(r-R_{N,j}(z,a))=0, \qquad (5)$$

where the derivative with respect to a has been calculated taking into account the relations

$$\frac{\partial J(z,a)}{\partial a}=-D(z,a),$$

$$\frac{\partial I(z,a)}{\partial a}=-2a\int_{z_{\min}}^{z_{\max}}\frac{dz}{\sqrt{n^2(z)-a^2}}.$$

So, the ray parameter $a_{N,j}$ is the function of both the parameters (N,j) determining the ''topology'' of the ray trajectory and the coordinates of the observation point (r,z).

B. Mode representation

The normal mode solution to the given problem in the far field is given by[2,18]

$$p(r,z)=\sum_{m}\sqrt{\frac{2\pi}{kra_m}}\,\varphi_m(z_0)\varphi_m(z)e^{ika_m r+i\pi/4}, \qquad (6)$$

where $a_m=k_m/k$, k_m is the horizontal propagation number, and the sum runs over the set of normal modes, $\varphi_m(z)$.

In the WKB approximation the values a_m satisfy the equation

$$kI(a_m) = 2\pi m - \phi_s - \phi_b . \tag{7}$$

The function $\varphi_m(z)$ between the turning points can be expressed as[2]

$$\varphi_m(z) = \varphi_m^+ + \varphi_m^- , \tag{8}$$

where

$$\varphi_m^{\pm}(z) = \frac{1}{\sqrt[4]{n^2(z) - a_m^2}} \sqrt{\frac{a_m}{D_m}} \exp[\pm i(kJ(z,a_m) + \phi_s(a_m)/2)],$$

with $D_m \equiv D(a_m)$. The functions φ_m^+ and φ_m^- present upgoing and downgoing waves (the so-called Brillouin waves) forming the mth mode.

Within the context of the WKB approximation the mode representation of the wave field (6) takes the form

$$p(r,z) = \sum_{j=1}^{4} \sum_m B_{m,j} e^{i\Phi_{m,j}} \Big|_{a=a_m} , \tag{9}$$

where

$$B_{m,j} = \sqrt{\frac{2\pi a}{kr}} \frac{1}{D(a)} \frac{1}{\sqrt[4]{(n^2(z_0) - a^2)(n^2(z) - a^2)}},$$

$$\Phi_{m,j} = k[\alpha_j J(z_0,a) + \beta_j J(z,a) + ar] + \frac{\alpha_j + \beta_j}{2} \phi_s + \frac{\pi}{4}.$$

Here α_j and β_j are the same constants as those in Eq. (4).

II. SOME RELATIONS BETWEEN RAYS AND MODES IN A RANGE-INDEPENDENT WAVEGUIDE

The mode representation (9) can be converted into the ray representation (3) and vice versa by the Poisson summation formula.[2,5] Moreover, the Poisson summation formula converts each part of one representation formed by terms with the given index j into the part of other representation formed by terms with the same index j. In both representations the terms with the given index j describe waves with the same pair of the launch and arrival angle signs.

However, there exists a much more detailed relationship between rays and modes.[3,5,6] Let us consider the wave field in an arbitrary vertical plane containing the source. The mth mode constructively interferes (adds in phase) with neighboring modes, i.e., the modes with numbers close to m, along the trajectories of two geometric rays leaving the source and having the same ray parameters $a = a_m$. The launch angles of these rays are equal in absolute value, opposite in sign, and coincide with the grazing angles (at the horizon $z = z_0$) of the Brillouin waves forming the mth mode. These rays, playing a great role in what follows, we shall call the mode rays and denote their trajectories as $z = z_m^{\pm}(r)$, where the superscript shows the sign of the launch angle. The same trajectories are determined by Eq. (2) with $a = a_m$ and appropriate parameters N and j. The parameters N and j are varying along the trajectories.

Let us compare the phases of two terms in Eq. (9) having the same index j and corresponding to the mth and the $(m+\mu)$th modes. Following Ref. 3 and using the formulas from the previous section, it can be shown that along the trajectory of the mode ray the approximate relation

$$\Phi_{m+\mu,j} - \Phi_{m,j} = -2\pi N\mu + \left(\frac{2\pi}{D_m}\right)^2 \frac{1}{2k} \frac{\partial R_{N,j}(a_m)}{\partial a} \mu^2,$$

takes place for $|\mu| \ll m$. While the last term in the above equation remains much less than π, the phase difference is approximately equal to some multiple of 2π and the modes interfere constructively. The value of $|\mu|$ at which the last term becomes equal to π can be chosen as an estimation of the number of modes constructively interfering with the mth mode along one of its mode rays. Denoting these estimations as $M_m^{\pm}$ (they are functions of range and the superscripts indicate the corresponding mode rays) we get

$$M_m^{\pm} = \sqrt{\frac{kD_m^2}{2\pi |\partial R_{N,j}/\partial a|_{a=a_m}}} = \sqrt{D_m \left| \frac{\tan \theta_m^{\pm}}{\partial z_m^{\pm}/\partial m} \right|}, \tag{10}$$

where the derivatives $\partial R_{N,j}/\partial a$ are calculated for the corresponding mode rays, $\partial z_m^{\pm}/\partial m = z_{m+1}^{\pm}(r) - z_m^{\pm}(r)$, and $\theta_m^{\pm}$ are the grazing angles of the mode rays. Equation (10) cannot be used near caustics where $M_m^{\pm}$ goes to infinity. The latter qualitatively reflects the fact that in terms of normal modes, caustics are the places where many modes add in phase. Note that far from caustics the derivatives $|\partial R_{N,j}/\partial a|$ are almost linearly increasing with distance. According to Eq. (10) it means that $M_m^{\pm}$ are decreasing with distance, on the average as $r^{-1/2}$, and at large enough ranges each mode ceases to add in phase with neighboring modes.

As it has been shown in Ref. 3, the group of constructively interfering modes form the contributions to the total wave field from the individual rays. For $M_m^{\pm} \gg 1$, the values $M_m^{\pm}$ can be considered as the numbers of modes forming the rays with the ray parameters equal to $a = a_m$ at the range r. More rigorous and detailed analysis of these issues has been done in Refs. 5–7. In what follows, we restrict ourselves to analysis of the wave field at short enough distances where the inequality $M_m^{\pm} \gg 1$ remains valid.

A cruder, but simpler, estimation of $M_m^{\pm}$ can be obtained by replacing the tangents of grazing angles and the derivatives with respect to mode number in Eq. (10) by their values averaged over the cycle of oscillation of the mode rays. Averaging on the cycle we shall denote by the horizontal line over the corresponding quantity. The averaged values of M_m^+ and M_m^- are approximately equal and the superscripts $\pm$ in the following two formulas are omitted. Taking into account the relation

$$\overline{|\tan \theta_m|} = \frac{2}{D_m} \int_{z_{\min}}^{z_{\max}} \left| \frac{\partial z_m}{\partial r} \right| dr = \frac{2(z_{\max} - z_{\min})}{D_m},$$

obtained under assumption that $D_m \gg L(a_m)$, we get

$$M_m = \sqrt{\frac{2(z_{\max} - z_{\min})}{\overline{|\partial z_m/\partial m|}}}. \tag{11}$$

Note that we employ the same designations for the coordinates of the mode and ray turning points. It is seen that the mth mode constructively interferes with the neighboring modes while the vertical distance between the two neighboring mode rays with equal signs of the launch angles remains

much less than the vertical distance between the mode turning points.

III. THE NOTION OF THE FRESNEL ZONES FOR MODES

A. Fresnel zones on cylindrically symmetric surface

Let us consider a cylindrically symmetric surface $r=r_s$ surrounding the point source. The total field complex amplitude $p(r,z)$ at $r \geqslant r_s$ are connected with the amplitude at the cylindrical surface by the approximate relations

$$p(r,z)=\sum_m U_m \varphi_m(z) e^{ika_m(r-r_s)};$$
$$U_m=\int_0^{z_b} p(r_s,z)\varphi_m(z)dz. \tag{12}$$

In the expression for U_m we have neglected the ''tails'' of the normal modes exponentially decaying in the half-space $z>z_b$.

Dividing $\varphi_m(z)$ into two terms according to Eq. (8) and using the ray representation for $p(r_s,z)$ [see Eq. (3)], we get the following expression for the mode amplitudes U_m:

$$U_m=\sum_{N,j} (G^+_{N,j}+G^-_{N,j}), \tag{13}$$

where

$$G^{\pm}_{N,j}=\int dz\ A_{N,j} e^{iF^{\pm}_{N,j}},$$
$$F^{\pm}_{N,j}=\Psi_{N,j}\pm[kJ(z,a_m)+\phi_s/2]. \tag{14}$$

The amplitudes in the integrands in Eqs. (14) are slowly varying functions of z (except the vicinities of caustics, which we do not consider), while their phases are rapidly varying functions of z. The main contributions to these integrals are given from the neighborhood of stationary phase points, where the derivative of $F^{\pm}_{N,j}$ with respect to z vanishes. Since the second term on the right-hand side of

$$\frac{d\Psi_{N,j}}{dz}=\frac{\partial\Psi_{N,j}}{\partial z}+\left.\frac{\partial\Psi_{N,j}}{\partial a}\right|_{a=a_{N,j}}\frac{\partial a_{N,j}}{\partial z}$$

according to (5) is identically zero, the expression for the phase derivative can be written as

$$\frac{dF^{\pm}_{N,j}}{dz}=-k\beta_j\sqrt{n^2(z)-a^2_{N,j}}\pm k\sqrt{n^2(z)-a^2_m}.$$

It vanishes at $z=z^{\pm}_m(r_s)$, i.e., at the points where the trajectories of the mode rays cross the surface $r=r_s$, since only at these points the condition $a_{N,j}=a_m$ is fulfilled. It means that only two out of all the integrals in sum (13) have the stationary points between the integration limits. These two integrals give the main contribution to U_m. The phases of their integrand we denote as $F^{\pm}_m$, where the superscripts coincide with the superscripts of the corresponding mode rays.

The second derivatives of the phases $F^{\pm}_m$ with respect to z are equal to

$$\left.\frac{d^2F^{\pm}_m}{dz^2}\right|_{z=z^{\pm}_m(r_s)}=k\beta_j\left.\frac{\partial a^2_{N,j}/\partial z}{2\sqrt{n^2(z)-a^2_m}}\right|_{z=z^{\pm}_m(r_s)}$$
$$=\frac{k}{\partial R_{N,j}/\partial a|_{a=a_m}\tan^2\theta^{\pm}_m}.$$

Approximating the phases in the neighborhood of the stationary points by parabolas, we get

$$F^{\pm}_m=\alpha_j(kJ(z_0,a_m)+\phi_s/2)+\frac{k}{2(\partial R_{N,j}/\partial a)|_{a=a_m}\tan^2\theta^{\pm}_m}(z-z^{\pm}_m(r_s))^2+\gamma^{\pm}_m\frac{\pi}{4},$$

where $\gamma^{\pm}_m$ show the signs of the derivatives $\partial R_{N,j}/\partial a$ calculated for the corresponding mode ray at the range r_s. Substitution of these expressions into Eqs. (12) and (14), replacing the preexponential factors in Eqs. (14) by their values at the stationary points $z^{\pm}_m(r_s)$, gives the simplified expression for the wave field at $r \geqslant r_s$.

According to the Huyghens–Fresnel principle[17] the field at $r \geqslant r_s$ is a result of interference of waves generated by secondary (virtual) sources on the surface $r=r_s$. The contribution from the sources located in the range $z_1 \leqslant z \leqslant z_2$ (the difference z_2-z_1 is assumed to be considerably greater than the wavelength) are given by Eqs. (12)–(14) with the integrals in Eqs. (14) running from z_1 to z_2. So, the constituent of the total field formed by the secondary sources located in the above range can be approximately represented as

$$p(r,z)=\sum_m \sqrt{\frac{2\pi}{ka_m r}}[\xi^+_m\varphi^+_m(z_0)+\xi^-_m\varphi^-_m(z_0)]\varphi_m(r)e^{ika_m(r-r_s)+i\pi/4}, \tag{15}$$

where

$$\xi^{\pm}_m=\frac{1}{\rho^{\pm}_m}\int_{z_1}^{z_2}dz\ \exp\left[i\gamma^{\pm}_m\left(\frac{\pi(z-z^{\pm}_m(r_s))^2}{(\rho^{\pm}_m)^2}-\frac{\pi}{4}\right)\right], \tag{16}$$

$$(\rho^{\pm}_m)^2=\lambda\left|\left.\frac{\partial R_{N,j}}{\partial a}\right|_{a=a_m}\tan^2\theta^{\pm}_m\right|=D_m\left|\frac{\partial z^{\pm}_m}{\partial m}\tan\theta^{\pm}_m\right|. \tag{17}$$

This result shows that comparatively small intervals surrounding the depths $z^{\pm}_m(r_s)$ on the surface $r=r_s$ can be indicated, where the secondary sources are located which play the main role in forming the field of an mth mode. In the neighborhood of each of these two depths, it is worthwhile to introduce a system of zones inside of which the phases of the exponents in Eq. (16) differ by not more than π. The boundaries of the νth zone ($\nu=1,2,...$) are defined by the condition

$$\nu-1 \leqslant \left(\frac{z-z^{\pm}_m(r_s)}{\rho^{\pm}_m}\right)^2 \leqslant \nu.$$

In analogy with the ray theory let us call these zones the Fresnel zones of the mth mode. The quantities $\rho^{\pm}_m$ defined by Eq. (17) represent the radii of the first Fresnel zones. The

zones are not circles, but it is convenient to use the term ''radii'' to describe their distances.

The notion of the Fresnel zones for modes cannot be used at caustics (where the ray representation fails) and at the mode ray turning points (where the WKB approximation to the eigenfunction fails). At these places the radii of the first Fresnel zones formally vanish.

Clearly, all the reasoning offered above can only be applied to the case when the radii of the first Fresnel zones are much less compared with the vertical distance between the mode turning points. (By the way, the same can be said about the usual Fresnel zones for rays.) Under that condition the integration in Eq. (16) over the whole range $0 \leqslant z \leqslant z_b$, i.e., summation over all the secondary sources, is practically the same as integration in infinite terms. Then, as it must be, Eq. (15) reduces to Eq. (6). The condition under consideration implies the constructive interference of the mth mode with a large number of its ''neighbors:'' $M_m \gg 1$. It can be easily shown by taking into account (10) and (11) and transforming Eq. (17) to

$$\rho_m \simeq \frac{D_m \overline{|\tan\ \theta_m|}}{M_m} \simeq \frac{2(z_{\max} - z_{\min})}{M_m}. \tag{18}$$

The above derivation of Eqs. (15)–(17) assumes that the secondary sources are located far from waveguide boundaries. Nevertheless, a simple additional analysis shows that the same relations remain valid for secondary sources located at arbitrary small depth. In particular, the lower limit in Eq. (16) can be equal to zero. This can be proved proceeding from suggestions similar to those that are often used to introduce the concept of image sources for accounting for influence of a pressure release boundary on sound propagation.[18]

So, the basic role in forming the field of the given mode is that of the secondary sources located in the neighborhood of its mode rays. This fact allows simple physical interpretation. According to the results discussed in Sec. II each Brillouin wave adds in phase (constructively interferes) with the Brillouin waves of neighboring modes only along the trajectory of its mode ray. Because of this, in the mode ray neighborhood, as distinct from the other waveguide portions, the field structure and, hence, the secondary source distribution repeat the structure of the corresponding Brillouin wave. In other words, secondary sources located in these places match with the mode field efficiently exciting the mode. Inasmuch as each mode is formed by two Brillouin waves, it conforms to two mode rays and, hence, to two Fresnel zone systems.

In the Kirchhoff approximation,[17] Eq. (16) formally describes the wave field behind an opaque screen with the slit between the horizons z_1 and z_2. In the same way diffraction by the screen with a number of slits can be described: integration in Eq. (16) should be carried out over all the slits. The notion of the Fresnel zones for modes can be used to analyze diffraction by screen in the same way as the usual Fresnel zones of rays are used for the similar purpose. On the other hand, the situation with modes is a little bit more complicated because there are two systems of Fresnel zones for each mode and these systems may overlap. Nevertheless, using the Fresnel zones for modes it is easy to predict variations of the coefficients $\xi_m^{\pm}$ due to blocking the part of the wave guide cross section. The mode amplitudes are defined by a linear combination of these coefficients [see Eq. (15)].

The analysis is especially simple when the vertical distance between the points $z_m^+(r_s)$ and $z_m^-(r_s)$ considerably exceeds the radii of the first Fresnel zones, i.e., the two Fresnel zones systems do not overlap. For example, the screen blocking the whole wave guide cross section, except two slits coinciding with the first Fresnel zones of the given mode, ''passes'' the mode only slightly distorting its amplitude. The screen blocking only the odd (or only the even) Fresnel zones of the given mode ''amplifies'' the mode amplitude. The similar effects for rays are well known.[17] It is not difficult to give an example of a pure ''mode'' effect which does not have a ''ray'' analog. Let us consider a case when the point source depth, z_0, coincides with the zero of an eigenfunction $\varphi_m(z)$. Then the mth mode is not excited by the source: despite each of the coefficients $\xi_m^{\pm}$ in the absence of the screen being equal to 1, the combination $\xi_m^+ \varphi_m^+(z_0) + \xi_m^- \varphi_m^-(z_0)$ defining the mode amplitude vanishes. Blocking one of the mode rays by a small opaque screen and, thus, changing one of the coefficients $\xi_m^{\pm}$, one can obtain the efficient excitation of this mode behind the screen. All the effects enumerated above are different manifestations of energy redistribution between modes as a result of scattering by screen.

B. Fresnel zones on plane

Let us consider the wave field on the plane $x = \text{const} > 0$ and expand its dependence on the transverse coordinate y into plane waves e^{iqy}. At large enough ranges from the source the field in the area $|y| \ll x$ is mainly formed by the waves propagating at small, with respect to the x-axis, angles, and Eq. (6) reduces to

$$p(x,y,z) = \sum_m \frac{i}{ka_m} \int_{-\infty}^{\infty} dq\ \varphi_m(z_0)\varphi_m(z) \times \exp\left(iqy - i\,\frac{q^2 x}{2ka_m} + ika_m x \right).$$

This expression, like Eq. (6), can be treated as the mode representation of the wave field, considering the product $\varphi_m(z)e^{iqy}$ as the eigenfunction defined by the parameters (q,m). The mode with the parameters (q,m) constructively interferes with modes having close parameters along the trajectories of geometric rays $z = z_{m,q}^{\pm}(x)$ lying in the vertical plane $y = y_{m,q}(x)$, where $y_{m,q}(x) = (q/ka_m)x$. The functions $z_{m,q}^{\pm}(x)$ coincide with the earlier defined functions $z_m^{\pm}(r)$ with r replaced by x (the subscript q identifies the vertical plane that the ray trajectories lie in).

In the plane $x = x_s$ we consider the secondary sources located in the area $z_1 \leqslant z \leqslant z_2$; $y_1 \leqslant y \leqslant y_2$. Using arguments similar to those used in the previous subsection, it can be shown that the contribution from the chosen secondary sources to the total wave field at $x \geqslant x_s$ is equal to

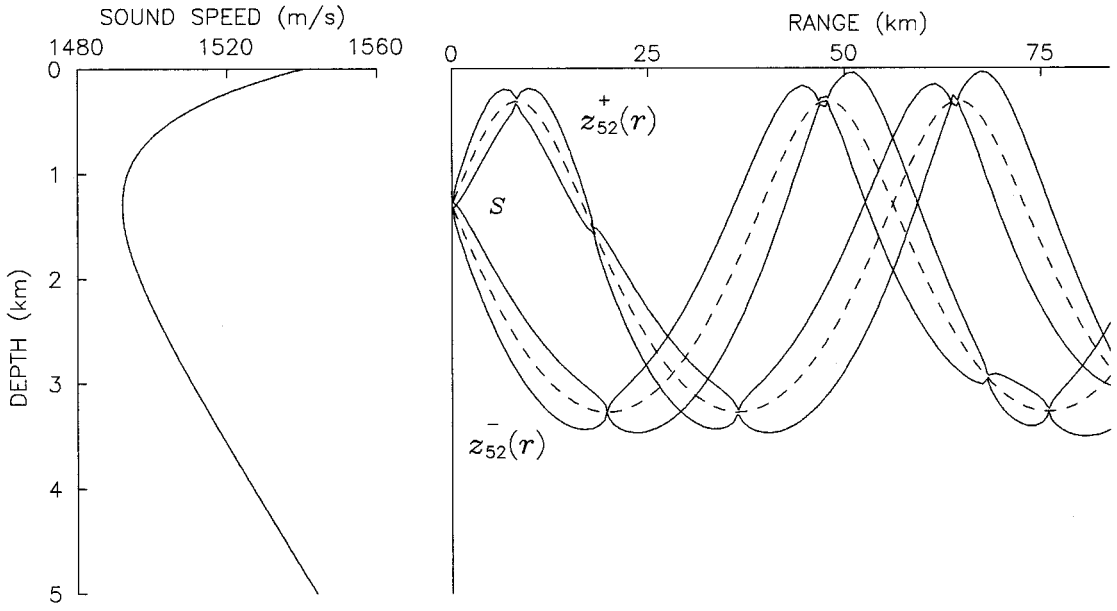


FIG. 2. Canonical sound-speed profile. The field is excited by a 100-Hz source. The areas of the waveguide covered by the first Fresnel zones strung on the mode rays corresponding to the 52nd mode (dashed curves) outlined by the solid curves.

$$p(x,y,z)=\sum_m \frac{i}{ka_m}\int dq[\xi^{+}_{m,q}\varphi^{+}_m(z_0) + \xi^{-}_{m,q}\varphi^{-}_m(z_0)]\varphi_m(z)\exp\left(iqy-i\frac{q^2x}{2ka_m}+ika_mx\right), \tag{19}$$

where

$$\xi^{\pm}_{m,q}=\int_{z_1}^{z_2}\int_{y_1}^{y_2} dz\ dy\ \Delta^{\pm}_{m,q}, \tag{20}$$

$$\Delta^{\pm}_{m,q}=\frac{1}{\rho_{m,y}\rho^{\pm}_{m,z}}\exp\left\{i\pi\left[\frac{(y-y_{m,q}(x_s))^2}{\rho^2_{m,y}}+\gamma^{\pm}_{m,q}\frac{(z-z^{\pm}_{m,q}(x_s))^2}{(\rho^{\pm}_{m,z})^2}\right]-i\frac{\pi}{4}(1+\gamma^{\pm}_{m,q})\right\},$$

$$\rho_{m,y}=\sqrt{\frac{2\pi x_s}{ka_m}}, \tag{21}$$

$\rho_{m,z}$ is defined by Eq. (17) with $r=x_s$, and $\gamma_{m,q}$ is a complete analog of γ_m in Eq. (16). The quantities $\rho_{m,y}$ and $\rho_{m,z}$ in analogy with the previous subsection we shall call the radii of the first Fresnel zones of the mth mode in horizontal and vertical planes, respectively.

In the scope of the Kirchhoff approximation Eqs. (19)–(21) describe the wave field behind the rectangular opening in an opaque screen. The same relations can be used to describe diffraction on the opening of an arbitrary shape with the diameter considerably exceeding the wavelength. Integration in Eq. (20) should be carried out over the opening.

IV. FRESNEL ZONES AND DIFFRACTION BY INHOMOGENEITIES OF REFRACTIVE INDEX

The relations similar to those presented above can be obtained for description of mode amplitude variations due to the influence of large-scaled inhomogeneities of the refractive index. We shall define the latter by adding in Eq. (1) a small perturbation $\delta n(x,y,z)$ depending on all three coordinates, to the unperturbed refractive index $n(z)$. In the case when all characteristic scales of the function δn are much greater than the wavelength, the wave field description is often performed in the scope of the ray theory.[2,19] Diffraction effects can be taken into account in the scope of the Rytov method generalizing the geometrical optics approximation.[19] In Refs. 12 and 13 it has been shown that there exist analog of these approaches to describe mode amplitude variations. In particular, an approximate solution to the Helmholtz equation can be expressed by Eq. (19) with $\xi^{\pm}_{m,q}$ equal to

$$\xi^{\pm}_{m,q}=\exp(iX^{\pm}_{m,q}), \tag{22}$$

where

$$X^{\pm}_{m,q}(x)=\frac{k}{a_m}\int_0^x dx'\int\int dz\ dy \times \delta n(x',y,z)\Delta^{\pm}_{m,q}(x',y,z). \tag{23}$$

The set of Eqs. (19), (22), and (23) is a complete analog for modes of the formulas obtained in the framework of the Rytov method for rays.

These relations show that despite the fact that every mode spans the whole space between its turning points, the most significant influence on its amplitude variations have the inhomogeneities located in the vicinities of the mode rays. So, each mode is most sensitive to the inhomogeneities located in two volumes surrounding its mode rays. A visual

representation of these volumes give two sets of the first Fresnel zones strung on the mode rays. Figure 2 presents examples of such volumes in the case when the field is excited by a 100-Hz source in the so-called canonical underwater sound channel.[19] The boundaries of the volumes corresponding to the 52nd mode are shown by the dashed curves.

The Rytov method reduces to the geometrical optics approximation when the radii of the first Fresnel zones of rays are small as compared with characteristic scales of inhomogeneities. Under the similar condition for the radii of the first Fresnel zones of modes, the function $\Delta_{m,q}$ reduces to

$$\Delta^{\pm}_{m,q}(x,y,z)=\delta(z-z^{\pm}_{m,q}(x))\,\delta(y-y_{m,q}(\mathbf{x})),$$

and Eq. (23) can be written as

$$X^{\pm}_{m,q}=\int_{\Gamma^{\pm}_{m,q}} ds\,\delta n, \tag{24}$$

where the integrals run along the mode ray trajectories $\Gamma^{\pm}_{m,q}$ and ds is the arc length.

The above expression and Eqs. (19) and (22) put together yield an approximate solution to the Helmholtz equation that may be called the geometrical optics for modes. Note that as in the framework of the usual geometrical optics,[19] the effect of refractive index inhomogeneities is allowed for by introducing an additional phase increment along the unperturbed geometric ray trajectories (this time along the trajectories of the mode rays).

The relations under consideration in this section are as simple and statistical averaging suited as their ray prototypes. Moreover, the statistical moments of mode complex amplitudes are expressed through exactly the same statistical characteristics of random phase increments along the ray trajectories, as the statistical moments of the ray complex amplitudes. So, to analyze the mode amplitude fluctuations and, hence, the total field fluctuations, one can use the formulas that have already been derived in works on studying the statistic of rays. For example, the expressions obtained in Refs. 19 and 20 for the phase increment variance and the phase-structure function combined with the relations presented above immediately yield the formulas to calculate the mean mode complex amplitude and the mutual correlations between the different mode complex amplitudes.[14]

The regions of validity of the approximate relations presented in this section in terms of frequency and range are practically the same as the regions of validity of the geometrical optics and the Rytov method. In the criteria given in Ref. 19 one should simply substitute the parameters of the mode rays for the parameters of the ray trajectories and replace the radii of the first Fresnel zones of rays with the radii of the first Fresnel zones of modes.

There is an additional point that deserves to be stressed when speaking about the validity region. Unlike their ray prototype, the approach under consideration remains valid not only at the regular waveguide points but on caustics as well.

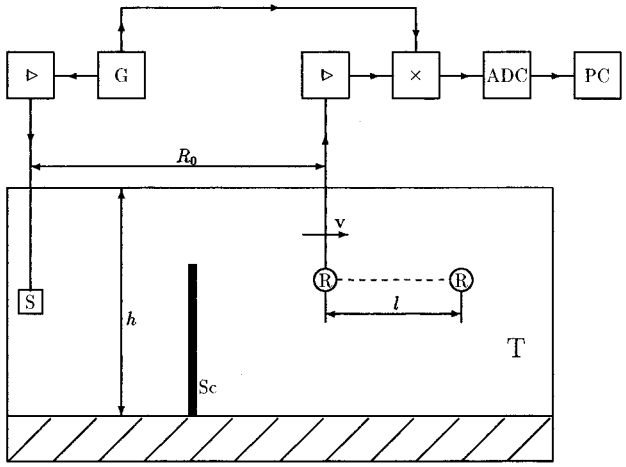


FIG. 3. The experimental setup. See text for details.

V. TANK EXPERIMENT

To demonstrate the use of the notion of the Fresnel zones for modes when describing the diffraction effects, we carried out a tank experiment on sound scattering by an opaque screen in an acoustic waveguide. To analyze the modal structure of the field, the sound pressure was registered with a point receiver uniformly moving along the waveguide with a subsequent Fourier analysis of the received signal. The idea was in selecting modes by their different Doppler frequency shifts caused by the receiver motion.[21]

A. Experimental setup

The experiment has been carried out in a laboratory tank of the dimensions $5\times0.65\times0.55$ m. The waveguide is a homogeneous water layer with a constant sound speed 1468 m/s and thickness $h=300\pm2$ mm lying over the suspended rubber bottom 4 cm thick. Our previous experiments performed in the same tank[22,23] show that this waveguide can be adequately described in the framework of the Pekeris model, where the rubber bottom is replaced with an effective fluid half-space with density 1.13 g/cm^3, sound speed 1778 m/s and attenuation of 2.8 dB/λ.

The simplified system block diagram is shown in Fig. 3. The generator (G) forms pulses at a carrier frequency $f_0=140$ kHz with duration $\tau_0=300$ μs and repetition frequency $F=63$ Hz. A carrier phase at the beginning of each pulse has had the same value. A ceramic piezoelectric transducer having the shape of rectangular parallelepiped has been used as an emitter (S). An emitting side has the dimensions: 2 mm in the vertical plane and 8 mm in the horizontal plane. The effective horizontal beamwidth is about 8° that diminishes the signal reflected from the lateral sides of the tank. In the vertical plane the radiation pattern is practically omni-directional. The pulse duration satisfies the condition $f_0\tau_0\gg1$, i.e., the radiation is quasi-monochromatic. The signal is registered by an omni-directional ceramic piezoelectric receiver (R), that is fixed on a rod which can move uniformly along the tank (T) at a constant depth. After passing through the amplifier (▷) the signal goes to the input of the

multiplier ($\times$). At the output of the multiplier we have two quadrature components of the received signal. After sampling in the analog-to-digital converter (ADC), the quadrature components are stored on the hard disk of the personal computer (PC) for further processing.

The waveguide height, h, i.e., the distance between the upper boundary of the rubber bottom and the water surface, can be measured to an accuracy of $\Delta h=2$ mm. The error is caused by the meniscus that appears after submerging the ruler in the water. The amplitude of the rubber boundary roughness does not exceed $\delta h_{\max}=0.5$ mm.

The emitter has been fixed at depth $z_0=150$ mm. The receiver has been moving along the waveguide with the velocity $v=1.800\pm0.005$ cm/s.

An opaque screen (Sc) made of rubber has blocked a part of the waveguide. The sound field has been registered behind the screen.

B. Spatial spectra of the wave field in a horizontal waveguide section

The mode amplitude attenuation caused by absorption in the bottom can be accounted for in Eq. (6) by adding the small imaginary parts $i\beta_m$ to the horizontal mode wave numbers ka_m. Then the registered signal can be represented as

$$\tilde{p}(r,z_0,z,t)=\frac{C}{\sqrt{r}}\sum_m Q_m(z_0,z)e^{-\beta_m r}\cos(\xi'_m r-\omega t), \tag{25}$$

where $Q_m(z_0,z)=\varphi_m(z_0)\varphi_m(z)$, $\omega=2\pi f$, f is the carrier frequency, t is the time, and the constant C depends on the emitter power. The product $Q_m(z_0,z)e^{-\beta_m r}$ gives the mth mode amplitude at the observation point (r,z).

To obtain the complex demodulate of the received signal, we have multiplied the latter by $\sin(\omega t)$ and $\cos(\omega t)$ and after low-pass filtering have the two quadrature components. According to Eq. (25) they are equal to

$$p_c(r,z_0,z)=\frac{C_c}{\sqrt{r}}\sum_m Q_m(z_0,z)e^{-\beta_m r}\cos(ka_m r), \tag{26}$$

$$p_s(r,z_0,z)=\frac{C_s}{\sqrt{r}}\sum_m Q_m(z_0,z)e^{-\beta_m r}\sin(ka_m r), \tag{27}$$

where C_c and C_s are the constants which can be varied by changing amplification coefficients of the equipment. In the experiment, by special choice of the amplification coefficients we have obtained equality of these constants.

The receiver has been moving with the constant velocity v from the point (R_0,z) in the direction away from the emitter. The time dependence of the quadrature components are given by Eqs. (26) and (27) with r replaced by R_0+vt.

The quadrature components obtained in such a way give the complex amplitude of the pressure field, $p(r,z_0,z)$, along the waveguide section $z=$const, where z is the vertical coordinate of the receiver. The spatial spectrum amplitude along this section is given by

$$G(\sigma,R_0)=\left|\frac{1}{l}\int_{R_0}^{R_0+l}p(r,z_0,z)e^{-i\sigma(r-R_0)}\,dr\right|. \tag{28}$$

Here l is the distance passed by the receiver during the data recording. This parameter defines the resolution in the spatial frequency (σ) domain. The spatial spectrum (28) shows the horizontal wave numbers present in the data and with a sufficient data window, l, it would resolve individual modes (a more sophisticated procedure to study the mode structure of the total field, based on the Hankel transform technique, has been described in Ref. 24).

Since the quantity $r^{-1/2}e^{-\beta_m r}$ changes little at the interval l, the above equation reduces to

$$G(\sigma,R_0)=\left|\frac{\text{const}}{\sqrt{R^*}}\sum_m Q_m e^{-\beta_m R^*+i\xi_m R_0}\frac{\sin(\sigma-ka_m)l/2}{(\sigma-ka_m)l/2}\right|, \tag{29}$$

where R^* is the distance between the source and the middle of the waveguide part passed by the receiver during the data recording.

Under the condition[21,25]

$$l\gg\frac{2\pi}{\min|k(a_{m_1}-a_{m_2})|},\quad m_1\neq m_2, \tag{30}$$

the mth mode is resolved in the spectrum (29). In our experiment this condition has not been fulfilled because of the limitations imposed by an insufficient length of the tank. Nevertheless, the spatial spectrum has allowed us to compare the contributions to the total field from different small groups of modes with neighboring numbers.

When calculating the spectrum (28) we have used an algorithm of the fast Fourier transform and to diminish sidelobes have applied the Blackman spectral window.

C. Results of tank experiment

Our analytical approach assumes smallness of the wavelength as compared with the radii of the first Fresnel zones and smallness of the latter as compared with the water layer height, h. In the experiment the wavelength was equal to 10 mm, the radii of the first Fresnel zones were of order of 100 mm, and h was equal to 300 mm. When calculating the mode parameter the algorithm proposed in Ref. 23 has been used.

The screen has been placed at a range $r_s=400$ mm from the source. It has blocked two-thirds (i.e., 200 mm) of the waveguide. Figure 4 shows the initial parts of trajectories of the mode rays corresponding to all 32 propagating modes. Note that all mode rays taking off downwards hit the screen. Since the sizes of the first Fresnel zones of all the modes are less than the screen vertical size, it means that the coefficients ξ_m^- in Eq. (15) for all modes should be much less than unity. The coefficients ξ_m^+ are maximum for the modes with numbers close to 15 and 26 (the mode rays corresponding to the 15th and 26th mode are shown by thicker solid lines than the others). The point is that the first Fresnel zones for the 15th and 26th modes (their radii are equal to 70 and 80 mm, respectively) better than others fit the aperture of the slit between the upper boundary of the screen and the water surface. Therefore the screen does not significantly weaken the

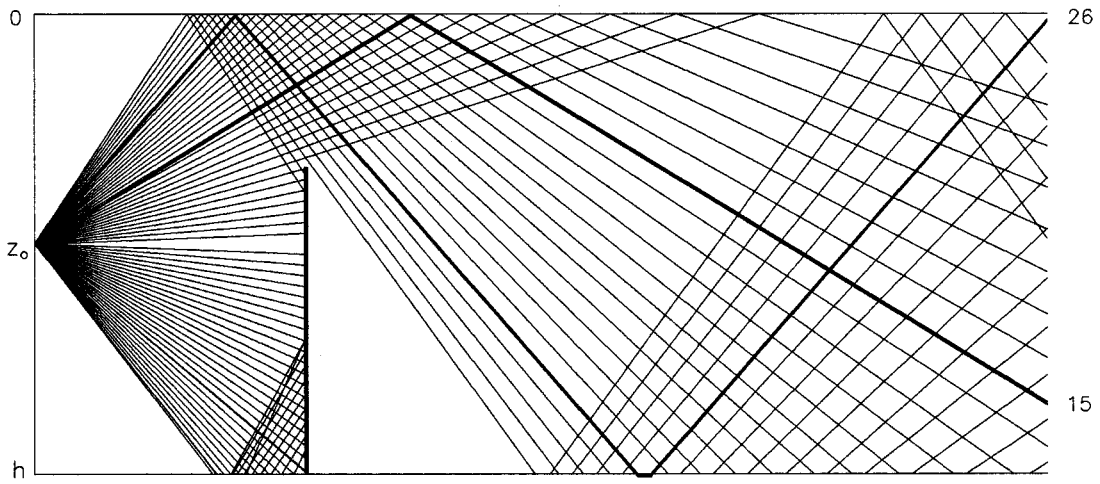


FIG. 4. The initial parts of the trajectories of the mode rays in the laboratory waveguide with the opaque screen. The mode rays corresponding to the 15th and 26th modes and taking-off upwards are shown by more thick curves than the others.

amplitudes of modes having the numbers close to 15 and 26. The amplitudes of some such modes may even increase behind the screen.

The results presented in Fig. 5 confirm this prediction. Here the spectra of two horizontal sections of the pressure field [see Eq. (29)] measured at the depth $z=30$ mm along the same interval 2900 mm in length with (solid curve) and without (dashed curve) the screen are shown. Unfortunately, the length of the horizontal section, $l=2900$ mm, used to calculate the spatial spectrum has not been large enough (due to limitations imposed by the size of the experimental tank) to resolve individual modes. The value of each spectrum at any spatial frequency is defined by contributions from a few unresolved modes. We can clearly see two intervals of spatial frequencies (shown by thick solid segments on the abscissa axis) where the field constituents, on the average, are less reduced in amplitude behind the screen than the constituents at other frequencies. Since these intervals contain the horizontal wave numbers of 15th and 26th modes (shown by arrows), we can conclude that the two groups of modes with the numbers close to 15 and 26 are indeed less weakened than other modes.

Being unable (due to the above-mentioned limitations) to estimate how accurate our approach can predict the variation of the amplitude of an individual mode, we, nevertheless, can compare the measured spectrum of the horizontal field section behind the screen and the transmission loss along the same section with the theoretical predictions of the same dependencies made using Eqs. (15)–(17). Such comparisons have been carried out in Figs. 6 and 7. Note that a discrepancy between the theoretical and experimental curves in Fig. 6 at high spatial frequencies (near $\sigma=60$ mm^{-1}), apparently is caused by the fact that the WKB approximation

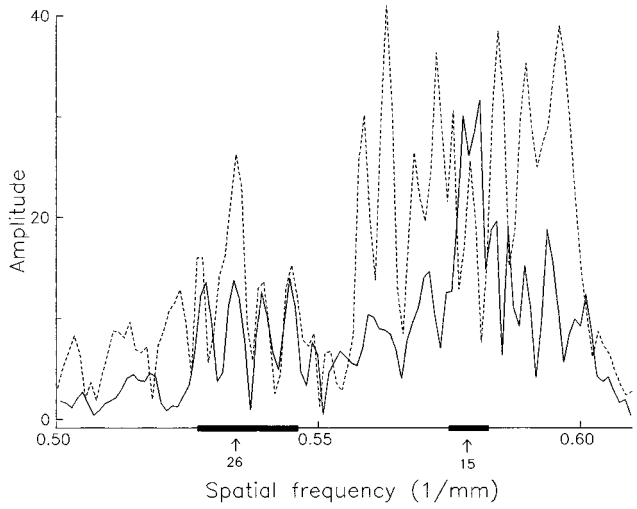


FIG. 5. Spatial spectra of the horizontal section of the waveguide without (dashed curve) and with (solid curve) the opaque screen. The arrows show the horizontal wave numbers corresponding to the 15th and 26th modes.

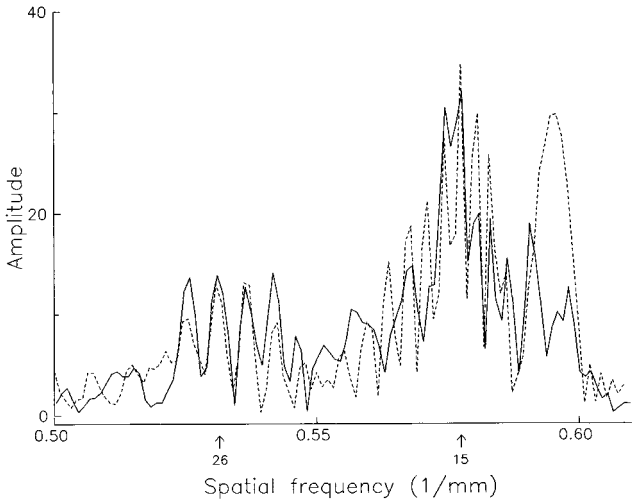


FIG. 6. Measured (the same solid curve as the one presented in Fig. 5) and predicted (dashed curve) spectra of the horizontal section of the wave field behind the screen.

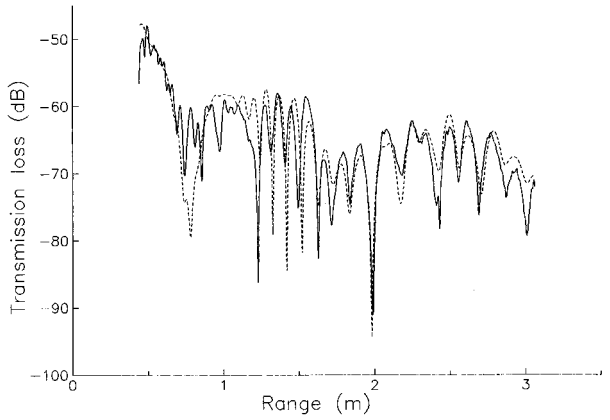


FIG. 7. Measured (solid curve) and predicted (dashed curve) transmission loss behind the screen.

used when deriving Eqs. (15)–(17), poorly describes the contributions from the low-order modes which form this part of the spectrum.

Thus, there is a reasonable agreement between the theory and the tank experiment. The discrepancies (mainly in small-scale details of the theoretical and experimental curves) are caused by (i) errors resulting from an approximate character of the theory and (ii) incomplete adequacy of the Pekeris model used in our calculations to the real waveguide where the tank experiment has been carried out. It cannot be precluded that just the second factor prevails. Numerous calculations of the sound field in the scope of the Pekeris model (that we have performed using a mode code) in the absence of the screen show that the small-scale details of the wave picture are extremely sensitive even to small variations of the parameters of the model. For example, changes of the source depth by 1 mm or the sound speed in the water layer by 1 m/s lead to considerable variations in the spatial spectrum of the pressure field. Our estimations show that the small random variations of the receiver velocity (their amplitude has been pointed out above) also has affected the results.

VI. CONCLUSIONS

The main results obtained in the framework of the approach under consideration in this paper are as follows. In each cross section of a waveguide, for the given mode, the two spots neighboring the mode rays can be pointed out where the secondary sources playing the principal role in forming the mode field are located. At short enough ranges the vertical sizes of such areas are small as compared with the distance between the mode turning points. When screens or the other inhomogeneities are placed in these spots the mode amplitude is changed more pronouncedly than in the case when the same inhomogeneities are placed far from them. In the aforementioned spots the given mode adds in phase with the neighboring modes and therefore energy exchange with the latter is most efficient when scattering takes place at large-scale inhomogeneities located there. To quantitatively describe the appropriate effects, notions of Fresnel zones for modes are introduced.

Simple and statistical averaging suited formulas are deduced to calculate mode amplitude variations under scattering at large-scale inhomogeneities which are ''mode analogs'' of well-known formulas entering the Rytov method and the method of geometrical optics. For example, the mutual correlation of different mode amplitudes is expressed through almost the same relations as the mutual correlation of ray amplitudes at different points.[13,14]

Equations (15)–(17) and (22)–(24) can be considered as the mode representations of the sound field in the waveguide with the large-scale inhomogeneities (an opaque screen and the refractive index inhomogeneities, respectively). The relations allowing for influence of the same inhomogeneities on the ray amplitudes (namely, the Huyghens–Fresnel–Kirchhoff approach and the Rytov method) can be regarded as the corresponding ray representations. These ray and mode representations can be converted into each other, but discussion of this subject is beyond the scope of the present paper.

Experimental studies performed in the laboratory to test the concept of the Fresnel zones for modes provide evidence that the theoretical predictions made in the scope of the approach under consideration are in good agreement with the experimental data.

In conclusion, note that the results analogous to the aforementioned can be obtained when describing mode scattering by surface inhomogeneities.[15,16] This subject has not been broached here. We are planning to consider it elsewhere.

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Fundamental Research (Grants No. 95-02-04565 and 94-02-04545) and by INTAS Grant No. 93-2492.

[1] I. Tolstoy and C. T. Clay, *Ocean Acoustics* (McGraw-Hill, New York, 1966).
[2] L. M. Brekhovskikh and Yu. Lysanov, *Fundamentals of Ocean Acoustics* (Springer-Verlag, Berlin, 1991).
[3] C. T. Tindle and K. M. Guthrie, ''Rays as interfering modes in underwater acoustics,'' J. Sound. Vib. **34**, 291–295 (1974).
[4] K. M. Guthrie and C. T. Tindle, ''Ray effects in the normal mode approach to underwater acoustics,'' J. Sound. Vib. **47**(3), 403–413 (1976).
[5] L. B. Felsen, ''Hybrid ray-mode fields in inhomogeneous waveguides and ducts,'' J. Acoust. Soc. Am. **69,** 352–361 (1981).
[6] A. Kamel and L. B. Felsen, ''On the ray equivalent of a group of modes,'' J. Acoust. Soc. Am. **71,** 1445–1452 (1982).
[7] T. Gao and E.-C. Shang, ''The transformation between the mode representation and the generalized ray representation of a sound field,'' J. Sound. Vib. **80**(1), 105–115 (1982).
[8] S. R. Rutherford, ''An examination of multipath processes in a range dependent ocean environment within context of adiabatic mode theory,'' J. Acoust. Soc. Am. **66,** 1482–1486 (1979).
[9] J. M. Arnold and L. B. Felsen, ''Rays and local modes in a wedge-shaped ocean,'' J. Acoust. Soc. Am. **73,** 1105–1119 (1983).
[10] S. A. Ozaki, ''A relation between normal mode and rays in underwater sound field,'' J. Acoust. Soc. Jpn. (E) **3**(1), 27–32 (1982).
[11] A. L. Virovlyansky and A. G. Kosterin, ''Fresnel volumes of modes in multimode wave-guides,'' Izv. Vyssh. Uchebn. Zaved. Radiofiz. **32**(4), 478–486 (1989).
[12] A. L. Virovlyansky and A. G. Kosterin, ''Method of smooth perturbation for the description of fields in multimode wave-guides,'' Sov. Phys. Acoust. **33**(4), 351–354 (1987).
[13] A. L. Virovlyansky, A. G. Kosterin, and A. N. Malakhov, ''Fresnel zones for modes and analysis of field fluctuations in random multimode

waveguides,'' Waves Random Media **1**(4), 409–418 (1991).
[14] A. L. Virovlyansky, A. G. Kosterin, and A. N. Malakhov, ''Mode fluctuations in a canonical underwater sound channel,'' Sov. Phys. Acoust. **35**(2), 229–235 (1989).
[15] A. L. Virovlyansky, A. G. Kosterin, and A. N. Malakhov, ''Scattering of the modes in the rough-surfaced wave-guide,'' Izv. Vyssh. Uchebn. Zaved. Radiofiz. **32**(8), 979–984 (1989).
[16] A. L. Virovlyansky, ''The mode amplitude transformation in a multimode waveguide with a quasisinusoidal boundary,'' Izv. Vyssh. Uchebn. Zaved. Radiofiz. **35**(5), 440–450 (1992).
[17] M. Born and E. Wolf *Principles of Optics* (Pergamon, Oxford, 1968).
[18] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1994).
[19] S. M. Flatte, R. Dashen, W. Munk, K. M. Watson, and F. Zachariasen, *Sound Transmission through a Fluctuating Ocean* (Cambridge U. P., Cambridge, 1979).
[20] R. Esswein and S. M. Flatte, ''Calculation of the phase-structure density from oceanic internal waves,'' J. Acoust. Soc. Am. **70,** 1387–1396 (1981).
[21] N. V. Gorskaya, G. N. Nicolaev, T. A. Rychova, and B. M. Salin, ''Spectral analysis in the investigation of fields of harmonic source in acoustic wave-guides,'' Sov. Phys. Acoust. **27**(2), 110–112 (1981).
[22] V. V. Kurin, O. V. Lebedev, and N. V. Pronchatov-Rubtsov, ''Experimental acoustic studies in wave-guides weakly inhomogeneous along the track,'' Acoust. Phys. **40**(3), 435–436 (1994).
[23] O. V. Lebedev, N. V. Pronchatov-Rubtsov, and S. I. Simdyankin, ''To calculation of sound field in layered acoustic waveguide with absorbing bottom,'' Acoust. Phys. **42**(1), 76–82 (1996).
[24] G. V. Frisk and J. F. Lynch, ''Shallow water waveguide characterization using the Hankel transform,'' J. Acoust. Soc. Am. **76,** 205–216 (1984).
[25] V. A. Eliseevnin, ''Operation of a horizontal line array in a water layer,'' Sov. Phys. Acoust. **25**(2), 123–126 (1979).

Direct optimization methods, ray propagation, and chaos. I. Continuous media

Martin A. Mazur and Kenneth E. Gilbert
The Applied Research Laboratory and the Graduate Program in Acoustics, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

(Received 19 February 1996; accepted for publication 25 July 1996)

In classical ray tracing, eigenrays between a source and receiver are determined by an initial value or ''shooting'' approach. The launch angles of rays from a source point are varied until the rays intersect the receiver endpoint. In nonseparable range-dependent environments, the ray paths can be chaotic, putting a fundamental limit on tracing rays by shooting methods. In the present paper, an alternative approach based on Fermat's principle of minimum propagation time is discussed. Rather than minimizing the travel time integral indirectly, by deriving the Euler–Lagrange equations for the ray paths, so-called ''direct methods'' of minimization, taken from the calculus of variations, are employed. Previous authors have demonstrated that simulated annealing, a direct method, can be used to find eigenrays in a particularly chaotic ray tracing problem. In the present paper, two direct methods are applied to the calculation of eigenrays in continuous media: the Rayleigh–Ritz technique, a classical direct method, and simulated annealing, a Monte Carlo direct method. Direct methods are compared to shooting techniques, and some of the advantages and drawbacks of both methods are shown using both nonchaotic and chaotic examples. © *1997 Acoustical Society of America.* [S0001-4966(97)04312-9]

PACS numbers: 43.30.Cq [MBP]

INTRODUCTION

In classical ray tracing, eigenrays between a source and receiver are determined by an initial value or ''shooting'' approach. The launch angles of rays from a source point are varied until the rays intersect the receiver endpoint. In range-dependent environments the ray paths can be chaotic.[1,2] Under chaotic conditions, the rays determined by shooting are extremely sensitive to initial conditions and the number of eigenrays connecting two points grows exponentially with range.[3,4] These two facts put fundamental limits on the accuracy and even the feasibility of using shooting to find eigenrays. An alternative approach has been suggested by the original work of Collins and Kuperman[5,6] and further examined by us in Ref. 7. In this alternative approach, the problem of finding eigenrays is cast in terms of Fermat's principle of minimum propagation time. Rather than minimizing the travel time integral indirectly by deriving the Euler–Lagrange equations for the ray paths, we use so-called *direct methods* of the calculus of variations.[8,9] Collins and Kuperman demonstrated that a direct method, simulated annealing, could be used to find eigenrays in a particularly chaotic ray tracing problem. Our work investigates the ideas of Collins and Kuperman within the general framework of direct minimization algorithms. Direct methods include classical techniques such as the Rayleigh–Ritz method, and Monte Carlo techniques such as simulated annealing[10,11] and genetic algorithms.[12] In the present paper, we apply the Rayleigh–Ritz method and simulated annealing to a variety of problems. We compare direct methods to shooting techniques, and show some of the advantages and drawbacks of both approaches in both nonchaotic and chaotic problems.

Direct methods have in common the following two properties: (1) Direct methods attempt to minimize the travel time integral directly, rather than by finding the minimizing path by way of a set of derived differential equations. The direct minimization is accomplished by reducing the inherently infinite-dimensional optimization problem to a finite-dimensional problem. The solution of this finite-dimensional problem approximates the eigenray path, and the accuracy of the approximation can be improved as desired. (2) Direct methods require a trial solution to begin the optimization procedure. In cases where multiple local minima (multipaths) exist, the minima can, in principle, be found by varying the trial solution.

In the present paper we limit our analysis primarily to continuous media. With modifications, direct methods can be applied to problems with discontinuities or singularities in the properties of the medium or in the boundaries. A companion paper[13] deals with a subset of such problems that can be analyzed in terms of discrete maps.

Theoretically, for continuous media, the integral and differential formulations of Fermat's principle are equivalent.[8] In problems with discontinuities or singularities in the properties of the medium or in the boundaries, however, there may exist ray paths that cannot be found by shooting. We have used direct methods to find these rays. In the present paper, we will give an example where there exist rays that do not satisfy the Euler–Lagrange equations everywhere, and that, in fact, are the envelopes of the rays that do satisfy these equations. Other examples of arrivals that do not satisfy the ray equations are the rays associated with ''head waves.'' More discussion of ray propagation in discontinuous media, including head wave phenomena, is given in the companion paper.[13]

In practice, the advantages of direct methods over shooting are that direct methods are insensitive to small perturba-

tions in the endpoints and in the parameters of the intervening environment, and that direct methods tend to produce the earliest arrivals. However, in our experience, direct minimization methods can reproduce the full arrival structure only with good trial solutions, presupposing knowledge of the arrival structure.

In range-dependent environments where the Euler–Lagrange equations for the rays cannot be solved by separation of variables, the equations may not be integrable and the ray paths are generally chaotic.[1,2] Henceforth, we refer to such range-dependent environments as *nonseparable* environments. In nonseparable environments, chaotic ray paths can exist even if the range dependence is weak. The calculation of chaotic ray paths using the shooting method is extremely sensitive to initial conditions. In addition, as mentioned above, the number of rays connecting two points under chaotic conditions grows exponentially with range. Consequently, with the shooting method in such environments, there is a fundamental limit on the accuracy of ray tracing, even with perfect knowledge of environmental conditions. In practice, in sufficiently benign chaotic environments, shooting methods, used with sufficient computer word length and care, can be used to the find the entire multipath structure between two points that are not spaced too far apart.[4] With a fixed word length and increasing range, it generally becomes impossible to find even one eigenray using the shooting method.

Recently, Collins and Kuperman[5,6] demonstrated in a numerical example that eigenrays can be found when ray chaos exists if the problem is recast in terms of Fermat's principle of minimum propagation time. The problem then becomes amenable to techniques of the theory of two-point boundary-value problems and of optimization theory. Working with a particularly chaotic example problem, Collins and Kuperman showed that shooting methods could not successfully find an eigenray between two points when specifying initial conditions with the computer word length available to them. A relaxation technique for the solution of two-point boundary-value problems, the ''bending'' method,[14] was also investigated. The relaxation technique was shown not to converge unless a very good trial solution was provided by a direct optimization technique. It was only when the travel time integral was minimized using a direct optimization technique called simulated annealing[10,11] that an eigenray was produced for their example.

By investigating a number of test problems, we have determined that direct methods, because of their insensitivity to endpoint conditions, can be used to produce earliest arrivals under chaotic conditions.[7] On the other hand, in chaotic problems the number of rays connecting two points under chaotic conditions grows exponentially with range.[3,4] It thus becomes difficult to reproduce the complex arrival structure because the required trial solutions presuppose knowledge of it.

In Sec. I, we discuss the various approaches to the calculation of ray paths, particularly the approaches derived from Fermat's principle. We describe two direct methods, Rayleigh–Ritz and simulated annealing. In Sec. II, we apply direct methods to some nonchaotic problems in both continuous media and to an example with a singularity. We show the advantages and limitations of the direct approach. In Sec. III, we briefly discuss the application of direct methods to nonseparable environments, using the example problem of Collins and Kuperman. Finally, in Sec. IV we draw some conclusions regarding the use of direct methods in continuous media. The application of direct methods to chaotic environments will be explored in more depth in the companion to this paper,[13] where propagation in environments with discrete transitions is investigated.

I. FERMAT'S PRINCIPLE AND DIRECT OPTIMIZATION METHODS

A. Approaches to the calculation of ray paths

In all of the variety of ways to derive the differential equations for acoustic ray paths, there are two basic approaches. The first is via the wave equation. The derivations using the wave equation apply the method of characteristics to the eikonal (infinite frequency) solution of the wave equation. This technique yields differential equations for the paths that are normal to the local wavefronts. The paths are the rays. The second approach is through Fermat's principle, which states loosely that in optics, of all possible paths connecting two points, light takes the path requiring the least time. At high frequencies, Fermat's principle can be applied to sound as well. The propagation time along an arbitrary path γ is

$$T(\gamma)=\int_{\gamma}\frac{ds}{c}, \tag{1}$$

where c is the speed of sound and s is arc length. A more precise statement of Fermat's principle is that T is stationary with respect to path perturbations if γ is a ray path. The differential equations for the ray paths are the Euler–Lagrange equations derived from Eq. (1) using techniques of the calculus of variations. A number of forms of these equations can be derived using various choices of path parametrization. The various forms are related to each other and identical to those that can be derived from the wave equation.

In some literature, γ is called an *extremal* of T if the Euler–Lagrange equations are satisfied along γ. It is often pointed out that this is a confusing terminology, since γ being an extremal is not sufficient to prove that T is a local *extremum* on γ with respect to nearby paths. In fact, there is some confusion in the literature over whether Fermat's principle is to be interpreted as requiring (local) *minimum* time or *stationary* time for ray paths. We make the following remarks: (1) The classical equations for rays derived as the characteristics of the wave equation are equivalent to the Euler–Lagrange equations as derived from the travel time integral using Fermat's principle. (2) The ray equations yield extremals, paths of stationary travel time, which are not guaranteed to locally extremize travel time. (3) The approach we have explored, i.e., direct minimization, is designed to find rays for which the travel time is a minimum, either local or global. In our experience, we have always been able to reproduce by direct methods rays that have been calculated using methods based on the Euler–Lagrange equations. We

do not explore here the nature of the extremal paths (i.e., local extremum, saddle curve) found by either set of methods. (4) In many physical problems, the intersection of these two sets, extremizing and extremal ray paths, is large or exhaustive. (5) Under chaotic conditions, the size of the intersecting set is unclear and a matter for study.

Strategies for obtaining ray paths have generally centered on the ordinary differential equations described above. The equations are usually solved using an initial value or ''shooting'' technique, even when eigenrays between two boundary points are desired. In the shooting technique, a fan of rays is propagated from a source point by varying the launch angle of the rays. Those rays that intersect the receiver endpoint are eigenrays. The shooting method is computationally fast and reliable for finding the spread of ray paths from a source in a range-independent environment, and can be efficient for finding eigenrays between two points under nonchaotic conditions. However, the ray equations are nonlinear, and in range-dependent environments they are generally nonseparable. In such environments, shooting techniques are prone to chaos. At shorter distances, it is possible to calculate a complete set of eigenrays between two points via the shooting method.[4] However, for any fixed available computer word length, the calculation of eigenrays in nonseparable environments becomes intractable at long ranges due to the extreme sensitivity of chaotic ray paths to initial conditions.

In Ref. 6, Collins and Kuperman explored solving the eigenray problem as a two-point boundary-value problem. When the eigenray problem is cast as a two-point boundary-value problem, the Euler–Lagrange differential equations can be solved numerically using methods such as relaxation or collocation. In these boundary-value techniques, the differential equations are replaced by linearized finite difference equations over a grid of points. A trial solution is specified at the grid points, and the values at these grid points are successively adjusted until the solution agrees with the finite difference equations and the boundary conditions. Collins and Kuperman used a relaxation technique, the bending method,[14] in attempting to calculate eigenrays. They found that although the technique is fast, it requires a very good trial solution in nonseparable environments, and can fail to converge. They were successful only when a direct optimization technique, the simulated annealing method discussed below, was used to provide a good trial solution.

The eigenray problem can also be attacked, as suggested by the work of Collins and Kuperman,[6] by considering so-called *direct methods* for minimizing the integral of Eq. (1). Direct methods attempt to find paths that minimize the integral of Eq. (1) directly, rather than through the Euler–Lagrange equations. Direct methods replace the infinite-dimensional minimization problem with a finite-dimensional one by restricting the class of functions under consideration to a class that can be represented using a finite number of adjustable parameters. A minimizing sequence of such representations is then constructed whereby the number of adjustable parameters is increased and the infinite-dimensional minimum is attained to the desired accuracy through a limiting process.

B. Classical direct minimization: The Rayleigh–Ritz method

In classical direct minimization, the idea is to reduce the infinite-dimensional minimization problem from the calculus of variations to a finite-dimensional problem whose solution approximates that of the original problem. The finite-dimensional problem is solved by techniques of finite-dimensional calculus. The accuracy of the approximation can be increased by considering a higher finite-dimensional parameter space. Below we discuss the Rayleigh–Ritz method.

In the travel time integral in Eq. (1), we assume that γ can be parametrized as a curve $[r(s),z(r(s))]$, with r a monotonically increasing function of arc length s, and z a continuous function of r. Henceforth, we shall refer to the continuous depth functions $z(r)$ that satisfy $z(r_0)=z_0$ and $z(r_f)=z_f$ as *admissible* functions. Then Eq. (1) can be rewritten as

$$T[z]=\int_{r_0}^{r_f}\frac{\sqrt{1+(dz/dr)^2}}{c(r,z(r))}\,dr. \tag{2}$$

The problem is to minimize T over all admissible $z(r)$. In the Rayleigh–Ritz method, we assume that we can rewrite admissible functions in the form

$$z(r)=\sum_{i=1}^{\infty}a_i\phi_i(r), \tag{3}$$

where the ϕ_i form a complete set of basis functions.

The Rayleigh–Ritz method consists of first truncating the series to N terms

$$\hat{z}_N(r)=\sum_{i=1}^{N}a_i\phi_i(r), \tag{4}$$

and then minimizing $T[z]$ over the subset of admissible functions defined by Eq. (4):

$$T_N=\min_{\{a_i\}}T[\hat{z}_N]. \tag{5}$$

The problem is now one of adjusting N constants a_i, and hence is finite dimensional. A necessary condition for the travel time T_N to be a minimum is that the set of coefficients $\{a_i\}$ be a solution of the N simultaneous nonlinear equations

$$\frac{\partial}{\partial a_i}T[\hat{z}_N]=0,\quad i=1,\ldots,N. \tag{6}$$

It can be shown (see Ref. 8) that if the ϕ_i form a complete basis set for the class of admissible functions, then T_N is a decreasing function of N and

$$\lim_{N\to\infty}T_N=T_{\min}, \tag{7}$$

where $T_{\min}$ is the (local) minimum travel time for any ray path. The number of terms, N, necessary for an adequate solution depends on the problem and the accuracy desired.

In a numerical implementation of the Rayleigh–Ritz method, we would write the admissible functions in the form

$$\hat{z}_N(r)=mr+b+\sum_{i=1}^{N}a_i\phi_i(r). \tag{8}$$

Here m is the slope and b the z-intercept of the line connecting the endpoints of the eigenray, and the ϕ_i are from a set of basis functions satisfying $\phi(r_0)=\phi(r_f)=0$. Removing the line from the series representation improves the convergence properties of the series and converts the optimization problem to an unconstrained problem: the basis functions automatically satisfy the boundary conditions.

Two basis sets suggest themselves. One is the so-called ''hat functions,'' defined as

$$\phi_i(r)=\begin{cases} 0, & r_0 \leqslant r \leqslant r_{i-1} \\ \dfrac{r-r_{i-1}}{\Delta r}, & r_{i-1} < r \leqslant r_i \\ \dfrac{r_{i+1}-r}{\Delta r}, & r_i < r \leqslant r_{i+1} \\ 0, & r_{i+1} < r \leqslant r_f, \end{cases} \tag{9}$$

where r_0 is the initial range, r_f is the final range, and Δr is the spacing between grid points. Because the line connecting the boundary points is removed from the basis function expansion, there is no need for a hat function at either boundary grid point.

Another convenient basis set is given by the sinusoids:

$$\phi_i=\sin\left(\frac{2\pi i}{\Lambda}\, r\right), \tag{10}$$

where $\Lambda=2(r_f-r_0)$.

The advantage of the hat functions is that the trial solution for the ray is easily specified at the grid points. The disadvantages are: (1) the optimization converges slowly with a poor trial solution; (2) a large number of grid points must be specified in most problems, resulting in high dimensionality and slow convergence.

The advantage of the sinusoidal basis is that a relatively small number of basis functions will often reveal the structure of the eigenray, decreasing dimensionality and computation time. The main disadvantages are: (1) the correspondence between the Fourier coefficients and the ray path shape is not easy to visualize. Thus, construction of a trial solution is difficult, since the trial ray path must be Fourier decomposed to obtain the trial coefficients. (2) The calculation of the travel time requires more complex and time-consuming numerical integration than for the hat functions.

There are some subtle and delicate mathematical questions left open by the above presentation. For example, there may be no *admissible* ray path for which the minimum $T_{\min}$ is actually attained. (The limit of T_N is guaranteed to achieve $T_{\min}$, but the limiting ray path that attains this travel time may be discontinuous.)

Some illustrative examples of the Rayleigh–Ritz method will be given in Sec. II. The Rayleigh–Ritz method, although appealing theoretically, is much slower and more difficult to work with than the Monte Carlo technique described next.

C. Monte Carlo direct minimization: Simulated annealing

A number of modern optimization techniques attempt to mimic certain natural processes that optimize a state of nature. In Ref. 6, Collins and Kuperman showed that one such direct method, simulated annealing, can be used to find eigenrays in a nonseparable environment. Simulated annealing is a Monte Carlo technique that is a numerical analogue to the physical process of the slow cooling of materials. When some materials are allowed to cool slowly, they settle on the lowest attainable energy state because random molecular motion at intermediate temperatures always allows the system to be able to jump out of local energy minima. In numerical simulated annealing, the variables in the space over which minimization is desired are varied randomly from an initial state. The maximum size of the allowable variations is an adjustable parameter. Variations that produce a decrease in the system energy (here identified with travel time) are allowed. On the other hand, those variations that increase the system energy are also allowed, but with a probability given by an exponential distribution that has a thermodynamic analog. The probability depends on the energy difference between the states and the current ''temperature'' of the system. System ''temperature'' is decreased at a rate that can be adjusted as a parameter of the algorithm, and the system is generally attracted to a globally minimum energy state.

Collins and Kuperman used ''quenched'' annealing, which speeds the minimization process, but allows the system to settle into local minima. Quenching is useful if the locations of the various local minima (in this case, acoustic multipaths) are sought. Simulated annealing was used by Collins and Kuperman to calculate eigenrays in a range-dependent sound-speed profile (described in Sec. III below) that produces particularly chaotic ray paths. In the work of Collins and Kuperman with simulated annealing, the parameter space is a grid of ray depths at equally spaced ranges. In order to calculate longer-range eigenrays using simulated annealing with such a grid, the dimension of the parameter space must be increased in proportion to the distance between the endpoints.

In the next section, we will discuss some of the practical differences between the simulated annealing algorithm and the Rayleigh–Ritz method, using nonchaotic examples. There are other recently developed Monte Carlo direct methods, such as genetic algorithms,[12] for the solution of minimization problems. There has also been significant recent research on classical minimization methods.[15] In the present work we limit our analysis to simulated annealing and the Rayleigh–Ritz method.

II. DIRECT OPTIMIZATION IN NONCHAOTIC EXAMPLES

We now apply direct optimization techniques to a few simple, nonchaotic example problems. First, consider a point source located at the origin of a coordinate system in an unbounded medium with a constant sound speed c_0. If r and z are the axes of the coordinate system and θ is the angle with respect to the r axis, the Euler–Lagrange equations are simply

$$\frac{dz}{dr}=\tan\,\theta_0, \tag{11}$$

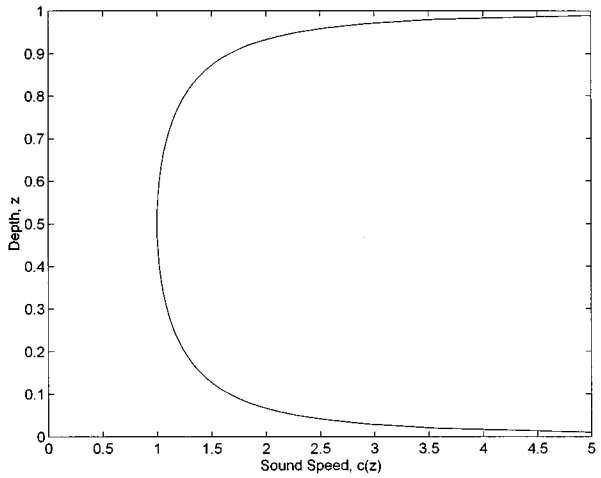


FIG. 1. Sound-speed profile $c(z)$ for the harmonic oscillator waveguide in Eq. (13). Here the waveguide parameters are $c_a=1$ and $z_a=1/2$.

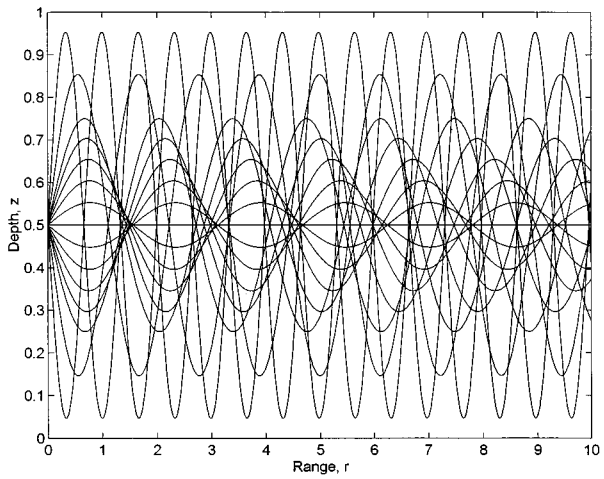


FIG. 2. A sample of rays in the harmonic oscillator waveguide. Here $c_a=1$, $z_a=1/2$ and $\theta_0=0°$, $\pm 6°$, $\pm 12°$, $\pm 18°$, $\pm 24°$, $\pm 30°$, $\pm 45°$, $\pm 65°$.

$$\frac{d\theta}{dr}=0. \tag{12}$$

Thus for each initial angle θ_0 the equations define a line whose angle is the initial angle. A simple heuristic argument shows that the travel time integral is minimized by such a line as well. Suppose we desire to find an eigenray connecting the origin to a point (r,z). Looking at Eq. (1), we see we can factor c out of the integral, since it is a constant. Then, since the shortest distance between two points is a straight line, the travel time integral is minimized by the line connecting (r,z) to the origin.

Next, we consider a more complex problem. In the following example we will illustrate some of the complications that a singular or discontinuous medium can cause. We consider a medium whose sound speed is given as

$$c=\frac{c_a}{\sqrt{1-(1-z/z_a)^2}}. \tag{13}$$

The sound-speed profile for $c_a=1$ and $z_a=1/2$ is illustrated in Fig. 1. Note that the medium is undefined outside $0<z<2z_a$ and is singular at $z=0$ and $z=2z_a$. The singular sound-speed profile will trap all rays between $z=0$ and $z=2z_a$. This duct is known as the *harmonic oscillator waveguide*.[16] The ray equations in Hamiltonian form are

$$\frac{du}{dr}=\frac{-p/z_a}{(1-u^2-p^2)^{1/2}}, \tag{14}$$

$$\frac{dp}{dr}=\frac{u/z_a}{(1-u^2-p^2)^{1/2}}, \tag{15}$$

where $u(r)=1-z/z_a$ and p is the conjugate variable to the depth z.

For rays launched from $r=r_0$, the solution of the above second-order system can be found after some algebra. It consists of a family of sinusoids

$$z(r)=z_a\left[1+\sin\,\theta_0\,\sin\left(\frac{r-r_0}{z_a\,\cos\,\theta_0}\right)\right], \tag{16}$$

parametrized by θ_0, the initial launch angle of the ray. A sample of these rays is shown in Fig. 2.

The harmonic oscillator waveguide was specifically chosen to demonstrate some of the properties of direct methods and contrast them with properties of shooting methods. The ray structure for a harmonic oscillator waveguide can be described analytically, and the shooting technique can be used to easily find eigenrays. Direct methods, on the other hand, are very sensitive to the presence of the ''fast'' boundaries. Suppose we are trying to find an eigenray between two points on the axis of the sound channel, spaced some distance apart. Given an arbitrary trial solution, the direct methods will generally return an eigenray that abruptly (i.e., as a step function in z) heads for one or the other of the boundaries (depending on the trial solution), stays at the boundary until the range of the far endpoint is reached, then heads abruptly for the sound channel axis. Figure 3 shows the two ''fast'' eigenrays. These ''rays'' are not derivable from the ray equations, and result from the singularities in the sound speed at the boundaries of the waveguide. The rays, in fact,

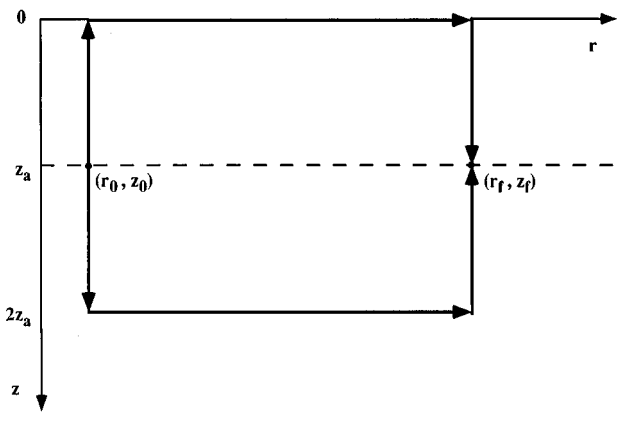


FIG. 3. Two ''fast'' eigenrays in the harmonic oscillator waveguide connecting two points on the axis, $z=z_a$, of the sound channel. Together, the two eigenrays form the envelope of the rays that can be found via the shooting method, but the eigenrays cannot themselves be calculated by shooting.

form an *envelope* of the rays produced by the Euler–Lagrange equations. They are similar in nature to the rays associated with the so-called ''head waves'' produced when sound propagates over a fast bottom.[17] The phenomenon of head waves is discussed in the companion paper,[13] in connection with ray propagation in discontinuous media.

In our example, the ''head wave'' solution is obstinate. It is the fastest ray between the endpoints. The other sinusoidal rays will not be produced unless fairly good trial solutions are provided. Thus, unless something of the ray structure is known *a priori*, only the fastest ray will be produced by direct methods. Producing other rays can become a guessing game.

We have applied direct methods to other somewhat more realistic cup-shaped ducts, such as the canonical Munk profile. In such cases, the fastest eigenray produced tends to be the one that quickly heads for the faster sound speeds and spends most of its trajectory there. Other rays are not as difficult to produce as with the harmonic oscillator duct, but the procedure for finding them tends to be trial and error unless a good trial solution is available.

III. DIRECT OPTIMIZATION IN AN EXAMPLE WITH CHAOS

In this section, we apply direct methods to the problem of finding eigenrays in a nonseparable environment. In order to apply direct methods to a simple example, and to have a basis for comparison with previous results, we use the range-dependent sound-speed profile investigated by Collins and Kuperman in Ref. 6. The profile, which applies in an unbounded domain, is given by

$$c(r,z)=c_0(1+\alpha\ \cos K_r r\ \cos K_z z), \tag{17}$$

where $c_0=1.5$ km/s, $\alpha=0.02$, $K_r=2\pi/20$ km^{-1}, and $K_z=2\pi/0.5$ km^{-1}. The profile consists of repeated ducts in depth, modulated by a sinusoidal variation in range. Our numerical calculations and those of Collins and Kuperman demonstrate the existence and degree of ray chaos for the profile in Eq. (17).

A. Application of direct methods

In our example problem, we attempted to find eigenrays in the sound-speed profile of Eq. (17) between the points (r_0,z_0) and (r_f,z_f), where $r_0=0$ km, $z_0=0.2$ km, $r_f=100$ km, and $z_f=0.2$ km. Collins and Kuperman treated this example in Ref. 6. With the computer word length available to them, Collins and Kuperman demonstrated that shooting methods could not successfully find an eigenray between these two points. However, using simulated annealing, they found an eigenray. The best solution they could achieve by shooting, obtained by using a launch angle of $-15.167\ 954\ 676°$, captured only the first part of the eigenray they had calculated with simulated annealing. Here, we first show our results using the Rayleigh–Ritz method, and later discuss our results using simulated annealing.

The solid curve in Fig. 4 shows the eigenray found by Collins and Kuperman using simulated annealing. We duplicated their calculation, and also found this ray using the Rayleigh–Ritz method. Also shown in Fig. 4 are a number of

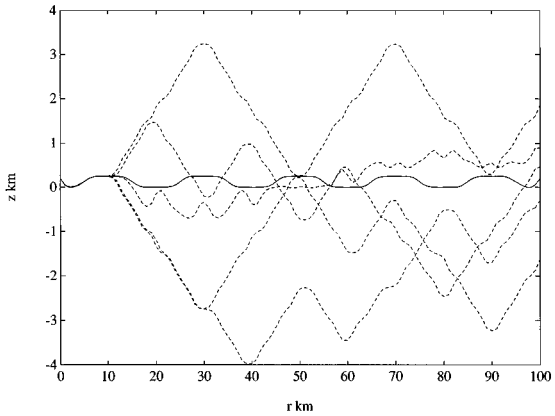


FIG. 4. Solid curve: Eigenray for the Collins–Kuperman example problem found using the Rayleigh–Ritz method with 51 sinusoidal basis functions. Dashed curves: Five rays calculated using shooting with launch angles within 10^{-5} deg of $-15.167\ 954\ 676°$. The dashed curves capture only the first part of the solid eigenray and are the best match we could achieve using shooting.

rays calculated using the shooting method with launch angles within 10^{-5} deg of the above mentioned launch angle. The figure illustrates the extreme sensitivity to initial conditions exhibited by rays in nonseparable environments and the difficulty that may be encountered in computing eigenrays using the shooting method.

For the first demonstration of the Rayleigh–Ritz method, we choose the basis functions ϕ_i to be the ''hat'' functions defined earlier. Figure 5 shows the eigenray calculated using the Rayleigh–Ritz method with the hat function basis over a grid of 51 points. We remark that despite the rather crude approach of using hat functions over only 51 points, the Rayleigh–Ritz method produced a reasonable approximation of the solid eigenray in Fig. 4. With enough points, the hat function duplicates the solution of Collins and Kuperman. We make the following further observations about the choice of hat functions: (1) Hat functions yield an eigenray whose first derivative is discontinuous at every grid point. (2) The basis functions are local descriptions of the rays. Consequently, to get the same resolution for a longer-range ray, proportionately more grid points must be taken. (3) There is no additional information about the ray to be had

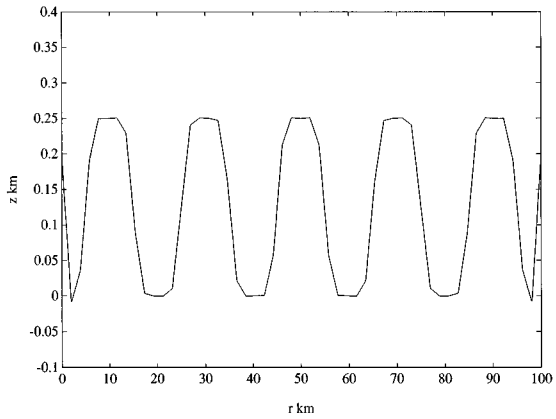


FIG. 5. Eigenray for the Collins–Kuperman example problem calculated using the Rayleigh–Ritz method with 51 ''hat'' basis functions.

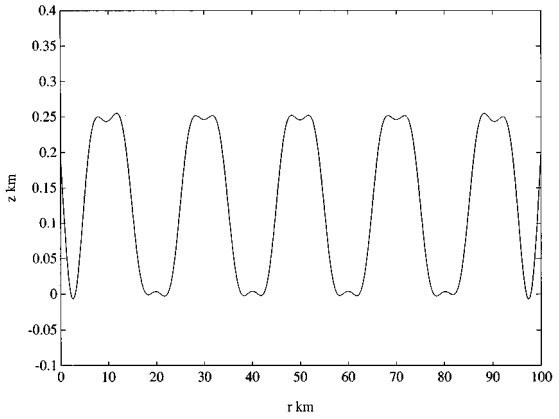


FIG. 6. Eigenray for the Collins–Kuperman example problem calculated using Rayleigh–Ritz method with 31 sinusoidal basis functions.

from the coefficients of the representation. The coefficients are simply the values of the ray depth at the grid points. The observations made here are all in contrast to the properties of the sine series representation, which we discuss next.

For the second demonstration of the Rayleigh–Ritz method, we use sinusoidal basis functions given by

$$\phi_i = \sin\left(\frac{2\pi i}{\Lambda}\, r\right), \tag{18}$$

where $\Lambda = 2(r_f - r_0)$. Figures 6 and 7 show the eigenray for the example as calculated using 31 and 91 coefficients, respectively. We can make the following observations regarding the use of sinusoidal basis functions: (1) The sinusoidal basis functions yield a ray that has continuous derivatives of all orders. (2) The sinusoidal basis functions are not local, i.e., each is nonzero everywhere but at a finite number of points. Therefore, global properties of the eigenray can begin to emerge for very low N. The lowest number of coefficients to yield a good representation of the eigenray is $N=31$, but certain characteristics (such as the five hump structure and the limitation of the ray in depth) start to appear with smaller N. (3) The coefficients of the sinusoidal representation comprise a spectrum of the spatial variation in the eigenray. Figure 8 graphs the coefficients of the eigenrays shown in Figs. 6 and 7. The information contained in these coefficients can

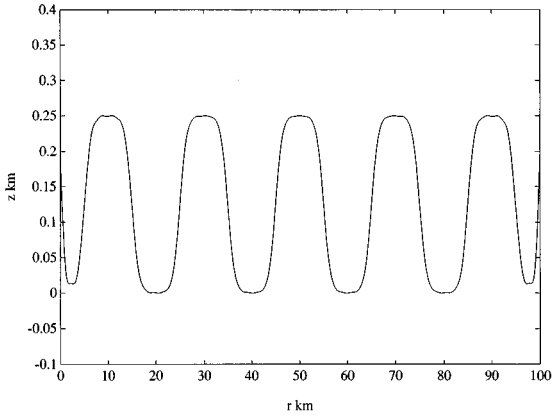


FIG. 7. Eigenray for the Collins–Kuperman example problem calculated using Rayleigh–Ritz method with 91 sinusoidal basis functions.

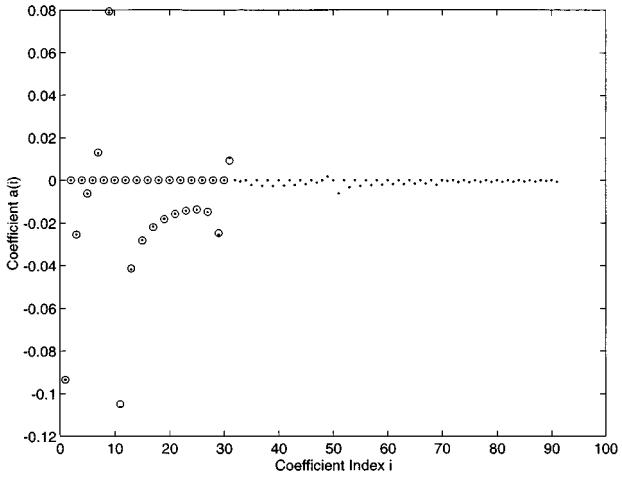


FIG. 8. Coefficients of sinusoids for an eigenray in the Collins–Kuperman example problem with $N=31$ ('$\circ$') and $N=91$ ('$\cdot$').

be helpful in generating trial solutions for finding longer-range eigenrays, or for finding eigenrays in slightly perturbed sound-speed profiles. In our example problem, for instance, we have determined that an oscillatory structure similar to that shown in Fig. 4 for the near-axis eigenray is maintained by near-axis eigenrays when r_f is increased. The large components of the spatial frequency of the near-axis eigenray (see Fig. 8) at odd multiples of 1/20 km^{-1} is also evident in longer-range eigenrays. In other problems with less evident symmetry, there are often oscillatory components to some of the eigenrays. There are also often spatial frequencies that predominate, and a ''bandwidth'' beyond which it is inefficient to calculate Fourier coefficients. Once the Fourier coefficients are established for a given problem, the information can be useful in choosing coefficients for trial solutions for longer-range rays or perturbed sound-speed profiles.

We have reproduced the eigenrays shown in this example using simulated annealing as well. The main advantages of simulated annealing over Rayleigh–Ritz are the simplicity of implementation of the algorithm, and its speed. Unlike Rayleigh–Ritz, no sophisticated numerical integration or gradient finding is required for simulated annealing. Depending on the ranges involved, simulated annealing can run one to several orders of magnitude faster than the Rayleigh–Ritz method.

B. Some questions concerning direct methods and shooting in nonseparable environments

For ray tracing in chaos-producing environments, the following questions have been explored: (1) What is the relationship between the chaotic rays produced by shooting and the eigenrays found by direct methods? (2) How do these methods compare in practice? (3) Do both of these methods yield ''true'' rays in nonseparable environments? Below, we discuss these questions in more detail.

To partly answer the first question, we note first that theoretically, in continuous media, the integral and differential formulations of Fermat's principle are equivalent. In practice, in every problem in continuous media that we in-

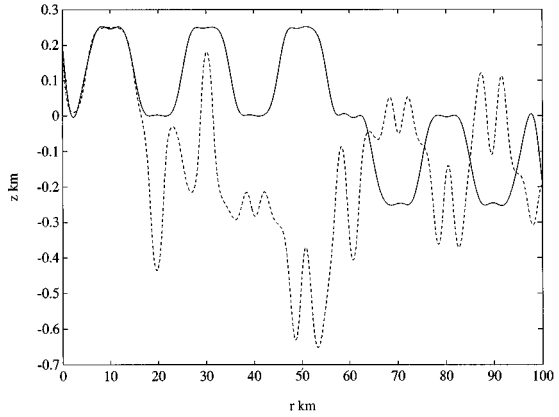


FIG. 9. Dashed curve: Ray for the Collins–Kuperman example calculated using shooting starting at $r_0=0.0$ km, $z_0=0.2$ km with a launch angle of $-15.167\,954\,676°$. Left endpoint is $r_f=100$ km, $z_f=-0.1945$ km. Travel time is 67.1622 s. Solid curve: Eigenray between same endpoints found using the Rayleigh–Ritz method with 51 sinusoidal basis functions. Travel time is 65.9498 s.

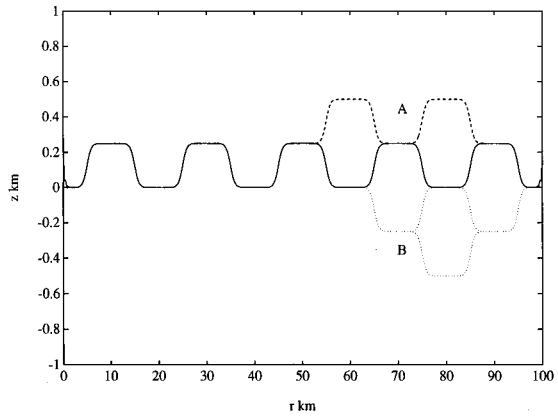


FIG. 10. There are three distinct eigenrays that follow a portion of the dashed path (A) for part of their journey between the endpoints, and the solid path for the remainder. Extrapolating leads to a mesh with exponential growth in the number of eigenrays with range. The dotted paths (B) show the beginnings of such a mesh. All rays shown have the same travel time to within roundoff error.

vestigated, every eigenray we could calculate using shooting was reproducible by direct methods. On the other hand, in problems with discontinuities or singularities in the properties of the medium or in the boundaries, the two formulations may not be equivalent, as we have seen. It is often the case in such problems that there exist rays that cannot be found by shooting which have shorter travel times than rays that can be found by shooting. In our experience, direct methods can calculate these eigenrays. More example problems in discontinuous media are given in the companion paper.[13]

In exploring the relationship between the two methods in practice, we compared the travel times of rays produced by shooting and those calculated by direct methods in the nonseparable environment of our example. Due to chaos, it can be difficult to calculate eigenrays between two *prespecified* widely spaced endpoints using the shooting method. In fact, in using the boundary conditions of our example above, we were unable to produce an eigenray by shooting with the available word length on our computer. Instead, we used the shooting method to propagate a ray to a specified range r_f, without concern for its final depth. We then used the final depth of this ray produced by shooting as the final boundary condition in calculating an eigenray using the Rayleigh–Ritz method. We first verified that the Rayleigh–Ritz method could reproduce the eigenray found by shooting. We then let the Rayleigh–Ritz method find an eigenray using an arbitrary trial solution. Figure 9 gives an example of a ray found by shooting, and an eigenray produced by the Rayleigh–Ritz algorithm using an arbitrary trial solution, connecting the same two points. In our example sound-speed profile, we invariably found that the eigenray produced by the Rayleigh–Ritz method using an arbitrary trial solution had an appreciably shorter travel time than the ray produced by shooting. Thus, direct methods tended to produce earliest arrivals. Furthermore, the eigenray produced by the Rayleigh–Ritz method was generally much less erratic than the ray found by shooting. Third, direct methods were observed to be stable with respect to the boundary conditions: small perturbations in the locations of the endpoints did not lead to great changes in the shapes of the eigenrays generated. On the other hand, direct methods were found to be quite sensitive to the trial solution chosen. Unless the trial solution was carefully chosen, the algorithm tended to converge to the earliest arriving ray.

As with shooting methods, the number of eigenrays that can theoretically be produced using direct methods grows exponentially with range. In practice it may be difficult to find eigenrays other than the fastest ones, due to the fact that arbitrarily chosen trial solutions tend to converge to the global minimum. On the other hand, for certain problems, finding even one eigenray using direct methods can often be helpful in finding more rays. An example of this can be seen in Fig. 10. There are three distinct eigenrays that follow a portion of the dashed path (A) for part of their journey between the endpoints, and the solid path for the remainder. They all have travel times that are equal, within roundoff error, to that of the solid ray. Each bifurcation is due to the fact that the ray paths are very unstable while riding on the parts parallel to the range axis. Excursions of the type shown are possible in any combination for each of the humps in the ray path, with immeasurably small differences in resultant travel time. In fact, due to the symmetry of the example duct, it is easy to construct a mesh of such rays, whose number grows exponentially with range. The dotted paths (B) in Fig. 10 show schematically the beginnings of such a mesh. All of the eigenrays in Fig. 10 have travel times that are equal, within roundoff error, to that of the solid eigenray.

The third question above, whether direct methods and shooting produce ''true'' rays when chaos is present, is the most difficult to answer, and is an area of ongoing research. To discuss shooting first, we note that the rays produced by shooting satisfy the Euler–Lagrange differential equations at each point to within the precision of the numerical method used. Because of the mirroring of the Euler–Lagrange equations by shooting methods, the rays produced by shooting methods in chaos-producing environments are extremely sensitive to initial conditions. By extension, since each point on a ray is an initial condition for subsequent points, shoot-

ing methods are extremely sensitive to numerical errors. Thus, even with perfectly given initial conditions, but with finite available numerical precision, rays produced by shooting will quickly diverge from the ''true'' ray with the same initial condition. The question then is whether and for how long a perturbed ray is near *any* ''true'' ray. This question introduces the subject of *shadowing*,[18,19] which is a subject of ongoing research and outside the scope of this paper. It is conjectured that for many classes of chaotic systems, for each numerically produced ray, there is a nearby initial condition that produces a ''true'' ray that is close to (or ''shadows'') the numerically calculated one. The shadowing conjecture has been proven for some restricted classes of chaotic systems.

We now discuss the third question above in relation to direct methods and show how our work with direct methods lends credence to the shadowing conjecture in chaotic ray tracing problems. We note that in practice, the ray paths produced by direct methods do not, in general, satisfy the Euler–Lagrange equations at each grid point. This is because the numerical optimization procedure is truncated before the limiting ray is actually achieved. On the other hand, when we have measured the amount by which rays numerically produced by direct methods fail to satisfy the Euler–Lagrange equations, we have found that the disagreement is small and can be made smaller by considering more coefficients or a more finely spaced grid. In addition, we have the observation of Collins and Kuperman[6] that simulated annealing can be used to facilitate the convergence of boundary-value methods, which numerically satisfy the Euler–Lagrange equations at the grid points. We also have our observation that for a given pair of boundary values and a given trial solution, direct methods tend to yield an approximate eigenray whose shape is relatively insensitive to small perturbations in those boundary conditions. Taken together, the above observations lend credence to the plausibility of the shadowing conjecture to ray tracing: somewhere in the neighborhood of the numerically calculated ray is a ''true'' ray.

Part of the key to understanding the success of direct travel time minimization in producing eigenrays may lie in the following observation: very little of the travel time of a ray comes from any local contribution. Therefore, it seems reasonable that a direct minimization method would be insensitive to perturbations in the boundary conditions of an eigenray. In fact, for grid-based direct methods (such as simulated annealing or Rayleigh–Ritz using ''hat'' functions), small perturbations at any single grid point should have little effect on the travel time. Only global perturbations, such as those introduced by varying a single coefficient of a sine basis function for the Rayleigh–Ritz method, should have a large effect on the travel time. For example, the initial angle of an eigenray calculated using the Rayleigh–Ritz method with sinusoidal basis functions varies greatly with the number of coefficients N. For $N>50$ in our example problem, the initial angle seemed to be approaching $-15°$, which is close to the launch angle for which the shooting method achieved the best match to the eigenray. However, the initial angle for $N=31$ was $-7.29°$, yet with 31 coefficients a good representation of the eigenray was obtained. The solutions to the ray equations generated by the shooting method vary tremendously for such a wide range of launch angles. The wide variation of the solution with launch angle is due to the fact that shooting methods mirror the dynamical behavior of the Euler–Lagrange equations, which are extremely sensitive to initial conditions in chaos-producing environments. On the other hand, direct methods were observed to be sensitive to the trial solution chosen, but to be relatively insensitive to errors in launch angles, and in the positions of the endpoints.

IV. SUMMARY AND CONCLUSIONS

We now summarize our experience using direct optimization methods in continuous media. First, given an arbitrary trial solution, direct methods always yielded at least one eigenray, even in chaos-producing conditions. Second, since direct methods strive to minimize travel time across the entire ray path at once, rays produced by direct methods were generally the earliest arriving rays when compared to those produced by shooting. Third, we have noted that direct methods are designed to find locally extremizing paths and, in theory, could miss saddle curves. We have not here investigated the nature of the extremal paths (i.e., local extremum or saddle curve) found by either method. In practice, however, we were able to reproduce using direct methods every ray we produced by shooting.

On the negative side, using known approaches for choosing trial solutions, direct methods could not be used to explore multipath structure when that structure was unknown. With an arbitrary trial solution, direct methods tended to converge to the fastest eigenrays. Reproducing other eigenrays usually required very good trial solutions, presupposing knowledge of the arrival structure. In fact, even when chaos was *not* present, direct methods required good trial solutions to find any but the earliest arrivals. When chaos is present, the multipath structure consists of rays whose number grows exponentially with range. Thus, it is clear that even with direct methods, ray chaos imposes fundamental limitations on the usefulness of ray concepts.

Comparing the direct methods to shooting, we can say the following: In practice, in sufficiently benign chaotic environments, shooting methods, used with sufficient computer word length and care, can be used to the find the entire multipath structure between two points that are not spaced too far apart.[4] The full multipath structure may include, for some types of problems, some eigenrays that cannot be found using shooting (e.g., the harmonic oscillator duct studied in this paper, and the head wave example developed in the companion paper[13]). Over longer ranges and with a fixed word length, it can quickly become impossible to find even one eigenray using shooting. In contrast, using direct methods, we were always able to find at least one eigenray, no matter what the range. In situations such as the example of Collins and Kuperman, where there is a good deal of structure or symmetry in the environment, direct methods were useful in finding some of the multipath structure. In order to fully exploit direct methods, however, it would be useful to develop a systematic strategy for choosing trial solutions to

map out an eigenray arrival structure in more general environments. Whereas the space of all shooting angles is one dimensional, the function space of possible paths when using direct methods is infinite dimensional. Thus, whether it is possible to develop a general strategy for generating good trial solutions for all the paths which produce local minima of the time delay integral is a question for further research.

In the companion to this paper,[13] we study ray propagation problems in discontinuous media that are amenable to analysis via discrete mappings. We examine some of the questions raised in the present paper without introducing the complicating and obscuring issues of sophisticated numerical analysis and infinite dimensionality. We also further explore a class of rays computable by direct methods that cannot be computed via shooting techniques.

ACKNOWLEDGMENTS

One of the authors (M. A. Mazur) would like to thank Dr. Diana F. McCammon of the Applied Research Laboratory at the Pennsylvania State University for her support and suggestions. This research was funded, in part, by the Office of Naval Research under the Fundamental Initiatives Program.

[1] D. R. Palmer, M. G. Brown, F. D. Tappert, and H. F. Bezdek, ''Classical chaos in nonseparable wave propagation problems,'' Geophys. Res. Lett. **15**, 569–572 (1988).

[2] K. B. Smith, M. G. Brown, and F. D. Tappert, ''Ray chaos in underwater acoustics,'' J. Acoust. Soc. Am. **91**, 1939–1949 (1992).

[3] M. G. Brown, F. D. Tappert, G. J. Goñi, and K. B. Smith, ''Chaos in underwater acoustics,'' in *Ocean Variability and Acoustic Propagation*, edited by J. Potter and A. Warner-Varnas (Kluwer Academic, Dordrecht, 1991), pp. 139–160.

[4] F. D. Tappert and X. Tang, ''Ray chaos and eigenrays,'' J. Acoust. Soc. Am. **99**, 185–195 (1996).

[5] W. A. Kuperman and M. D. Collins, ''Finding eigenrays by optimization with application to tomography and overcoming chaos,'' J. Acoust. Soc. Am. **93**, 2425(A) (1993).

[6] M. D. Collins and W. A. Kuperman, ''Overcoming ray chaos,'' J. Acoust. Soc. Am. **95**, 3167–3170 (1994).

[7] M. A. Mazur and K. E. Gilbert, ''Finding eigenrays in environments prone to numerical instability by directly optimizing the travel time integral,'' J. Acoust. Soc. Am. **96**, 3354(A) (1994).

[8] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1953), Vol. I, Chap. 4, pp. 164–274.

[9] I. M. Gelfand and S. V. Fomin, *Calculus of Variations* (Prentice-Hall, Englewood Cliffs, NJ, 1963), pp. 195–197.

[10] S. Kirkpatrick, C. D. Gellatt, and M. P. Vecchi, ''Optimization by simulated annealing,'' Science **220**, 671–680 (1983).

[11] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, ''Equations of state calculations by fast computing machines,'' J. Chem. Phys. **21**, 1087–1091 (1953).

[12] L. Davis and M. Steenstrup, ''Genetic Algorithms and Simulated Annealing: An Overview,'' in *Genetic Algorithms and Simulated Annealing*, edited by L. Davis (Morgan Kauffmann, San Mateo, CA, 1987), Chap. 1, pp. 1–11.

[13] M. A. Mazur and K. E. Gilbert, ''Direct optimization methods, ray propagation and chaos: II. Propagation with discrete transitions,'' J. Acoust. Soc. Am. **101**, 184–192 (1997).

[14] B. R. Julian and D. Gubbins, ''Three-dimensional seismic ray tracing,'' J. Geophys. **43**, 95–113 (1977).

[15] D. Cvijović and J. Klinowski, ''Taboo search: An approach to the multiple minima problem,'' Science **267**, 664–666 (1995).

[16] T. L. Foreman, ''An exact ray theoretical formulation of the Helmholtz equation,'' J. Acoust. Soc. Am. **86**, 234–246 (1989).

[17] A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (Acoustical Society of America, New York, 1989), pp. 376–378.

[18] E. Ott, *Chaos in Dynamical Systems* (Cambridge U.P., Cambridge, MA, 1993), p. 18.

[19] S. M. Hammel, J. A. Yorke, and C. Grebogi, ''Do numerical orbits of chaotic dynamical processes represent true orbits?'' J. Complexity **3**, 136–145 (1987).

Direct optimization methods, ray propagation, and chaos. II. Propagation with discrete transitions

Martin A. Mazur and Kenneth E. Gilbert
The Applied Research Laboratory and the Graduate Program in Acoustics, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

(Received 19 February 1996; accepted for publication 25 July 1996)

Eigenrays between a source and receiver are classically determined by an initial value or ''shooting'' approach. In range-dependent environments the ray paths can be chaotic, putting a fundamental limit on the accuracy of classical ray tracing methods. An alternative approach can be based on Fermat's principle of minimum propagation time. Using the so-called ''direct methods'' of the calculus of variations, the travel time integral can be minimized directly, rather than by means of the Euler–Lagrange differential equations used in shooting. In the current paper, examples are studied where the changes in a ray's trajectory occur at discrete points. Discrete mappings, analogous to the Euler–Lagrange equations used in shooting techniques in continuous problems, are introduced. It is demonstrated that direct methods can be applied to the travel time summations in discrete problems, just as they can be applied to the travel time integral in continuous problems. It is shown that direct methods can be used to calculate certain arrivals (e.g., ''head waves'') that cannot be produced by shooting. Using discrete examples, direct methods are compared to the more conventional discrete mapping (shooting) approach, without the complications of numerical analysis and infinite dimensionality found in continuous problems. Two examples are studied that are associated with a standard discrete mapping known to be chaotic. Direct methods are used to find eigenrays for these chaotic examples. The advantages and limitations of direct methods in discrete problems with chaos are discussed.
[S0001-4966(97)04412-3]

PACS numbers: 43.30.Cq [MBP]

INTRODUCTION

Classically, eigenrays between a source and receiver are determined by an initial value or ''shooting'' approach. The launch angles of rays from a source point are varied until the rays intersect the receiver endpoint. In the companion to this paper,[1] we investigated an alternative approach, suggested by the original work of Collins and Kuperman,[2,3] based on Fermat's principle of minimum propagation time. In the approach of Collins and Kuperman, the travel time integral is minimized using so-called ''direct methods'' of the calculus of variations,[4,5] rather than by solving the Euler–Lagrange differential equations. The Rayleigh–Ritz algorithm, a classical direct method, and simulated annealing,[6,7] a Monte Carlo approach, were considered. We applied these methods to examples in continuous media. In particular, direct methods were applied to an example of Collins and Kuperman for which the ray paths were known to be chaotic. In studying the example, we found that direct methods could find eigenrays when none could be found via shooting with the word length available to us to represent the initial conditions. When the travel times of arbitrary rays calculated by shooting were compared with those of eigenrays with the same endpoints found by direct methods, the rays found by direct methods were the earliest arrivals. On the other hand, in order to find rays other than the earliest arrivals using direct methods, good trial solutions were required. Hence, a great deal of prior knowledge about the arrival structure was needed, since the number of rays connecting two points is known to grow exponentially with range under chaotic conditions.[8,9]

In the present paper, we turn our attention to problems where the changes in a ray's trajectory occur at discrete points. A well-known example of such problems is that of finding the eigenrays connecting two points that lie over a planar interface separating two dissimilar fluids. If the lower fluid is a ''fast'' bottom, there is an eigenray associated with the so-called ''head wave'' that cannot be found by classical shooting techniques. However, such a ray can be found by directly minimizing the expression for the travel time of the ray.

In discrete problems with multiple points of transition, it is often possible to derive discrete mappings, which yield the ray path information between the discrete points by simple functional evaluation. Thus, there is no need for numerical integration or ordinary differential equation (ODE) solving algorithms as is the case with continuous problems. Discrete mappings are the analogs to the Euler–Lagrange differential equations used in the shooting methods for continuous problems. Another advantage of studying discrete problems is that, for them, the problem of minimizing the travel time summation is inherently finite dimensional. The discrete mappings used in shooting can be derived using principles from multivariate calculus. Alternatively, direct methods such as simulated annealing can be applied directly to the travel time summations.[10] Because of their conceptual and analytical simplicity, discrete problems can thus yield insight

into some environments where ray chaos is an issue. In the current paper, we study two examples that can be associated with the Taylor–Chirikov mapping,[11,12] a standard analytical problem known to be chaotic. We compare the shooting technique to direct methods for the chaotic examples.

By studying problems with discrete transitions, we have buttressed many of the observations and conclusions of our first paper, without introducing the potentially complicating and obscuring issues of numerical analysis and infinite dimensionality. Our experience has shown direct methods to have the following properties: (1) Direct methods can easily find certain eigenrays under chaotic conditions, whereas for shooting methods with a fixed accuracy in the initial angle, the ability to calculate eigenrays rapidly vanishes with range. (2) Direct methods are useful for finding the earliest arrivals, i.e., eigenrays with the global minimum travel time. (3) By varying the trial solution, direct methods can produce a multiplicity of eigenrays. We have been able to use direct methods to reproduce eigenrays produced by shooting. (4) Unlike shooting methods, direct methods are relatively insensitive to boundary conditions. (5) On the other hand, direct methods can show a high degree of sensitivity to the trial solution used in the optimization procedure. Under chaotic conditions, the number of eigenrays connecting two points increases exponentially with range in nonseparable environments. Direct methods can, in principle, find these exponentially many rays. However, unless the trial solution is carefully chosen, the optimization method tends to find a globally optimum solution (earliest arrival). Therefore, direct methods cannot, in practice, replicate a complex arrival structure because the required trial solutions presuppose knowledge of it. (6) Direct methods can be used to calculate certain rays that cannot be found by shooting techniques. These *diffracted rays*[13] may actually violate the Euler–Lagrange differential ray equations at particular points. Diffracted rays can occur in environments with discontinuities or singularities in the properties of the medium or in the boundaries. In the companion paper,[1] for instance, we gave an example where direct methods were used to find an eigenray that was the envelope of the refracted ray paths. In the present paper, we show that direct methods can yield rays at an interface between two media that are not products of the law of reflection, and hence cannot be found by shooting (e.g., see the ''head wave'' example below). In the abovementioned examples, the rays nonetheless obey Fermat's principle and are physically real.

In Sec. I we introduce the type of problem for which discrete mappings are useful, and we give a simple example. In Sec. II we derive the travel time summations for two somewhat more complex examples from acoustics and optics. We also derive discrete mappings for these examples from the travel time summations via Fermat's principal. Both of the examples exhibit chaotic ray propagation. In Sec. III we apply both discrete mappings (shooting) and direct methods to the chaotic examples. In Sec. IV we discuss the details of our findings and draw some conclusions regarding direct methods.

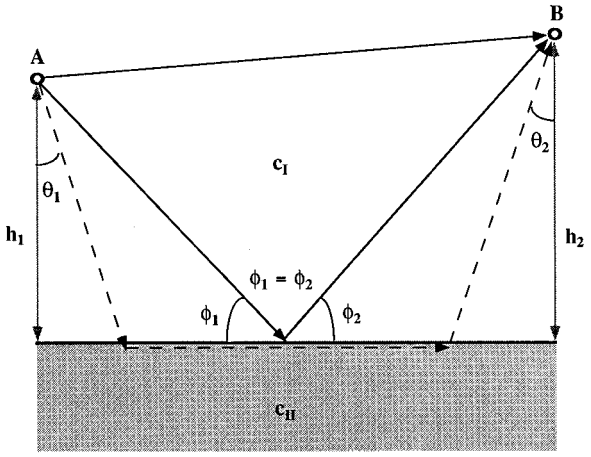


FIG. 1. A simple eigenray problem with discrete transitions: a source and receiver over a ''fast'' bottom. The solid rays can be found both by shooting and by directly minimizing travel time. The dashed ray, which is associated with the ''head wave'' solution of the wave equation, can be found by directly minimizing travel time, but not by shooting.

I. PROPAGATION WITH DISCRETE TRANSITIONS

In many simple ray propagation problems, a significant change in ray propagation occurs at discrete points in the ray path. In such problems, the use of a *discrete mapping* allows one to progress from one point of change to the next by simple functional evaluation. Discrete mappings can thus be employed to simplify the analysis by avoiding the numerical complications of solving a set of differential equations. On the other hand, discrete mappings, like differential equations for continuous systems, are subject to chaos in most range-dependent environments.

In problems with discrete transitions, as in continuous problems, we can also directly minimize the travel time summations. For example, techniques such as steepest descent and simulated annealing are minimization methods that can be employed.

A. A simple example

The type of problem shown in Fig. 1 is a simple example[13] of a problem with discrete transitions. In such problems, the changes in a ray path (or in the type of path a ray takes) occur at discrete points. The ray path between these points is known from general principles, and only a simple functional evaluation (in this case, the law of reflection) is necessary to determine the transition to the next path.

Consider a source A and receiver B located at heights h_1 and h_2, respectively, above a planar interface separating two dissimilar fluids, as shown in Fig. 1. Suppose further that the speed of sound in the upper medium is c_{I} and that in the lower is c_{II} where $c_{\mathrm{I}}<c_{\mathrm{II}}$. In an initial value, or ''shooting'' approach to determining the ray paths connecting A and B, we would immediately recognize the line-of-sight path. We would also be able to determine a reflected path by drawing a ray to the bottom, invoking the law of reflection and continuing the ray, and repeating this process until we had found the ray connecting the two endpoints.

There is a third possibility for a ray path, which cannot be found from classical ray theory. To find this ray, we em-

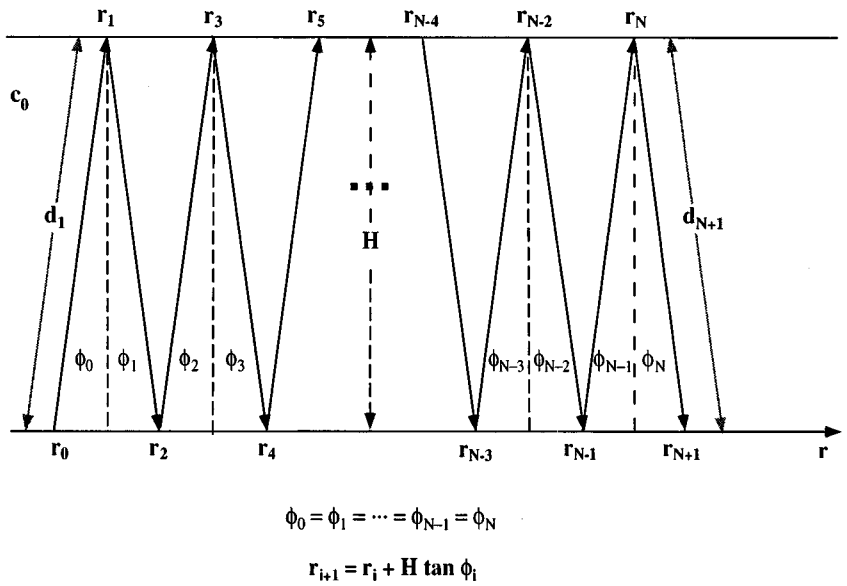


FIG. 2. Ray propagation in an isovelocity duct with perfectly reflecting boundaries.

ploy an approach that considers all combinations of allowable paths between the two endpoints, and selects those for which the travel time is locally minimum. Any path in medium I or medium II must travel in a straight line. Neglecting for a moment the law of reflection, we can posit a hypothetical path that leaves the source at an angle θ_1, intersects the bottom and proceeds in medium II along the interface between the two media, and emerges in medium I on a path making an angle θ_2 with the bottom to intersect the receiver. The travel time of such a path is given by

$$T_{\mathrm{AB}}=\frac{h_1}{c_{\mathrm{I}}\cos\,\theta_1}+\frac{r-h_1\,\tan\,\theta_1-h_2\,\tan\,\theta_2}{c_{\mathrm{II}}}+\frac{h_2}{c_{\mathrm{I}}\cos\,\theta_2}, \tag{1}$$

where r is the horizontal distance between A and B. If we require T_{AB} to be stationary with respect to variations in θ_1 and θ_2, we arrive at the conclusion that $\theta_1=\theta_2=\sin^{-1}(c_{\mathrm{I}}/c_{\mathrm{II}})$. Thus, the hypothetical ray enters and leaves medium II at the so-called *critical angle* θ_c determined by the ratio of the sound speeds of the two media. However, at the point where it reenters medium I, this hypothetical ray violates the law of reflection, as well as the Euler–Lagrange differential equations. However, that such a ray exists is confirmed by solution of the wave equation for this problem. The portion of the ray from the bottom to the receiver, called a *diffracted ray*,[13] is not explicable by classical ray theory, and is associated with the *head wave* solution of the wave equation. The reader may verify that the direct and reflected ray paths also minimize the travel time among all rays which connect A and B with zero and one bottom reflection, respectively. In the second case, the law of reflection is *derived from* Fermat's principle.

B. Derivation of discrete mappings

A slightly more complicated example can be given as follows: Consider an isovelocity channel of depth H bounded above and below by perfectly reflecting boundaries, as shown in Fig. 2. Suppose there is a source at (0,0). A ray leaving this source with an initial angle with respect to the vertical of ϕ_0 intersects the top boundary $z=0$ at $r_1=H\,\tan\,\phi_0$. Invoking the law of reflection, we see that $\phi_1=\phi_0$, and we can determine that the ray will intersect the bottom again at $r_2=r_1+H\,\tan\,\phi_1$. In this way, we develop an iterative discrete mapping for the coordinates (r_i,ϕ_i) of the ith intersection of the ray with the boundaries. For this simple example, the mapping is

$$\phi_{i+1}=\phi_i, \tag{2}$$

$$r_{i+1}=r_i+H\,\tan\,\phi_i, \tag{3}$$

with the initial condition $(0,\phi_0)$.

In an eigenray problem, we would wish to find the initial angles for which rays will intersect the bottom at $r=r_f$. In the shooting approach, we would employ the mapping above and vary the initial angle ϕ_0 until we found $r_{N+1}=r_f$ for some N. [In this simple example, we note that the mapping has a closed form solution: in practice we would only have to find all solutions ϕ_0 of $\tan\,\phi_0=(r_f-r_0)/(N+1)H$.]

We can also derive the mapping for the above problem from the travel time summation, without invoking the law of reflection. In fact, the law of reflection becomes a consequence of this derivation. We note that the length of the ray segment between boundary bounces is given by

$$d_i^2=H^2+(r_i-r_{i-1})^2, \tag{4}$$

and so the total travel time is

$$T_{\mathrm{tot}}=\sum_{i=1}^{N+1}\frac{d_i}{c_0}=\sum_{i=1}^{N+1}\frac{H}{c_0}\sqrt{1+\left(\frac{r_i-r_{i-1}}{H}\right)^2}, \tag{5}$$

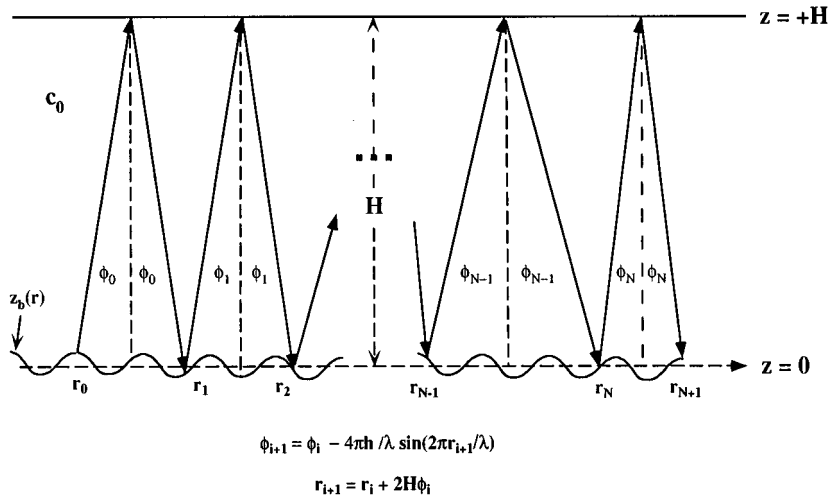


FIG. 3. An example of chaotic acoustic ray propagation: an isovelocity duct with perfectly reflecting boundaries. The upper boundary is $z=+H$. The lower boundary is given by $z_b(r)=-h\cos(2\pi r/\lambda)$.

where $r_{N+1}=r_f$. We can now derive equations that are the necessary conditions for the travel time to be minimum. The equations are analogous to the Euler–Lagrange differential equations used for continuous problems, and are derived by setting the gradient of T_{tot} with respect to the bounce ranges r_i to zero:

$$\frac{\partial T_{\text{tot}}}{\partial r_i}=\frac{1}{c_0}\left[\left(\frac{r_i-r_{i-1}}{d_i}\right)-\left(\frac{r_{i+1}-r_i}{d_{i+1}}\right)\right]=0,$$
$$i=1,\ldots,N+1.$$

Noting that

$$\sin\phi_i=\frac{r_{i+1}-r_i}{d_i},$$

we thus have that

$$\sin\phi_i=\sin\phi_{i-1}$$

or

$$\phi_i=\phi_{i-1}, \tag{6}$$

which is the law of reflection. The other half of the mapping,

$$r_i=r_{i-1}+H\tan\phi_{i-1} \tag{7}$$

is obtained by inspection. With the replacement of i with $i+1$ in Eqs. (6) and (7) we recover the mapping given by Eqs. (2) and (3). Hence, we see that the techniques of multivariate calculus can be applied to the travel time summation to find a discrete mapping. The law of reflection is a consequence of the mapping. The mapping can be used to find eigenrays via the shooting method.

We can also find eigenrays for the above problem by minimizing the travel time summation in Eq. (5) using direct methods such as simulated annealing. For this simple example, simulated annealing easily yields the same eigenrays as shooting. In the next section, we derive travel time summations for two somewhat more complex problems from acoustics and optics.

II. TWO EXAMPLES EXHIBITING CHAOS

A. Derivation of travel time summations and mappings

In most range-dependent environments, ray propagation is subject to chaos. Such environments are called *nonseparable*[14,15] because the equations for the ray paths cannot be solved by separation of variables. We now investigate two physical examples, which exhibit chaos: the first acoustical, the second optical. The optical example, although somewhat artificial, is easiest to visualize. These two examples both lead to the same discrete mapping, the so-called Taylor–Chirikov, or ''standard'' mapping, but each example yields its own insights. The Taylor–Chirikov mapping is a standard analytical mapping, known to exhibit chaos, that has been the subject of considerable study. See, for example, Refs. 11 and 12 for readable theoretical discussions, and Ref. 8 for another example from acoustics.

Consider the duct shown in Fig. 3 (a similar example can be found in Ref. 16). The top surface at $z=+H$ is flat and a perfect acoustic reflector. The bottom surface at $z=z_b(r)$ is again perfectly reflecting, but varies sinusoidally about $z=0$ with amplitude h and wavelength λ:

$$z_b(r)=-h\cos(2\pi r/\lambda). \tag{8}$$

Thus, the height of the duct is given by $H-z_b(r)$. We consider high grazing angle rays. With source and receiver on the bottom at ranges r_0 and $r_f=r_{N+1}$, respectively, the travel time taking $N+1$ surface bounces is given by

$$T_{\text{tot}}=\sum_{i=1}^{N+1}T_i, \tag{9}$$

where

$$T_i=\frac{1}{c_0}\{d_{i1}+\delta_{i1}+d_{i2}+\delta_{i2}\}, \tag{10}$$

and where the components d_{ij} and δ_{ij} are illustrated in Fig. 4. We have that

$$d_{i1}=d_{i2}=d_i=H\sec\phi_{i-1}$$

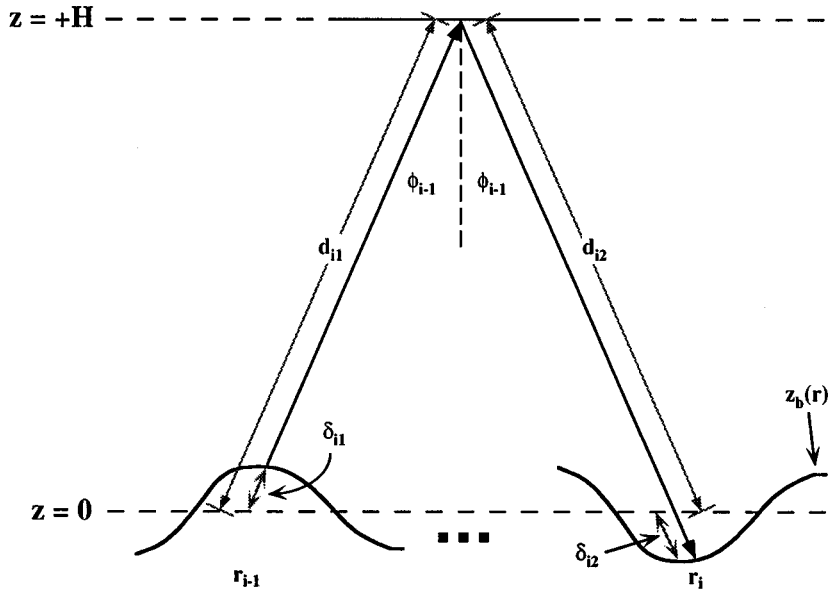


FIG. 4. Illustrative description of d_{ij} and δ_{ij}. The quantity d_{ij} measures the distance from the surface at $z=+H$ to $z=0$ along the direction of the ray. The distance from the bottom z_b to $z=0$ along the direction of the ray is given by δ_{ij}. Note that δ_{ij} is positive for $z_b<0$ and negative for $z_b>0$.

and

$$\delta_{i1}=z_b(r_{i-1})\sec\,\phi_{i-1}, \quad \delta_{i2}=z_b(r_i)\sec\,\phi_i.$$

For $\phi_i \ll 1$, we have

$$d_i \approx H\left(1+\frac{1}{2}\tan^2\,\phi_{i-1}\right)\approx H\left[1+\frac{1}{2}\left(\frac{r_i-r_{i-1}}{2H}\right)^2\right],$$

$$\delta_{i1}\approx z_b(r_{i-1}),$$

and

$$\delta_{i2}\approx z_b(r_i).$$

Using the approximate expressions for d_{i1}, δ_i and δ_{i2}, we have

$$T_i \approx \frac{2H}{c_0}\left[1+\frac{1}{2}\left(\frac{r_i-r_{i-1}}{2H}\right)^2\right]-\frac{1}{c_0}\left(z_b(r_{i-1})+z_b(r_i)\right)$$

so that

$$T_{\rm tot}\approx\sum_{i=1}^{N+1}\left\{\frac{2H}{c_0}\left[1+\frac{1}{2}\left(\frac{r_i-r_{i-1}}{2H}\right)^2\right]-\frac{2}{c_0}\,z_b(r_i)\right\} + \frac{z_b(r_{N+1})-z_b(r_0)}{c_0}. \tag{11}$$

We wish to minimize $T_{\rm tot}$ with respect to the bounce ranges r_i. To find a discrete mapping, as before, we set the gradient of $T_{\rm tot}$ with respect to the bounce ranges to zero. Thus, the necessary conditions for a (local) minimum become

$$\frac{\partial T_{\rm tot}}{\partial r_i}=0\approx\frac{1}{c_0}\left[\left(\frac{r_i-r_{i-1}}{2H}\right)-\left(\frac{r_{i+1}-r_i}{2H}\right)\right] - \frac{2}{c_0}\,z_b'(r_i), \quad i=1,\ldots,N+1, \tag{12}$$

where z_b' is the derivative of z_b with respect to r. Substituting the following identity

$$\frac{r_{i+1}-r_i}{2H}\approx\tan\,\phi_i\approx\phi_i \tag{13}$$

into Eq. (12), we have

$$\phi_i\approx\phi_{i-1}-2z_b'(r_i). \tag{14}$$

Rearranging Eq. (13) and substituting Eq. (8) for z_b, we have a two-dimensional discrete mapping for generating a succession of points (r_i,ϕ_i) given an initial range r_0 and launch angle complement ϕ_0:

$$r_{i+1}\approx r_i+2H\phi_i, \tag{15}$$

$$\phi_{i+1}\approx\phi_i-\frac{4\pi h}{\lambda}\sin(2\pi r_{i+1}/\lambda). \tag{16}$$

The mapping is known as the Taylor–Chirikov, or ''standard'' mapping, or the ''kicked undamped rotator'' mapping,[11,12] and it exhibits chaotic behavior. The mapping can occur in the context of other examples from acoustics, and its properties have been explored in considerable depth in Ref. 8. Note also that in Eq. (14) the law of reflection is a consequence of Fermat's principle.

Before discussing shooting and direct methods in the context of this example, we introduce another example, from optics, that again yields the mapping in Eqs. (15) and (16) above. We consider a second example for the mapping because the acoustical example does not yield immediate insight into the shapes of the ray paths. In the acoustical example, the mapping is in range-angle space with unequal bounce ranges, and consequently the ray path is difficult to visualize. In the optical example, in contrast, the ray path is in range-altitude space and can be immediately seen.

Consider a series of thin optical lenses, equally spaced in range r, and separated by spacing Δx, as shown in Fig. 5. Between the lenses, the speed of propagation is a constant,

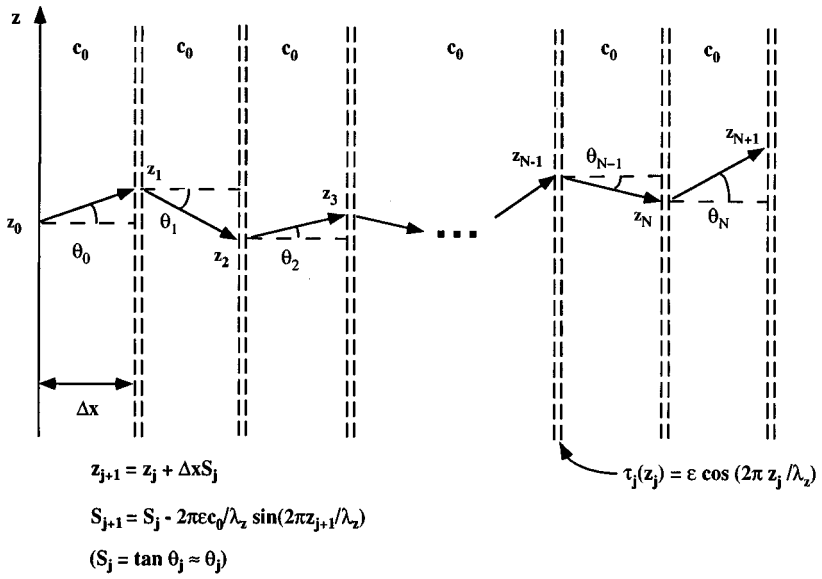


FIG. 5. An example of chaotic optical ray propagation: a series of lenses with associated propagation delays that vary with z. The space between the lenses is isovelocity.

c_0. The lenses are such that they can be described by an associated travel time delay τ, which varies with altitude z. If τ is chosen to be

$$\tau = \epsilon \cos(2\pi z/\lambda_z),$$

where ϵ and λ_z are, respectively, the amplitude and wavelength of the vertical variation, then, as shown below, we obtain the Taylor–Chirikov mapping. We consider rays propagating at low angles to the horizontal. Then the travel time of a ray originating at altitude z_0 and progressing to the $(N+1)$st lens is given by an expression very similar to Eq. (11):

$$T_{\text{tot}} \approx \sum_{j=1}^{N+1} \frac{\Delta x}{c_0} \left\{ 1 + \frac{1}{2} \left(\frac{z_j - z_{j-1}}{\Delta x} \right)^2 \right\} + \tau(z_j), \tag{17}$$

where z_j is the altitude at which the jth lens intercepts a ray. We have used the low angle or ''parabolic'' approximation to simplify the expression. We identify the local slope of a ray as

$$S_j = \frac{z_{j+1} - z_j}{\Delta x} = \tan\ \theta_j \approx \theta_j, \tag{18}$$

where θ_j is the propagation angle with respect to the horizontal, as shown in Fig. 5. By setting the gradient of T_{tot} with respect to the altitudes z_j equal to zero and using Eq. (18), we can derive the following mapping, which is the Taylor–Chirikov mapping:

$$z_{j+1} \approx z_j + \Delta x S_j, \tag{19}$$

$$S_{j+1} \approx S_j - \frac{2\pi\epsilon c_0}{\lambda_z} \sin(2\pi z_{j+1}/\lambda_z). \tag{20}$$

The equivalence of the mapping in Eqs. (19) and (20) to that in Eqs. (15) and (16) is easily seen. Equation (20) has the same form as the law of reflection [Eq. (16)] for the acoustical example, and can be looked on as a ''law of transmission'' for this optical example.

III. ITERATIVE MAPPINGS AND DIRECT MINIMIZATION APPLIED TO CHAOTIC EXAMPLES

In media with discrete transitions, as in continuous media, we can produce eigenrays either by applying shooting techniques or by using direct minimization methods. With shooting, an eigenray is found first by choosing an initial condition (e.g., altitude and launch angle in the chaotic optical example) and then using the iterative mappings derived for the problem to propagate the ray. The initial conditions are varied until an eigenray is obtained. In contrast, direct methods are applied by numerically minimizing the travel time summations. In either case, the calculations involved at each grid point are simple functional evaluations. Thus, with discrete transitions, the calculations are straightforward compared to those with continuous media, where numerical ODE solvers are required for shooting and numerical integrators are required for direct methods.

If we examine the simple, nonchaotic example given for the isovelocity channel (Sec. I B), we see by inspection that the mapping yields straight-line propagation in the direction given by the initial launch angle. The ray follows the law of reflection at the boundaries. It is not apparent by inspection, but is true, that direct methods such as simulated annealing will also yield straight-line propagation and obey the law of reflection for this example. Thus, when there are no discontinuities in the sound speed (such as would occur with a fast bottom), both shooting and direct methods yield the same rays.

In the acoustical example in Sec. II that yielded chaotic ray paths, we again specified that the bottom is a perfect acoustical reflector. Thus, because there are no discontinuities in the sound speed, such as would occur with a fast bottom, we expect rays to obey the law of reflection at both boundaries. Hence, under the conditions of straight-line propagation and the law of reflection, there is no mechanism for ''head waves'' and their associated rays. For our ex-

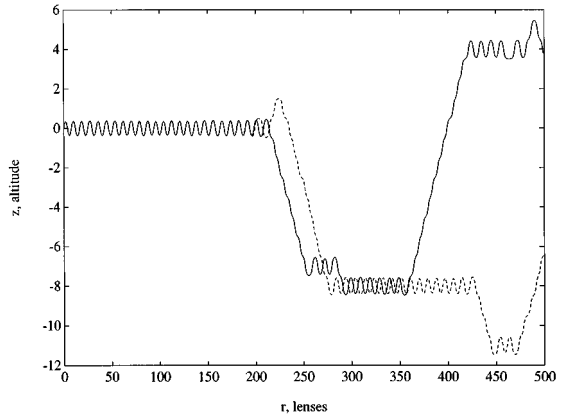


FIG. 6. Chaotic rays produced by the shooting method using the Taylor–Chirikov mapping. The rays were launched with initial position $r_0=0$, and initial angle $\theta_0=15.0°$ (solid ray) and $\theta_0=15.0°-10^{-9°}$ (dashed ray).

amples, we would thus expect there to be a one-to-one correspondence between rays that could be generated by shooting and those that could be found by direct methods. However, due to the exponential proliferation of rays with range, this is very difficult to verify exhaustively. Nevertheless, we can compare shooting and direct methods using examples. In our examination, we first rely on the optical example because the graph of altitude, z, versus lens number, i, is identical to the ray path. We will return to the acoustical example near the end of this section.

We have already noted that the mapping given by Eqs. (19) and (20) can be chaotic under certain conditions. An example of this chaotic behavior is shown in Fig. 6. Here we have $\epsilon=0.02$, $\lambda_z=1$, $\Delta x=1$, and $c_0=1.5$. Two rays were launched with initial conditions $r_0=0$ and $\theta_0=15.0°$ (solid ray) and $\theta_0=15.0°-10^{-9°}$ (dashed ray). The travel times for these two rays are 336.98 and 337.21 units, respectively. Both rays are very erratic and diverge from each other. Other rays with nearly identical initial conditions could also be plotted, and these would also be very erratic and divergent.

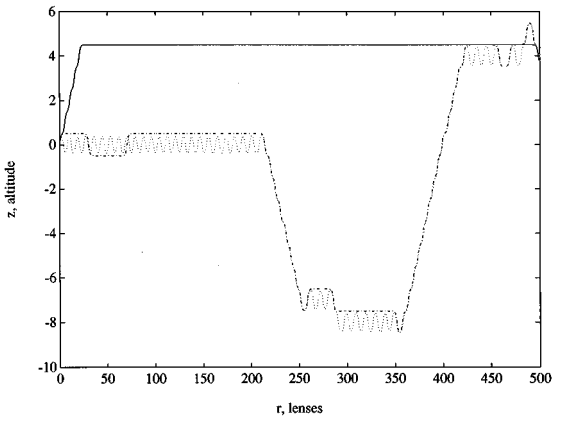


FIG. 7. Eigenrays produced by simulated annealing for the example in Fig. 6. The dotted eigenray reproduces, using simulated annealing, the solid ray in Fig. 6 that was calculated by shooting. A trial solution very close to the solution was needed to produce the dotted eigenray. The travel time is 336.98 units. The dot–dash eigenray was produced using a trial solution somewhat perturbed from the dotted eigenray. The travel time is 327.51 units. The solid eigenray is the ''fast'' eigenray. Most arbitrarily chosen trial solutions led to the ''fast'' eigenray. The travel time is 323.45 units.

Figure 7 shows some results when a direct method, simulated annealing, is applied to the problem investigated by shooting in Fig. 6. In Fig. 7 we use the same physical parameters, but treat the problem as a boundary value problem. The initial point is $r_0=0$, $z_0=0$, and the final point was chosen, for comparison to the solid ray in Fig. 6, to be $r_f=500$, $z_f=3.8066$, which is where the solid ray in Fig. 6 ended up. We note first that the ray calculated by shooting (solid ray in Fig. 6) was reproduced using direct methods (dotted eigenray in Fig. 7). However, we note that for this ray, the direct method is extremely sensitive to the trial solution chosen: if we use the ray derived by the shooting method as a trial solution to simulated annealing, or a trial solution only slightly perturbed from this, we get back the original ray. Otherwise, we get a different ray, usually one with a shorter travel time. Figure 7 also shows two other eigenrays produced by simulated annealing with the same boundary conditions, but somewhat different trial solutions. The dot–dash eigenray that looks somewhat like an ''envelope'' of the ray derived by shooting was obtained by choosing a boundary condition that was considerably perturbed from the original ray. The travel time for the dot–dash ray was 327.51 units. Finally, the ray path with the least variation (solid eigenray in Fig. 7) had a travel time of 323.45 units. This was the ''fastest'' eigenray we were able to obtain. Most randomly chosen trial solutions led to this eigenray. Thus, without some prior knowledge of the arrival structure, it is difficult to find any but the fastest rays, unless, as we show later, some symmetry in the problem can be exploited.

It is worthwhile to ask if the converse is true: can shooting methods reproduce the ''fast'' rays obtained by direct methods? As we saw in the ''fast bottom'' example, if the environment has discontinuities in the propagation speed, reproducing the ''fast'' rays by shooting may be impossible. In the chaotic examples given here, there is no discontinuity that would cause head waves. However, due to the extreme sensitivity of shooting to initial conditions, it is very difficult to reproduce the ''fast'' rays via shooting. In the optical example with a small number of lenses, we were able to use shooting to reproduce the fast ray with some difficulty. However, the accuracy required to reproduce the ''fast'' ray grew exponentially with the number of lenses, and after ≈ 50 lenses was impossible to achieve with the word length available to us to represent initial conditions.

Figure 7 also draws our attention to another feature of direct methods. Although direct methods generally produce only the ''fastest'' rays given an arbitrary trial solution, in some cases direct methods can give some insight into the arrival structure, and make clear how more eigenrays may be constructed. The dot–dash eigenray that appears somewhat like an ''envelope'' of the ray calculated by shooting is actually one of a family of such rays. Due to the symmetry and structure of the problem, rays that travel along the levels $z_n=\pm n/2$, $n=1,3,5,...$ contribute least to the travel time. The fastest eigenrays between two points will tend to spend much of their time near the levels z_n. They do not actually ride exactly on these levels since, due to the ''law of transmis-

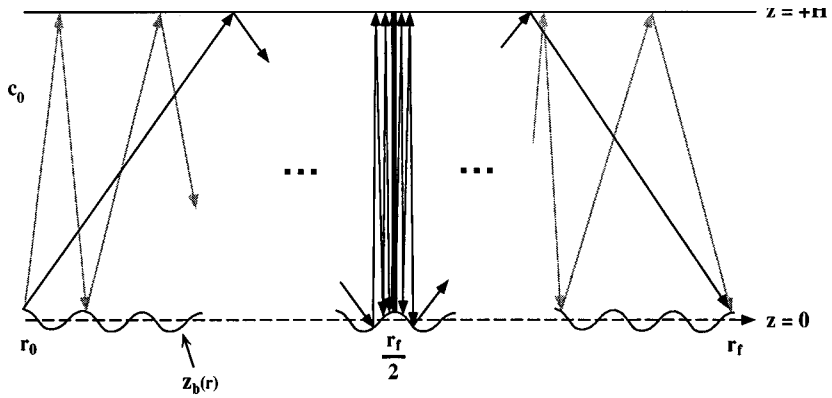


FIG. 8. A schematic representation of an arbitrary eigenray (shaded) and a ''fast'' eigenray (dark) for the chaotic acoustical example. Both eigenrays have the same number of bounces between r_0 and r_f. The ''fast'' ray is the ray with the shortest travel time for a given number of bounces. The thick black line at $r_f/2$ schematically represents a large number of nearly vertical bounces packed closely together. The tightly grouped bounces are consistent with the law of reflection and minimize the travel time per bounce. The ''fast'' eigenray takes the least number of bounces consistent with the law of reflection to go from r_0 to $r_f/2$, and to go from $r_f/2$ to r_f. These two features of ''fast'' eigenrays, namely, hurrying to a state where the least travel time is expended, and staying there as long as possible, were also seen in the optical example (solid eigenray in Fig. 7).

sion'' [Eq. (20)], once a ray is at one of the levels z_n, it is never reflected away. However, due to the instability in the problem, the slightest departure from the paths that stay near these levels will cause the ray to meander. Thus, between any two points, there is a family of eigenrays that meander about on a mesh consisting of the levels z_n. The height of the mesh increases with the distance between the endpoints, and an exercise in counting shows that the number of possible eigenrays in the family increases exponentially with range. The set of eigenrays on the mesh of levels z_n is only a small subset of the total set of chaotic ray paths between the two endpoints. In fact, in our experience, arbitrary rays calculated using shooting spend the *least* amount of their travel time at the ''fast'' levels z_n, since these rays are not calculated with a goal of reducing travel time to a minimum. In problems lacking the symmetry of the present example, direct methods may not provide the same degree of insight into the arrival structure.

Are the ''fast'' rays of any consequence? In some problems, ''fast'' rays are associated with head waves, and cannot be reproduced using shooting methods. In ocean environments, head waves are usually weak arrivals. However, ''fast'' rays such as those associated with head waves are not always weak arrivals, and in chaotic problems it can be impossible, with finite word length, to find the fastest arrivals using shooting. Consider our acoustical example in Fig. 3. Since we allowed no losses associated with boundary contact, and no evanescent propagation, all loss was due to geometrical spreading. Thus, for a given number of bounces, the strongest rays connecting two points were the fastest rays.

Again returning to our chaotic acoustical example (Fig. 3), we show another way of looking at the difference between arbitrary eigenrays calculated by shooting, and ''fast'' eigenrays found by direct methods. Figure 8 shows schematically two eigenrays. The shaded ray is an arbitrary eigenray calculated by shooting. The ray meanders seemingly randomly as it traverses the duct. The dark ray is the ''fast'' eigenray *with the same number of bounces as the shaded ray*, calculated by a direct method. Note that most of the bounces are spent at the halfway point in the duct, bouncing from the surface to near one of the peaks of the sinusoidal bottom. Obviously, such bounces are the least time costly of all possible bounces. The ray takes the widest angle path consistent with the law of reflection to reach this state. These two features of ''fast'' eigenrays produced by direct methods, namely, hurrying to a state where the least travel time is expended and staying there as long as possible, were also seen in the optical example (solid eigenray in Fig. 7).

IV. SUMMARY AND CONCLUSIONS

In this section, we first note some similarities between the results derived from discrete mappings and the observations we made for continuous media. We then draw our conclusions.

In problems in continuous media and in problems with discrete transitions, there were advantages and drawbacks associated with direct methods. First, on the positive side, even in chaotic problems, direct methods always yielded at least one eigenray between two points. With shooting methods, it becomes more difficult to find eigenrays with increasing range, and for a fixed word length eventually becomes impossible. Second, direct methods yielded the earliest arriving rays, even in chaos-producing conditions, since the methods strive to minimize travel time across the entire ray path at once. Third, in problems containing environmental discontinuities that would lead to diffractive propagation, direct methods yielded ''nonclassical rays'' (such as diffracted rays) that could not be calculated using shooting methods.

On the negative side, given arbitrarily chosen trial solutions, the most likely solutions were the ''fast'' eigenrays. Direct methods generally *did not* reproduce a ray produced by shooting *unless* a trial solution very close to this ray was employed. Direct methods therefore could not be used to explore multipath structure when that structure was unknown. This was true even in cases where chaos was not present. When chaos *is* present, the multipath structure can include geometric and diffracted rays whose numbers grow

exponentially with range. Thus, it is clear that even with direct methods, ray chaos imposes fundamental limitations on the usefulness of ray concepts.

We have left a number of questions not only unanswered, but in some cases unexamined. The most important of these questions are connected with fully characterizing the behavior of direct methods with respect to the type of problem. For example, are there any chaotic problems where direct methods fail to find eigenrays? What is the relationship between the location in *phase space* (q.v., Refs. 8, 14, and 15) of a chaotic eigenray and the sensitivity of direct methods to the trial solution used to find that ray? Examination of these more detailed questions is a subject for further research.

ACKNOWLEDGMENTS

One of the authors (M. A. Mazur) would like to thank Dr. Diana F. McCammon of the Applied Research Laboratory at the Pennsylvania State University for her support and suggestions. This research was funded, in part, by the Office of Naval Research under the Fundamental Initiatives Program.

[1] M. A. Mazur and K. E. Gilbert, ''Direct optimization methods, ray propagation and chaos. I. Continuous media,'' J. Acoust. Soc. Am. **101**, 174–183 (1997).
[2] W. A. Kuperman and M. D. Collins, ''Finding eigenrays by optimization with application to tomography and overcoming chaos,'' J. Acoust. Soc. Am. **93**, 2425(A) (1993).
[3] M. D. Collins and W. A. Kuperman, ''Overcoming ray chaos,'' J. Acoust. Soc. Am. **95**, 3167–3170 (1994).
[4] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1953), Vol. I, Chap. 4, pp. 164–274.
[5] I. M. Gelfand and S. V. Fomin, *Calculus of Variations* (Prentice-Hall, Englewood Cliffs, NJ, 1963), pp. 195–197.
[6] S. Kirkpatrick, C. D. Gellatt, and M. P. Vecchi, ''Optimization by simulated annealing,'' Science **220**, 671–680 (1983).
[7] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, ''Equations of state calculations by fast computing machines,'' J. Chem. Phys. **21**, 1087–1091 (1953).
[8] M. G. Brown, F. D. Tappert, G. J. Goñi, and K. B. Smith, ''Chaos in Underwater Acoustics,'' in *Ocean Variability and Acoustic Propagation*, edited by J. Potter and A. Warner-Varnas (Kluwer Academic, Dordrecht, 1991), pp. 139–160.
[9] F. D. Tappert and X. Tang, ''Ray chaos and eigenrays,'' J. Acoust. Soc. Am. **99**, 185–195 (1996).
[10] M. A. Mazur and K. E. Gilbert, ''Finding eigenrays in environments prone to numerical instability by directly optimizing the travel time integral,'' J. Acoust. Soc. Am. **96**, 3354(A) (1994).
[11] H. G. Schuster, *Deterministic Chaos: An Introduction* (Weinheim, Germany, 1989), 2nd rev. ed., p. 19.
[12] M. Tabor, *Chaos and Integrability in Nonlinear Dynamics: An Introduction*, (Wiley, New York, 1989), pp. 134–135.
[13] A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (Acoustical Society of America, New York, 1989), pp. 376–378.
[14] D. R. Palmer, M. G. Brown, F. D. Tappert, and H. F. Bezdek, ''Classical chaos in nonseparable wave propagation problems,'' Geophys. Res. Lett. **15**, 569–572 (1988).
[15] K. B. Smith, M. G. Brown, and F. D. Tappert, ''Ray chaos in underwater acoustics,'' J. Acoust. Soc. Am. **91**, 1939–1949 (1992).
[16] S. S. Abdullaev, *Chaos and Dynamics of Rays in Waveguide Media*, edited by G. M. Zaslavsky (Gordon and Breach, Langhorne, PA, 1993), pp. 192–200.

Seafloor acoustic backscattering from different geological provinces in the Atlantic Natural Laboratory

Robert J. Greaves
Massachusetts Institute of Technology/Woods Hole Oceanographic Institute, Joint Program, Woods Hole, Massachusetts 02543

Ralph A. Stephen
Woods Hole Oceanographic Institution, Clark Laboratory, Woods Hole, Massachusetts 02543

(Received 13 October 1995; revised 30 May 1996; accepted 21 August 1996)

The characteristics of acoustic signals backscattered from inside-corner and outside-corner oceanic crust are investigated using acoustic reverberation data from the 1993 Acoustic Reverberation Special Research Program (ARSRP) Acoustics Cruise. Specifically, the seafloor dip distribution, as measured from Hydrosweep bathymetry data is compared in areas of each crustal type and a correlation between seafloor dip and seafloor scattering strength is sought. Beamformed and match-filtered acoustic data from the site A (run 1) monostatic, wideband, LFM (linear frequency modulated) experiment are used to find the scattering strength corresponding to specific areas of the seafloor. Scattering strength is determined as the average of intersecting beams from different source locations in order to reduce the left–right ambiguity inherent in beamsteering of a linear array. When overlaid on the bathymetry map, high-scattering strengths are found to correspond to steep flanks of seafloor features and can be used to determine their shape and orientation. Some of these feature shapes are characteristic of specific crustal regions. Cross-plotting scattering strength with true grazing angle or seafloor dip shows a trend increasing at a rate of about 0.1 dB/deg. This trend is not, however, sufficiently constrained to be a useful predictor. It is concluded that the seafloor dip, on the scale of a few hundreds of meters, that can be resolved with Hydrosweep bathymetry data, influences but does not determine, scattering strength. Since, in detail, the observed variations in scattering strength are larger than can be explained by data error or seafloor dip, it is suggested that other characteristics of steeply dipping areas, such as subsurface properties or smaller scale surface features, strongly affect the level of backscattered signals. [S0001-4966(97)05112-6]

PACS numbers: 43.30.Gv, 43.30.Hw, 43.30.Vh [JHM]

INTRODUCTION

In this study we compare acoustic reverberation signals backscattered from inside-corner and outside corner oceanic crustal regions, which border the trace of a second-order ridge discontinuity on the western flank of the mid-Atlantic ridge (MAR). Inside corner (IC) and outside corner (OC) refer to the place of origin of oceanic crust relative to the intersection of mid-ocean ridges (MOR) and ridge discontinuities, shown schematically in Fig. 1. MOR's with slow spreading rates (1–5 cm/yr), such as the MAR, characteristically produce IC and OC crust with distinct, and in some cases, dramatic, differences in seafloor structure and composition.[1–3] In particular, large scale features of OC crust produced at slow spreading ridges tend to be strongly lineated and low relief while IC features tend to be blocky and have higher relief.

The data we use are from the Acoustic Reverberation Special Research Program (ARSRP), 1993 Acoustics Cruise which was funded by the Office of Naval Research. During this cruise, a variety of acoustic reverberation experiments were conducted over several sites on the western flank of the MAR.[4] All of these sites lie along the trace of a second-order discontinuity and within a region on the west flank of the MAR referred to as the Atlantic Natural Laboratory.[5] We analyze beamformed and match-filtered monostatic reverberation data acquired during the site A (run 1) experiment. Site A and the locations from which these data were recorded are shown in Fig. 2. We chose, for comparison, the three seafloor areas labeled as, site A (OC crust), Syrian Knob (IC crust), and Casa Grande (IC crust).

A hypothesis of the ARSRP research community is that backscattered signals from IC crust are characteristically different than those from OC crust. Specifically, it has been suggested that OC crust will produce less monostatic, low-angle backscattering than IC crust. This hypothesis is primarily based on the ideas that, (1) for a fixed incident wave field grazing angle, the scattering strength increases with increasing seafloor dip (i.e., increasing true grazing angle) and (2) that IC crust has more steeply dipping features than OC crust. The hypothesis presumes that other seafloor properties, such as the thickness of sediment cover or subseafloor rock properties, have less effect, than seafloor morphology, on backscattered signals.

In our analysis, we look for a quantitative correlation between seafloor scattering strength, determined from low-grazing angle backscattered signals, and the measured seafloor dip, as a method for distinguishing IC and OC crustal areas. A preliminary step is to find the correlation between scattering strength and true grazing angle. True grazing angle

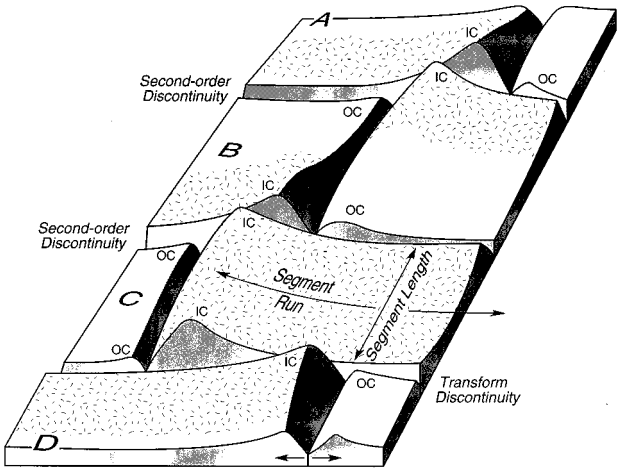


FIG. 1. Inside-corner (IC) crust and outside-corner (OC) crust are found on opposite sides of mid-ocean ridge (MOR) segments. In this figure there are four segments denote as A, B, C, and D. When the displacement between segment ends is greater than 30 km the discontinuity is usually the site of an active transform fault and is called a transform discontinuity. When the displacement is less than 30 km, it is referred to as a second-order discontinuity and may or may not be the site of an active transfrom fault. In either case, IC is used to describe the crust initiated between segment ends, i.e., adjacent to the active displacement zone of the discontinuity. OC crust is then the portion of the crust initiated adjacent to the passive trace of the discontinuity, and is coupled to the IC crust across the discontinuity soon after being formed in the ridge axial valley. This figure represents the crust formed at a slow spreading center, where IC crust is, usually, significantly higher than OC crust. (Figure from Tucholke and Lin.[1])

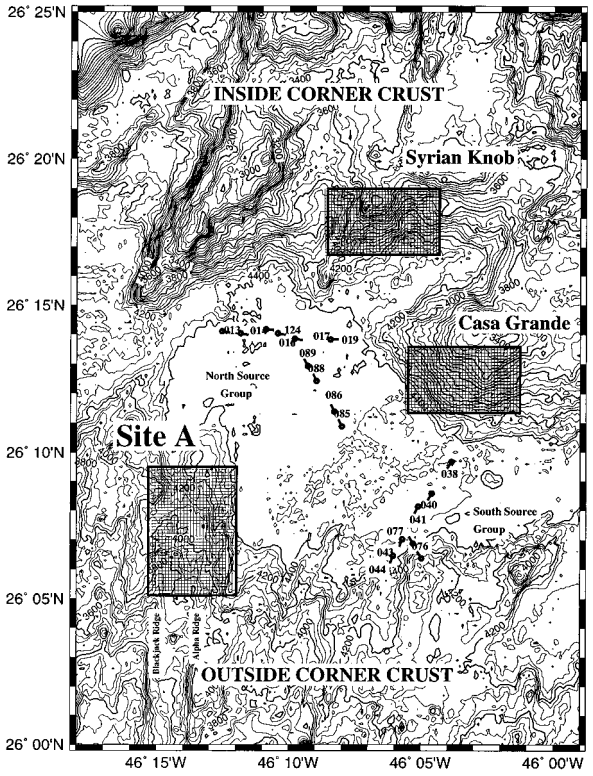


FIG. 2. Site A is located on OC crust on the edge of a sediment pond which marks the trace of a second-order discontinuity. Backscattered signal from two areas located on IC crust, labeled as Syrian Knob and Casa Grande, are compared to signals from site A. Source locations in two regions around site A, the north and south source groups, are analyzed separately. Within each study area, the ranges and depths to 200×200 m grid cells are used to find the corresponding acoustic backscatter signals. Circles with flags indicate the receiver array location and orientation at each source/receiver location.

is defined as the angle between the incident wave field propagation vector (incident at the grazing angle) and the plane defining the local seafloor dip. We use Hydrosweep bathymetry data, which has a nominal resolution of 200 m,[6] to define seafloor dip. Several studies have shown that there is a general correspondence between strong backscattered signals and large seafloor structural features. Makris *et al.*[7] have done detailed analysis of low-angle backscatter data at a site further to the west and found very good correlation of strong backscatter signals to seafloor ridges in the 1/2 CZ (convergence zone: a full CZ is the distance at which an acoustic wave field, refracted and turned at depth by the ocean acoustic velocity gradient, is returned to the sea surface) and 1 1/2 CZ ranges but did not determine an explicit relationship between intensity and seafloor dip. Preston *et al.*[8] found correlations between backscattered signals and large high-standing seafloor features for CZ's as distant as 400 km from the source and receiver arrays. Stephen *et al.*[9] mapped individual beam (the acoustic time series computed for a particular direction by beamsteering the linear receiver array) data sets onto one of our study areas (site A) and showed a similar correspondence between steep dip and high-intensity backscatter. Shaw *et al.*[10] found, using the same acoustic data, but using finer scale bathymetry data, that high-intensity backscatter mapped to areas with the strongest bathymetric gradient. However, they used data from single source locations and did not account for the left–right ambiguity of beamsteered monostatic reverberation data. In our analysis, we combine signals backscattered from each seafloor location from different source locations, before relating them to seafloor dip. This combination of intersecting beams acts to reduce directional ambiguity. Similar techniques have been shown to be effective in reducing left–right ambiguity in the analysis of monostatic reverberation data by Makris *et al.*[11] and Preston *et al.*[8]

In Fig. 3 we summarize some of the possible quantitative differences in scattering that could be observed and related to crustal regions. If we take the incident beam grazing angle to be constant, then changes in true grazing angle are due to changes in seafloor dip. If scattering strength increases with increasing dip, and IC and OC crust have distinct distributions of seafloor dip, then we might expect a result similar to Fig. 3(a). If the two crustal areas have similar seafloor dip distributions but have different scattering strengths we could see results similar to either Fig. 3(b) or (c). If the dip distribution and scattering strength are both similar, then we are not likely to be able to distinguish the crustal types on this basis, as suggested by Fig. 3(d). Other characteristics, such as seafloor feature shape, determined by simply mapping acoustic backscattering, are qualitative but can still be useful in identifying different crustal areas.

I. GEOLOGICAL SETTING

Site A lies on outside-corner (OC) crust formed about nine million years ago at a second-order discontinuity on the

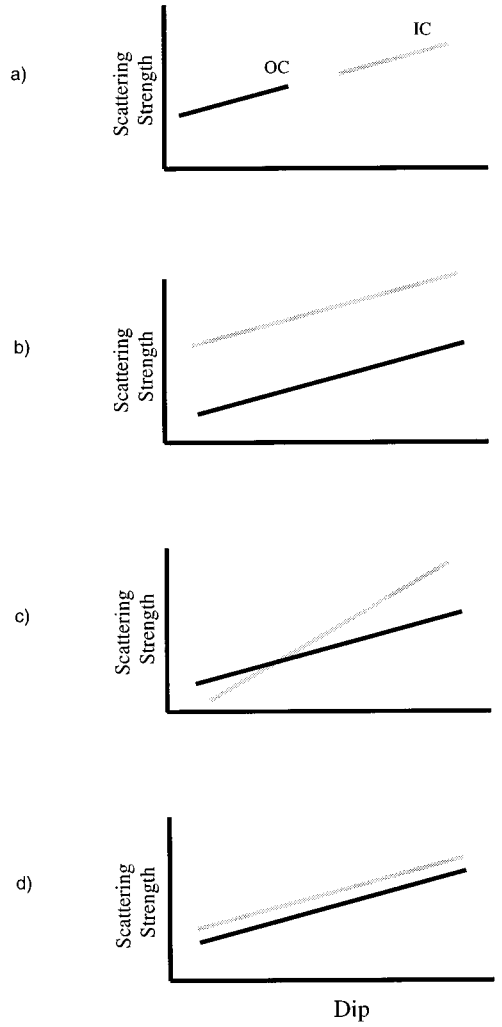


FIG. 3. If IC and OC crust have distinct populations of seafloor dip it may be possible to distinguish them on the basis of acoustic scattering strength, as in (a). If they have similar distributions of seafloor dip but have different scattering strength due to surface characteristics, then it may be possible to distinguish the crustal regions by differences in scattering strength levels (b) or in the shape or slope of a functional relationship (c). If both dip distribution and scattering strength are similar it may not be possible to distinguish the regions on any of these basis (d).

MAR about 400 km north of the Kane fracture zone.[12] Discontinuities occur frequently along mid-ocean ridges, dividing them into short ridge segments. Displacement of ridge segments occurs between segment ends, i.e., along the tectonically active portion of a discontinuity. The trace of each discontinuity is marked by a bathymetric low extending into the ridge flanks, the axis-of-maximum depth (AMD). In most cases, these are the deepest features along the ridge system and form the large sediment ''ponds.'' In general, offset displacement at a discontinuity of more than about 30 km is caused by the strike-slip motion of a transform fault and is referred to as a transform discontinuity. Smaller offsets are indicative of active displacement due either to a transform fault or to crustal deformation, and are referred to as second-order[1] or in some cases, zero-offset discontinuities.[13] Inside-corner (IC) crust refers to the crust found on the side of the ridge where the transform or discontinuity is active (i.e., between the ridge segment ends) (see Fig. 1). Crust on the opposite side of the ridge, where the discontinuity is passive, is called OC crust.

IC crust formed at slow spreading MOR's, is initially significantly higher than the adjacent OC crust. IC and OC crust formed at second-order discontinuities have elevation differences equal to, or greater than, those observed at transform discontinuities.[2,3] As young crust moves away from the spreading center it subsides as thermal cooling of the lithosphere takes place and uplifting forces are reduced. Along transform discontinuities, there may be little or no elevation difference by the time IC and OC crust become coupled in the fracture zone. Contrary to this, IC crust formed at second-order discontinuities spends relatively little time in the active displacement zone, so that much less subsidence of the original high elevation occurs before it becomes coupled to the OC crust. Once the crust moves away from the MOR axis, subsidence lowers both IC and OC regions but some difference in elevation between them remains locked in place as the plate moves far from the spreading center. This is important in our consideration of seafloor dip as a distinguishable characterstic of crustal type. If we lift a corner of two equal areas of seafloor, but lift one (IC) higher, we can predict that on average, the seafloor dip of the area lifted higher will be larger. For this reason, the OC crust around site A, which is located about 100 km west of the MAR, is expected to have lower elevations and, on average smaller seafloor dip than the adjacent IC crust.

A strong morphological distinction between IC and OC crust, is that OC crust tends to have highly lineated large scale features, while IC crust shows a more blocky structure.[3] The linear features on OC crust, are dominated by long ridges that persist for tens of kilometers (the half-length of the ridge segment) with strike parallel to the MOR. Over large areas, repetition of the ridges give OC crust a corrugated appearance. Where they intersect the AMD, they slope downward and end as distinctive ''toes'' along the edge of the sediment pond (e.g., Alpha Ridge and Blackjack Ridge in Fig. 2). These ridges are also asymmetric with the steep side facing the MOR. In general, ridges on OC crust are not single faults, but instead are made up of groups of normal faults formed in colinear sequences as blocks of new crust rise from the rift valley. As the blocks rise they are tilted back (away from the spreading center) in response to tectonic extension. Individual faults are quite steep, (70°–90°), but with short scarp heights, on the order of tens of meters.[1] Even though these individual faults are very steep, when averaged over hundreds of meters, the dip of the seafloor is much less, on the order of 20°–45°.

Long lineations are not characteristic of IC crust. Tucholke and Lin[1] observe that IC's have one grouping of very steep (70°–90°) and large displacement ($>$150 m) normal faults that face parallel to the rift valley but another grouping of somewhat less steep (20°–60°) faults facing either the rift valley or the ridge-discontinuity intersection. The combination of faulting leads to the blocky structure of IC crust. The region north of site A and the sediment pond, is dominated by large irregular blocks characteristic of IC crust.

Another important difference between IC and OC crust

is interpreted from descriptions of rock samples recovered from the seafloor. Tucholke and Lin[1] summarized rock sample data collected along the MAR. These data suggest that OC crust is almost uniformly covered by volcanic rock (basalt). In contrast, in IC areas, plutonic and ultramafic rocks, in particular gabbro and partially serpentinized peridotites, are frequently recovered in addition to volcanic rocks. They conclude that IC's have anomalously thin oceanic crust and that, to a large extent, the upper (volcanic) layer of normal oceanic crust is missing. Differences in scattering strength from IC and OC crustal regions that could not be associated with large scale seafloor morphology might be related to variations in basaltic versus gabbroic surface and volume properties.

Sediment thickness has a strong effect on the backscattered field. This has been shown clearly in model studies by Robertsson and Levander[14] and Stephen and Dougherty[15] where it is shown that increasing sediment thickness, decreases the backscattered signal. We know of no evidence suggesting that IC and OC have characteristically different sediment thickness. However, we might reason that more steeply dipping areas will hold less sediment on the steeply sloping regions and more in local bathymetric lows. If this is true then sediment distribution will act to enhance the relative scattering from more steeply dipping areas such as IC crust.

II. DATA

A. Acoustic data

The acoustic reverberation data used in this analysis were collected during the site A (run 1) monostatic experiment on board the R/V CORY CHOUEST. The ARSRP monostatic experiments only approximated the monostatic experiment geometry, since the source and receiver were separated by a horizontal distance of 1.174 km (Ref. 4, p. 106) and the research vessel from which they were deployed was underway ($\approx$3 knots) during the data acquisition. In the processing description, we show that accounting for the difference in source and receiver positions is important in analyzing 1/2 CZ data, but the correction for the moving ship is small and is neglected.

The acoustic source used for these experiments is a vertical line array (VLA) of ten piezoelectric flextensional acoustic projectors spaced 2.29 m apart. The depth to the center of the VLA for the site A experiment was 184 m. Time delays were applied to the ten sound projectors to effectively steer the central beam of the source radiation pattern downward at an angle of 9° from the horizontal. The receiver array consisted of 128 hydrophone groups spaced every 2.5 m in a horizontal line array (HLA). Two of the hydrophones were desensitized for source monitoring. The receiver array was towed at depths that varied from 130 to 170 m with the center of the array approximately 1 km from the stern of the ship (Ref. 4, pp. 278–284). The average depth of the receiver array was 155 m.

To generate a source pulse, the VLA is driven with an LFM (linear frequency modulated) 5-s sweep over the frequency band from 200 to 255 Hz. After the backscattered signals are received by the HLA, they are beamformed and match filtered on the ship. The beamforming process generates 126 directional beams each associated with a direction (actually two because of the beamforming left–right ambiguity) relative to the original receiver array axis. The 0° and 180° beams (end fire) are in-line with the receiver array axis and the 90° beam is the broadside beam. Angular beamwidth varies with beam direction, increasing toward the end-fire directions. Beams within 30° of the end-fire beams (0°–30° and 150°–180°) were not included in our analysis since they have much larger beamwidths and are strongly affected by array movement and ship noise.

Time series data recorded at each source/receiver location were reported, after shipboard processing, in units of the square of the pressure amplitude (μPa^2) observed at two-way travel times. These data are corrected in our processing for transmission loss and source beam pattern before they are treated as measurements of seafloor scattering strength. An important acquisition parameter is the consistency of the source strength. The calibrated source level at the time of the site A (run 1) experiment is specified as 230 dB (*re*:1 μPa–1 m) (Ref. 4, p. 217). Data from the source monitor hydrophone showed that the source level was very stable with a maximum observed variation of 0.4 dB from the mean but an average deviation of only $\pm$0.14 dB.

The HLA has a nominal broadside beamwidth of 1.1° at 250 Hz.[4] This is an important number since it is used to describe the azimuthal resolution of the beamforming system. It has been suggested[16] that the actual broadside beamwidth may be as large as 2°–3° due to a few inoperative hydrophones and bending of the HLA.[17]

The range (time) resolution of the beam data is estimated from the width of the central peak of the autocorrelation of the source function. The width of this peak is inversely proportional to the bandwidth of the source. For the bandwidth of these data (200–225 Hz), the range resolution is estimated to be 13.6 m (Ref. 4, p. 145).

B. Bathymetry data

The bathymetry data used in the analysis are from a Hydrosweep multibeam survey conducted as part of the 1992 ARSRP Geology and Geophysics Reconnaissance Cruise.[18] We use the bathymetry data as a grid of 200-m spaced samples interpolated from the original swath data.

Kleinrock[6] describes the resolution of the Hydrosweep multibeam bathymetry system as approximately 134$\times$134 m in the region nearest the nadir and 187$\times$256 m in regions reached by the outermost beams, for a water depth of 3500 m. Depth resolution is considerably better and is estimated to be about 20 m. Lateral variation in seafloor features on the scale of 200 m can be detected but practical feature resolution is on the order of 500 m. Grindlay *et al.*[19] also review high-resolution multibeam systems and conclude that multibeam data (such as Hydrosweep) cannot resolve individual seafloor features less than about 200 m and slopes of steep flanks that are greater than 45°. Tucholke and Lin[1] suggest that steep slope ($>$30°) areas, determined by multibeam data, are indicative of average slopes over hundreds of meters, which may include small scale features with steeper

dip. Based on these studies, we consider the nominal resolution of the Hydrosweep data to be 200 m, and treat each bathymetry value as the average depth value within a 200-m-diam circle on the seafloor.

III. DATA PROCESSING

The objective of the data processing is to transform the ''raw'' acoustic data into seafloor (interface) scattering strength. We also calculate the dip of the corresponding seafloor location and the true grazing angle of the incident acoustic wave field.

An initial review of the scattered signal levels from site A showed a strong dependence on source azimuth relative to the strike of large scale features. Therefore, we chose to separate the data into two groups having similar source azimuth relative to site A. One group is to the northeast of site A (see Fig. 2) and we call this the north source group. The other area is to the east of site A and we call this the south source group.

We make some approximations in our treatment of the backscatter and bathymetry data. First, an approximation is made to account for transmission loss and spatial beam pattern. The exact local beam ''footprint'' intensity variation is not accounted for, i.e., we assume that within each beam footprint, the seafloor is uniformly insonified. Second, by stacking intersecting beam signals we reduce the left–right ambiguity by enhancing repeated scattering. Stacking gives up some sensitivity of signal level to seafloor dip and removes details of signal variability as a function of source azimuth. Third, we do not account for the possibility of multiple scattering between patches of seafloor. Fourth, we approximate the seafloor dip within a particular beam footprint by a single seafloor dip value extracted from the Hydrosweep data. We also limit analysis to scattering from offset ranges between 10 and 25 km and depths from 3200 to 4500 m and, we do not use any beams within 30° of endfire.

The backscattering of cw (monofrequency continuous wave) acoustic signals from a nearly planar surface is described by the active sonar equation,[20]

$$\mathrm{RL}_s = \mathrm{SL} - \mathrm{TL} + \mathrm{TS}, \tag{1}$$

where RL_s is the observed signal level; SL is the initial source level; TL is the two-way transmission loss during propagation from source to target and target to receiver and includes the effect of the source beam pattern; and TS is the equivalent target strength of the scattering interface (seafloor). For broad bandwidth data, Eq. (1) is descriptively correct, but in application requires integration over the frequency band. If phase information is retained in the integration, we obtain the ''coherent'' broadband result; if phase information is ignored, we obtain the ''incoherent'' broadband result (Ref. 21, pp. 158–161).

In this study, the ''raw'' data are the match filtered beamformed time series which correspond to RL_s. Although there are no actual ''targets'' on the seafloor, seafloor scattering appears to be discrete (i.e., some seafloor features backscatter more than others). The equivalent ''target strength,'' TS, of specific patches of seafloor is defined as

$$\mathrm{TS} = S_s + 10\ \log_{10} A, \tag{2}$$

where S_s is the seafloor scattering strength and the approximate beam footprint area, A, is given by

$$A = r \Delta r \Delta \theta, \tag{3}$$

where r is the range offset, Δr is the range resolution and $\Delta \theta$ is the azimuthal resolution of the beam.

A. Transmission loss

We make two corrections for transmission loss. The first is a correction for the source spatial beam pattern and geometric spreading. The second, discussed further in a later section, is the removal of signals that appear to originate from areas shadowed from the incident beams by local bathymetric features.

For calculating the seafloor scattering strength (S_s) from the observed backscattered signal level (RL_s) we use the approximation

$$S_s \approx \mathrm{RL}_s - \mathrm{SL} + \mathrm{TL}_{\mathrm{cyl}}^{\mathrm{src}} + \mathrm{TL}_{\mathrm{cyl}}^{\mathrm{rec}} + \mathrm{TL}_0 - 10\ \log_{10} A, \tag{4}$$

where $\mathrm{TL}_{\mathrm{cyl}}^{\mathrm{src}}$ and $\mathrm{TL}_{\mathrm{cyl}}^{\mathrm{rec}}$ are the transmission losses predicted by cylindrical spreading relative to the source and receiver, and TL_0 is a static shift that, when combined with $\mathrm{TL}_{\mathrm{cyl}}^{\mathrm{src}}$ and $\mathrm{TL}_{\mathrm{cyl}}^{\mathrm{rec}}$ accounts for geometric spreading and the source spatial beam pattern. The derivation of this equation is given in the Appendix.

An estimate of the error in scattering strength values calculated using (4) is made by considering the error associated with each term on the right side of the equation. The background noise level of RL_s is estimated to be more than -10 dB down relative to the average calculated S_s and is negligible. We estimate this background noise level as the signal associated with water column scattering only, i.e., no seafloor scattering. As previously stated, the source level, SL, was observed to have a variation of only ± 0.14 dB. The difference between our approximation and the exact geometric spreading loss combined with the source spatial beam pattern is about ± 0.5 dB for the majority of the offset ranges used in our analysis (see the Appendix). We consider a net error of ± 1 dB to be a realistic estimate of the error in the calculation from the first five terms on the right of Eq. (4).

The potential error due to the area correction term is more problematic. For calculating the area in Eq. (3) we use a beamwidth of 1.1° and range resolution of 13.6 m. At 15 km the area correction ($-10 \log_{10} A$) is then -36 dB. However, if the beamwidth is 3°, the correction is -40 dB. The beamwidth is a function of the receiver array geometry and if this is continuously changing during the experiment then there could be as much as 4 dB of variation in S_s. If the array is stable but misshapen then this reduces to a static shift in the data. We have assumed that the receiver array was stable during the data collection since there is no evidence in the acquisition logs[4] suggesting that it was not. We assume this to be the case, so that the only ''error'' in the area term is an unknown static shift which, in our analysis, does not affect the conclusions. Therefore, we estimate that

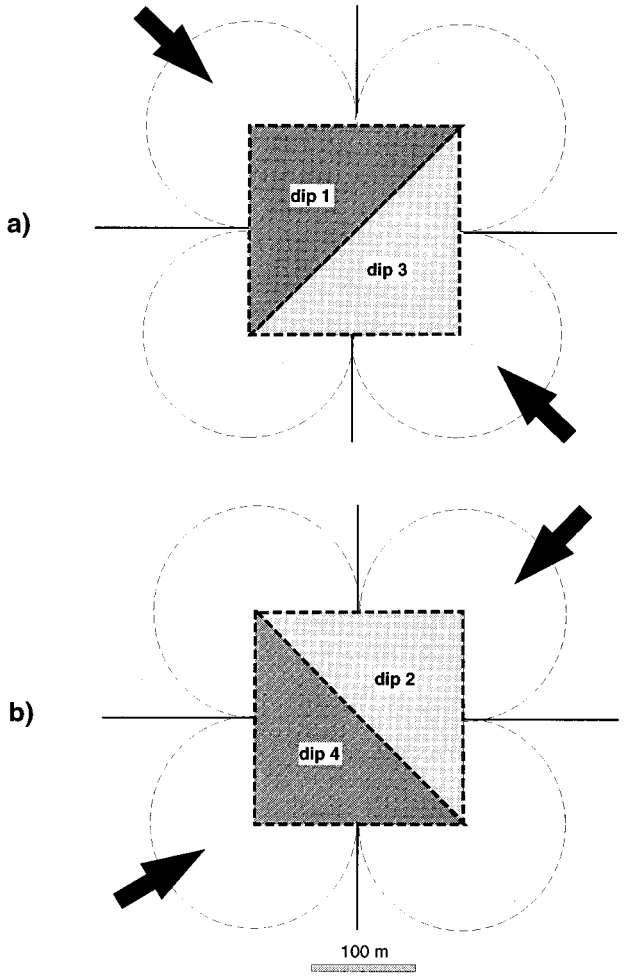


FIG. 4. Bathymetry values at each corner of the grid cell are considered to be average values within 200-m-diam circles (dotted circles) on the seafloor. Four triangles defined by the corners of the grid cell are used to define four directional seafloor dips. For each incident acoustic beam, two of the dips defined in this way are assigned to the beam. Which dips are assigned depends on incident beam direction. (a) Beams incident from the NW or SE are assigned dip 1 and dip 3 and (b) beams incident from the NE or SW are assigned dip 2 and dip 4.

the total error in the variation of the calculated scattering strength (S_s) is just the ± 1 dB due to the first five terms on the right of Eq. (4).

B. Seafloor locations and dip estimation

Given the nominal 200-m resolution of the Hydrosweep bathymetry data, we divided each area to be studied into 200×200-m grid cells (Fig. 2) and use each possible set of three cell corners to determine four dip planes, as shown in Fig. 4. Each incident beam is associated with a pair of dip planes as defined by two of the triangles, depending on the azimuth of the beam relative to the grid cell. For example, beams incident from the NE or SW quadrants are associated with the dip and dip azimuth determined for dip planes that fit the bathymetry values at the corners of triangle ''dip2'' and ''dip4.'' Similarly, beams incident from the SE or NW are associated with dip and dip azimuth calculated using values at the corners of triangles ''dip1'' and ''dip3.'' By defining dip in this way we are approximating the directional derivative of the bathymetry while explicitly utilizing every bathymetry value. In Fig. 5 we compare approximate beam footprints at 10-, 20-, and 30-km offset range to the area covered by 200-m-diam. circles at the corners of a grid cell. Within the 1/2 CZ, we estimate that the beam footprint area is on the order of 13 m $\times$ 200 m (range resolution $\times$ beamwidth) at 10 km, increasing to 13 m$\times$600 m at 30 km. Since three bathymetry values are required to define the dip and dip azimuth, the net area associated with each estimate of dip is quite large and fully includes the beam footprint up to 25 km in offset, even for beams intersecting the grid cell 100 m from the grid cell center. If beamwidth is on the order of $2°-3°$, this estimate of dip remains a crude, but reasonable, measure of the dip encountered by the acoustic beam footprint. We note here, that in final processing, we average data from seven consecutive (in range) beam footprints that fit within the 100-m range of each triangle.

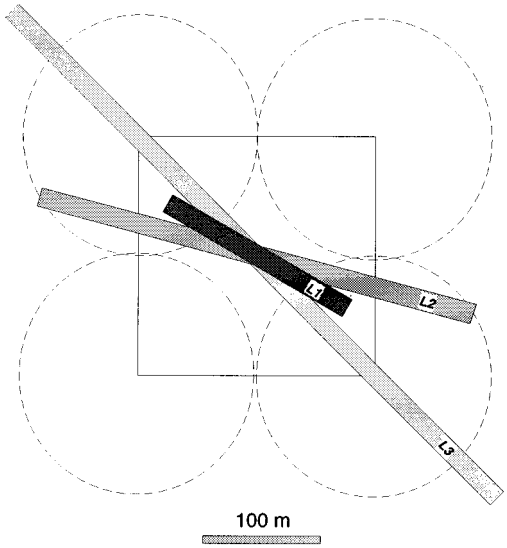


FIG. 5. Approximate beam footprints, for a beamwidth of 1.1°, at 10 km (L1), 20 km (L2), and 30 km (L3), in comparison to the area covered in the seafloor dip calculations (Fig. 4). For the ranges and depths considered in this study, the width of the beam footprint is the same scale as the dip resolution of Hydrosweep bathymetry data. In the analysis, seven consecutive beam footprints (covering a range of about 100 m) are associated with each dip value.

Figure 6 shows a histogram of seafloor dip as a percentage of total number of dip values calculated separately for each of the three target areas. The mean dip for the IC areas, Casa Grande and Syrian Knob, is a few degrees higher than for the OC area, site A. This agrees with the hypothesis that IC crust has characteristically steeper dip than OC crust, but is not as large a difference as expected. Also, it is clear from this plot that there is not a distinguishable difference in dip distribution that might have yielded the type of backscattering distinction suggested in Fig. 3(a). It must also be remembered that the dip distribution determined from Hydrosweep is skewed in that very steep dips (greater than about 45°) are not observed.

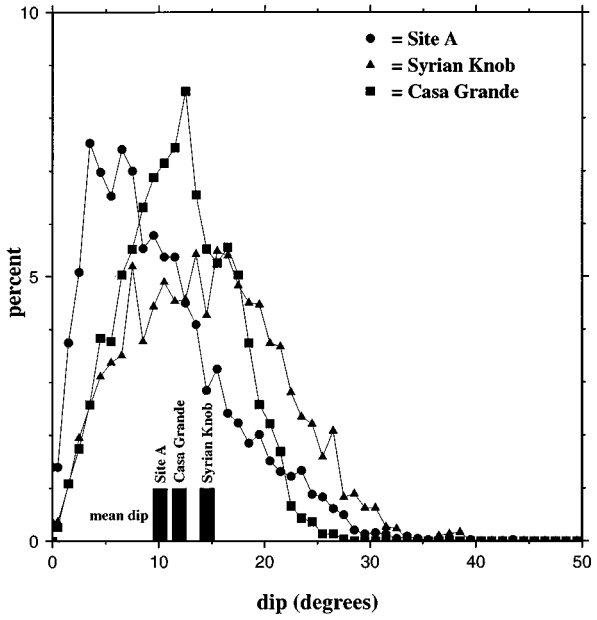


FIG. 6. Comparison of observed dips calculated from Hydrosweep data in each study area. The vertical axis is the percentage of calculated dips within each area that fall within each 1° interval. Note that the mean dip of site A, on OC crust, is smaller than the mean dip of both IC areas. Also, note that dip values greater than about 45° are not well resolved in Hydrosweep data.

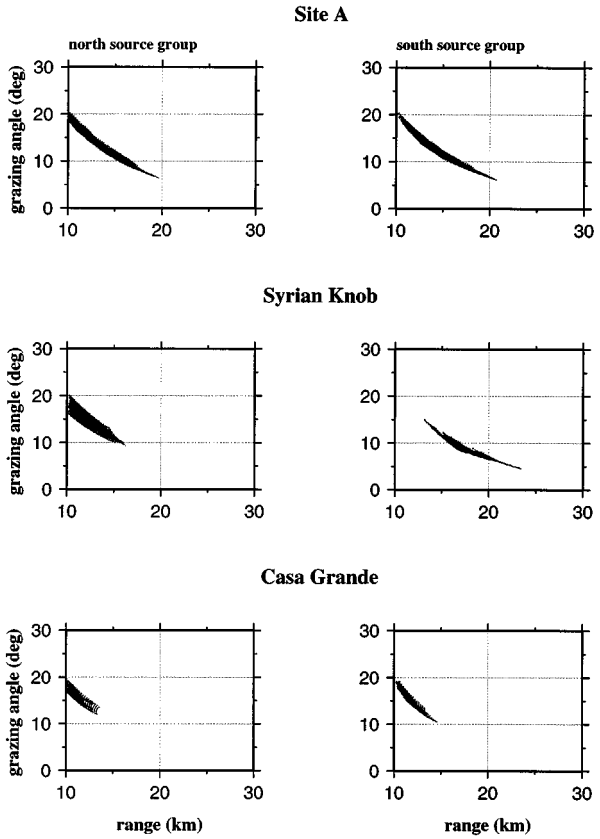


FIG. 7. The distribution of grazing angles versus range for each target area. These grazing angles are calculated using ray tracing in a layered ocean velocity model and assume a flat seafloor. The velocity profile used to generate this model is shown in the Appendix. The breadth of the curves indicate the range of observed depths. Note also, this plot shows that almost all of the horizontal offsets (range) actually used in our analysis lie within 10–20 km.

C. Source and receiver range

The fact that the source and receiver arrays are not exactly co-located is included in determining range (horizontal offset) and two-way travel time to each grid cell. Errors in travel time estimation can be as large as several hundred milliseconds if only the source, receiver, or their mid-point is used in calculating travel time. We use the ship location recorded on the acquisition logs[4] as the source location. From this location and the receiver array and ship headings, we calculate the position of the center of the receiver array. The locations shown in Fig. 2 correspond to the calculated receiver array locations.

Since the ship continues to make way during the time period between source firing and the arrival of the backscattered signals, there is a correction that could be applied to the receiver array position. This correction depends on ship speed, total signal travel time, and the beam direction. For example, the ship speed during the site A experiment was maintained at about 3 kn $\approx$ 1.5 m/s. The largest correction necessary would be for end fire beams, but in our study we exclude all beams within 30° of end fire. In our worst case, beams at 30°, the possible error in range at 10 km is about 17 m and at 20 km about 34 m. For broadside beams, near 90°, there is essentially no error. Since shifting beam footprint locations by such small amounts would not change the grid locations with which they are associated with, in most cases, we chose to neglect this correction.

D. Travel time and grazing angle

A Snell's law ray tracing routine was used to calculate tables of travel time and beam grazing angle for depth and range pairs. For this calculation we used the water velocity profile shown in Fig. A1. Calculated grazing angle is the angle between the ray (representing the beam or propagation vector) and a horizontal plane at each depth. The distribution of calculated grazing angles for each of the three study areas as a function of range is shown in Fig. 7. The nonuniqueness of grazing angles for each offset is due to the variation in observed depths. Separate travel time tables were generated for the source array at 184 m depth and the receiver array at 155 m depth. The offset of each source and receiver to the center of each grid cell and the depth of the grid cell are used to find, in the tables, the two-way (ray path) travel times to the center of each cell and the calculated beam grazing angle. Time windows of 133 ms (horizontal two-way travel time across a grid cell triangle) on either side of this center time are used to assign beam time series data from each source to the corresponding dip triangle.

At this point we also calculate true grazing angle, which is the angle between the incident beam propagation vector and the plane defining the local seafloor dip. This is calculated as the inverse cosine of the scalar product of the unit vector defined by the grazing angle and azimuth of the incident beams at each grid cell and the dip normal vector of each dip triangle. When the dip normal vector is in the same vertical plane as the beam propagation vector (i.e., the seafloor is ''looking'' directly back at or away from the source

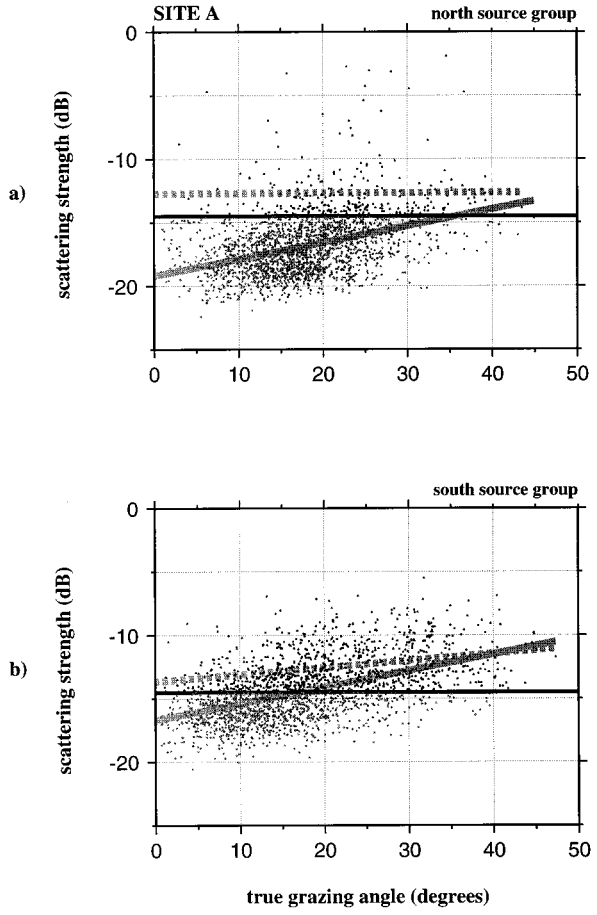


FIG. 8. Scattering strength of stacked intersecting acoustic beams, in site A, as a function of true grazing angle. The solid grey line is the linear regression fit, which for (a) the north source group data has slope of 0.13 ± 0.01 dB/deg with a correlation coefficient $r = 0.37$, and for (b) the south source group data is 0.13 ± 0.01 dB/deg and $r = 0.5$. Error of individual data values is estimated at ± 1 dB in calculated scattering strength and $\pm 2°$ in grazing angle. The standard deviation of the scattering strength when averaged over $1°$ intervals is about ± 2.5 dB (not shown) which suggests that there is variation on the order of ± 1.5 dB due to variations in seafloor properties other than true grazing angle. The dashed gray lines are the linear regression fits to the data that lies above a threshold (thin black line) chosen at -14.5 dB. In (a) the slope of this fit is 0.002 ± 0.02 dB/deg with $r = 0.01$ and in (b) the slope is 0.06 ± 0.01 dB/deg with $r = 0.3$. Note that thresholding the data in this way does not increase the slope of the trend in either case, nor does it improve the correlation coefficient.

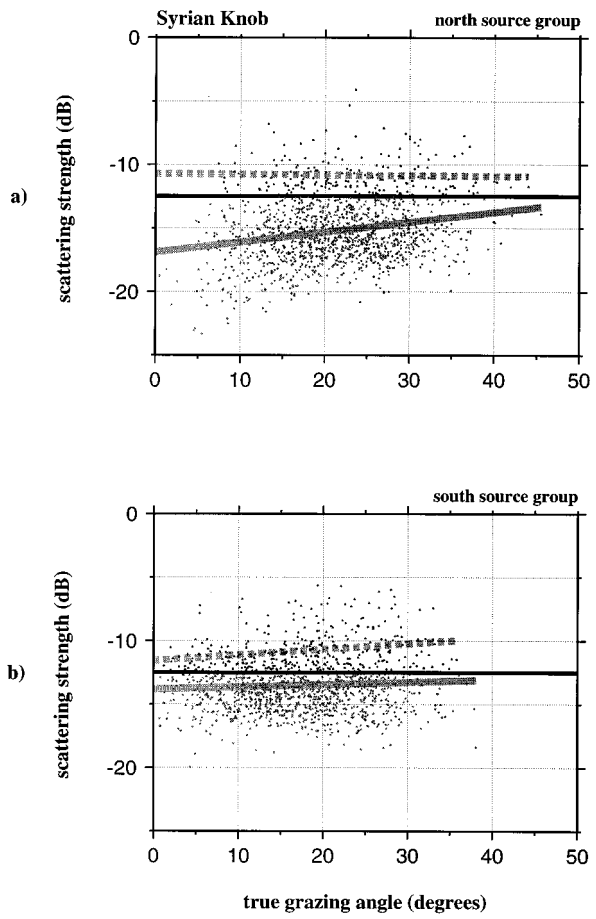


FIG. 9. Like Fig. 8 but for Syrian Knob. For the regression lines, in (a) slope $= 0.08 \pm 0.01$ dB/deg and $r = 0.22$ and in (b) slope $= 0.02 \pm 0.01$ dB/deg and $r = 0.06$. For signals above the threshold at -12.5 dB in (a) slope $= -0.005 \pm 0.01$ dB/deg and $r = -0.03$ and in (b) slope $= 0.05 \pm 0.01$ dB/deg with $r = 0.23$.

location), the true grazing angle is just the sum of the beam grazing angle and seafloor dip (or difference if the seafloor dip faces away from the source direction). For those cases where the seafloor dips away from the incident beam at an angle greater than the beam grazing angle, we treat the area as a shadow zone, and exclude the associated signal, corresponding to that dip triangle, from any further analysis. This is the second type of transmission loss factor that we account for.

E. Beam averaging and stacking

The next step in our processing is to stack (equal weight average) the time series data from intersecting beams, from the different source locations, at each dip triangle. We do all of the stacking and averaging after the signals have been converted to scattering strength in dB. We found that prestack scattering strength (in dB) are normally distributed, whereas the data in units of squared pressure are skew normally distributed. Averaging the normally distributed signals produces a better estimate of the true signal mean. Once the data are stacked we find the time average scattering strength over the time window (133 ms) of each triangle. Since in stacking we have averaged beams with different source azimuths, we also average the true grazing angles. We find that the range of true grazing angles at any particular grid cell is not large, since the variation in range and azimuth of individual sources within either source group is limited. The standard deviation of the mean true grazing angle for each location is about $\pm 2°$.

Figures 8–10 are cross plots of scattering strength at each dip triangle as a function of the average true grazing angle. This is plotted separately for north and south source groups. We also find the linear regression fit to the data, as well as the linear fit to the data when all values below some threshold level are eliminated.

In Figs. 11–13 we plot the data after we have constrained the stack such that only those beams with source azimuth and dip azimuth within $10°$ of being parallel are included, i.e., beams with source azimuth almost normal to

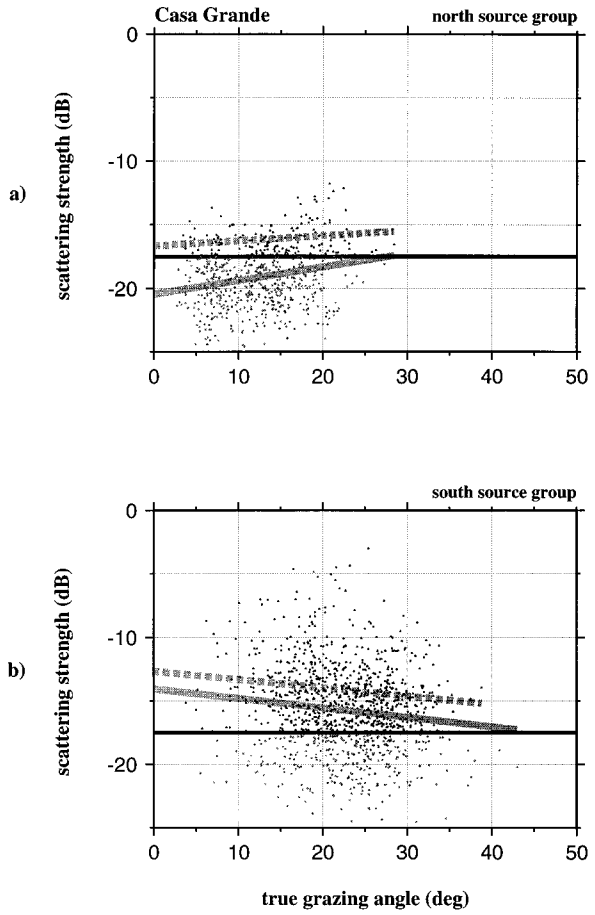


FIG. 10. Like Fig. 8 but for Casa Grande. For the regression lines, in (a) slope $=0.11\pm0.02$ dB/deg and $r=0.24$ and in (b) slope $=-0.08\pm0.02$ dB/deg and $r=-0.13$. For signals above the threshold at -17.5 dB in (a) slope $=-0.04\pm0.02$ dB/deg and $r=0.18$ and in (b) slope $=-0.07\pm0.01$ dB/deg with $r=-0.15$.

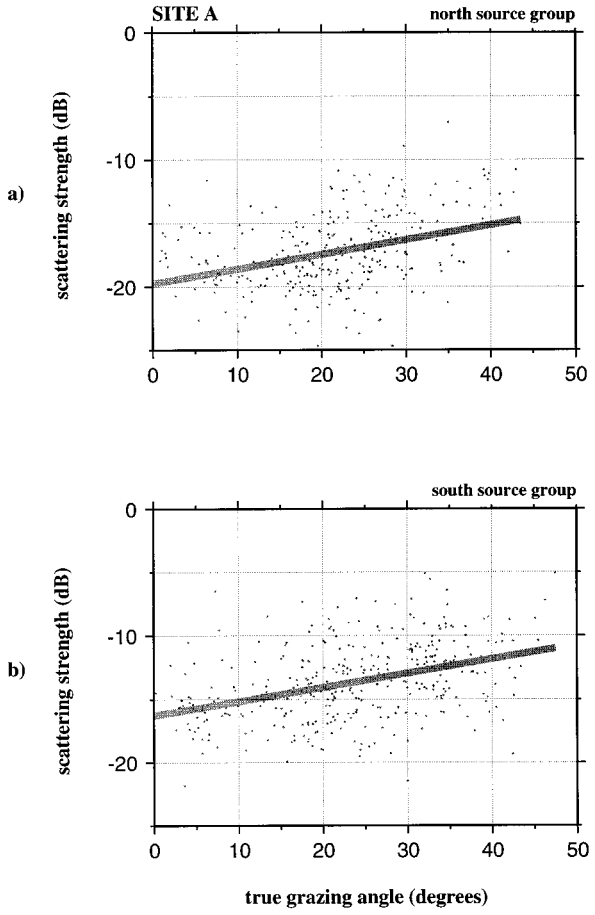


FIG. 11. Scattering strength of stacked intersecting beams, in site A, where only beams having source azimuth within 10° of being normal to the dip plane strike (''back-looking'' dips) are included in the stack. The solid gray line is the linear regression fit, which for (a) the north source group data, has slope of 0.11 ± 0.02 dB/deg with a correlation coefficient $r=0.36$, and for (b) the south source group data, has slope of 0.11 ± 0.01 dB/deg and $r=0.39$. Note that reducing the stack to include only scattering from ''back-looking'' seafloor areas does not improve the correlation nor does it significantly change the slope of the observed trend compared to Fig. 8.

the strike of the dip planes. In effect, this shows the scattering strength of the most directly ''back-looking'' seafloor surfaces. We look at this since it has been suggested,[7] that ''back-looking'' seafloor produce the strongest backscatter signals in monostatic scattering experiments.

We also find the weighted average scattering strength of the four dip triangles within each grid cell, where the weights are the number of beams that make up the stack within each triangle. We reject any stack with less than four beams in the stack. These average scattering strengths, per grid cell, are then overlaid onto the bathymetry contour maps for each of the study areas as shown in Figs. 14–16. In these plots we have chosen the threshold of the gray scale range to highlight the locations corresponding to just the highest scattering strengths.

IV. ANALYSIS

A. Scattering strength versus true grazing angle

It is clear in Figs. 8–10 that there is a trend toward increasing scattering strength with increasing true grazing angle for both IC and OC crustal areas. One exception to this is the case of the south source group data in Casa Grande [Fig. 10(b)]. This trend is quantified by the linear regression lines calculated for each data set. The specific slopes and correlation coefficients for each case are given in the figure captions. In general, the slope of this line is on the order of $0.1=\pm0.01$ dB/deg, but the correlation coefficients are small (<0.5). This implies that the functional relationship between scattering strength and true grazing angle is poorly described by a linear function. It is also clear that the broad distribution in scattering strength at each grazing angle would not yield a strong correlation for any functional description. We conclude from this that, at the scale of the resolution of Hydrosweep bathymetry data, there is not a simple functional relationship between scattering strength and true grazing angle.

In spite of this we can ask: Is the trend observed, sufficiently strong to differentiate IC crust from OC crust? The difference in mean seafloor dip of the IC crust area, site A, compared to mean dips of the IC crust areas (Fig. 6), is about 5°, which yields only a 0.5 dB difference in scattering strength between the two crustal areas. Based on our estimates of error in scattering strength calculation (±1 dB) and true grazing angle ($\pm2°$), we conclude that the trend is not

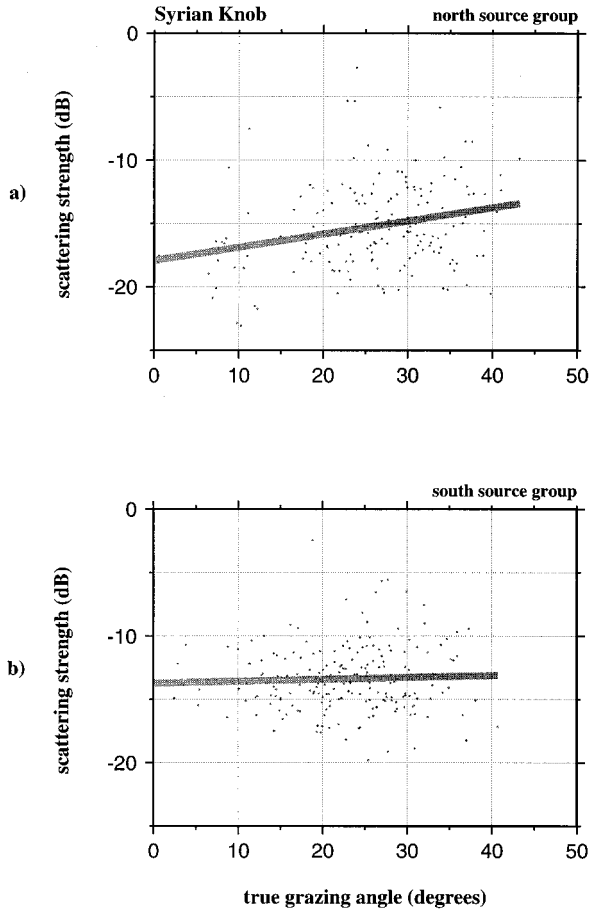


FIG. 12. Like Fig. 11 but for Syrian Knob. For the regression lines, in (a) slope $=0.10\pm0.03$ dB/deg and $r=0.26$ and in (b) slope $=0.02\pm0.03$ dB/deg and $r=0.04$. Note that the stack of data from only the ''back-looking'' seafloor areas does not improve the correlation nor does it significantly change the slope of the observed trend compared to Fig. 9.

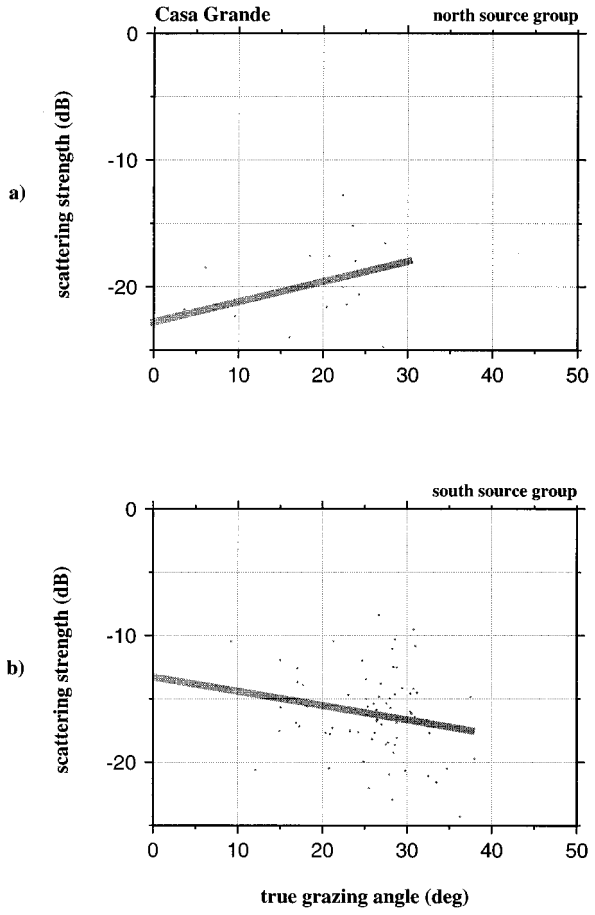


FIG. 13. Like Fig. 11 but for Casa Grande. For the regression lines, in (a) slope $=0.16\pm0.09$ dB/deg and $r=0.39$ and in (b) slope $=-0.11\pm0.06$ dB/deg and $r=-0.2$. Note that the stack of data from only the ''back-looking'' seafloor areas does not improve the correlation nor does it significantly change the slope of the observed trend compared to Fig. 10.

steep enough to distinguish between the crustal regions so that the idea expressed in Fig. 3(a) cannot be applied. We also calculated the averages of scattering strengths in 1° interval and found that the standard deviation is about ±2.5 dB. This suggests that there is about ±1.5 dB of variation in scattering strength not directly correlated to seafloor dip, which again is larger than the difference in strength that can be associated with the average difference in dip between the two crustal areas. The plots also show that there is no distinct difference in the trend or magnitude of scattering strength between the two crustal types, i.e., the relationships suggested in Fig. 3(b) and (c) are not found.

We considered the possibility that only very high scattering strengths show a strong correlation to grazing angle. By choosing a threshold [based on the highlighting of steep dip features in map form (Figs. 14–16)], we look for a more constrained relationship for signals above this threshold. The linear fit and correlation coefficients after including this constraint are also shown in Figs. 8–10. Again the results do not suggest a functional fit between the parameters. In this case, the fact that the Hydrosweep bathymetry data does not define seafloor dip above 30°–40° may be strongly affecting the result.

In Figs. 11–13 we test the idea that strong backscatter is produced primarily by the most directly ''back-looking'' seafloor areas. In these plots, only signals from areas having dip azimuth within 10° of the return direction to the source were included in the stacks. The fact that the slope of the regression lines and the breadth of the distributions in scattering strength do not differ for each case, compared to Figs. 8–10, suggests that ''back-looking'' seafloor areas do not dominate the signals observed in monostatic reverberation backscattering.

B. Scattering strength maps

In spite of the lack of a clear functional relationship in cross plots, there is a correspondence between high-scattering strength and steeply dipping seafloor features observed in all of the scattering strength maps (Figs. 14–16). These maps suggest that the scattering strength can be used to determine the shape of large scale high-standing seafloor morphology. The linear bathymetric ridges in site A, Fig. 14(b), correspond to linear scattering strength highs, but only when insonified from directions at near normal angles relative to the ridge strike. When insonified from highly oblique

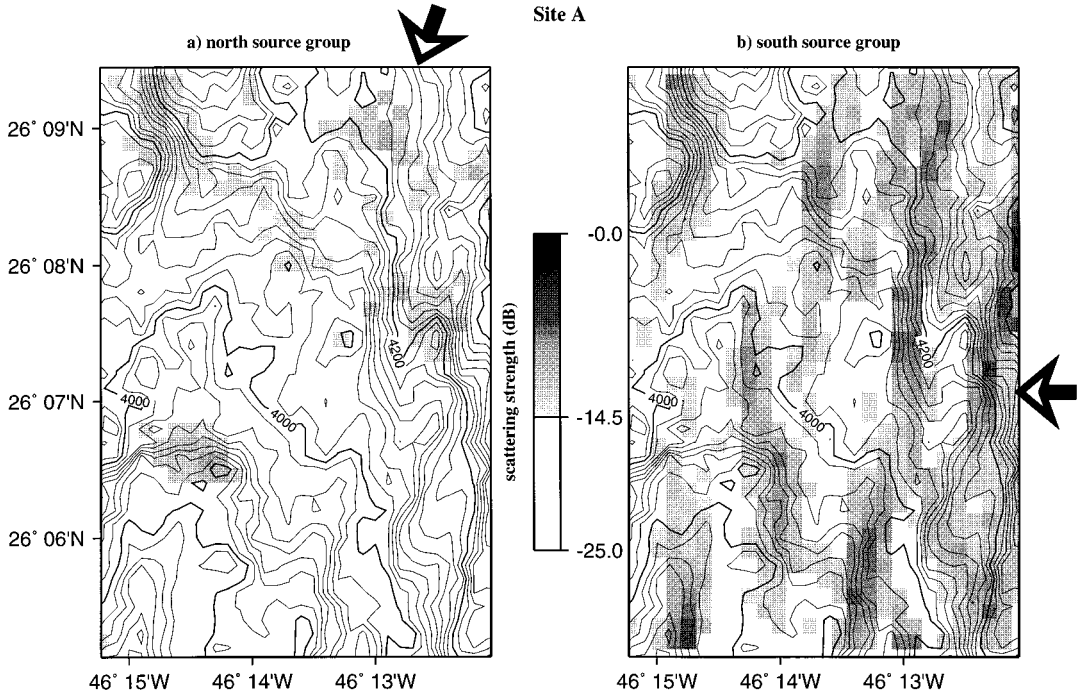


FIG. 14. Average (per grid cell) of stacked scattering strength of all intersecting beams mapped to site A. The gray scale shows the locations of just the highest values (above a threshold at −14.5 dB). The arrows show the general direction of insonification from each source group. In (a) the north source group data shows a correspondence to a few isolated topographic highs that have steep dip facing the source direction. When insonified from the east by the south source group, (b), the asymmetric linear ridges parallel to and facing the mid-Atlantic ridge (east) are highlighted. These ridges are characteristic of OC crustal areas.

angles, Fig. 14(a), only isolated peaks, in particular the peak at the end of Blackjack Ridge (26°06.5′N,46°14.5′W), stand out as a scattering strength highs. Such isolated highs, which seem to be associated with changes in dip direction, may be indicative of diffraction scattering from edges. The linearity and strong dependence of scattering strength on source azimuth, may be a useful qualitative indicator of OC crustal areas.

In Fig. 15(a), the high backscatter region wraps around the flank of the main peak of Syrian Knob, mimicking the shape of the topographic corner facing toward the north source group. A lower ridge, extending to the southwest is also delineated by high-scattering strength. In Fig. 15(b), the arcuate scattering strength high corresponds to a similar shape in the bathymetry facing the direction of the south source group. Since IC crustal areas have blocky, rather than linear, structures, strong scattering responses are observed from all source azimuths, although not necessarily corresponding to the same portion of a structure. When contrasted to the response observed for the OC crust at site A, we could conclude that the comparison of mapped spatial patterns of scattering strength for large changes in source azimuth, can be used to differentiate IC and OC crustal areas.

This conclusion must be tempered by the results of mapping in the Casa Grande area which is also described as being on the IC side of the discontinuity. The results for this area are somewhat problematic in that many of the beams from both source groups that intersected the area were within 30° of endfire and were rejected early in the processing. This resulted in fewer beams in both north and south source group stacks. For the north source group, Fig. 16(a), the strong scattering strength maps, generally, to a topographic corner, i.e., a change in dip azimuth. Even though Casa Grande is a steep feature, it does not have the blocky structure that is expected for IC crust. Since the dip azimuth of the seafloor structure facing the north source group is relatively oblique to the direction of insonification, the results are similar to those for oblique insonification at site A [Fig. 14(a)].

The mapping of the south source group data, Fig. 16(b), suggests that the whole flank of Casa Grande has high scattering strength when insonified from the direction near normal to the strike of the seafloor dip. There is some striping which may indicate that some areas of the slope are more responsive than others. Higher-resolution bathymetry data would help to resolve whether there is detailed structure associated with this striping.

The map results for Casa Grande appear to be more similar to the response observed at site A than at Syrian Knob. This seems to be the result of Casa Grande being a less blocky structure than Syrian Knob. Therefore, before using the difference in acoustic response of lineated versus blocky structure to differentiate IC and OC crust, some further consideration of the scales of these features should be made.

C. Scattering strength versus seafloor dip

We estimate the direct relationship between seafloor dip and scattering strength by using the linear regression fit to each data set to remove the approximate effect of beam grazing angle from the scattering strength. In Figs. 17–19, we plot the resulting scattering strength versus seafloor dip,

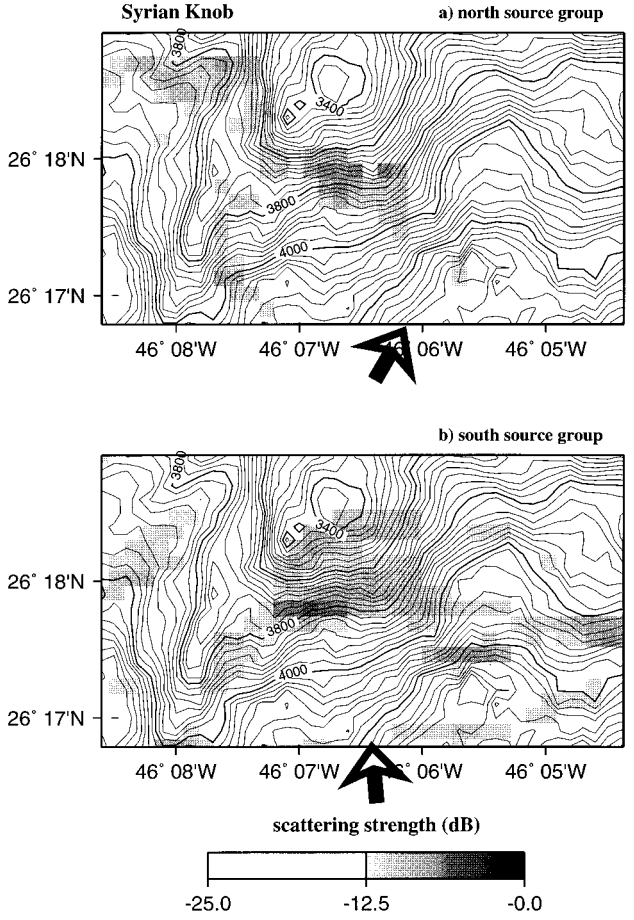


FIG. 15. Average (per grid cell) of stacked scattering strength of all intersecting beams mapped to Syrian Knob. The gray scale shows the locations of just the highest values (above a threshold at −12.5 dB). The arrows show the general direction of insonification from each source group. (a) When insonified by the north source group the steep flanked corners facing the sources are highlighted. The shape of the peaks in the backscattered signals mimic the shape of the seafloor feature. (b) When insonified by the south source group, the south facing cols and back wall of the block produce a similar arcuate-shaped scattering strength high.

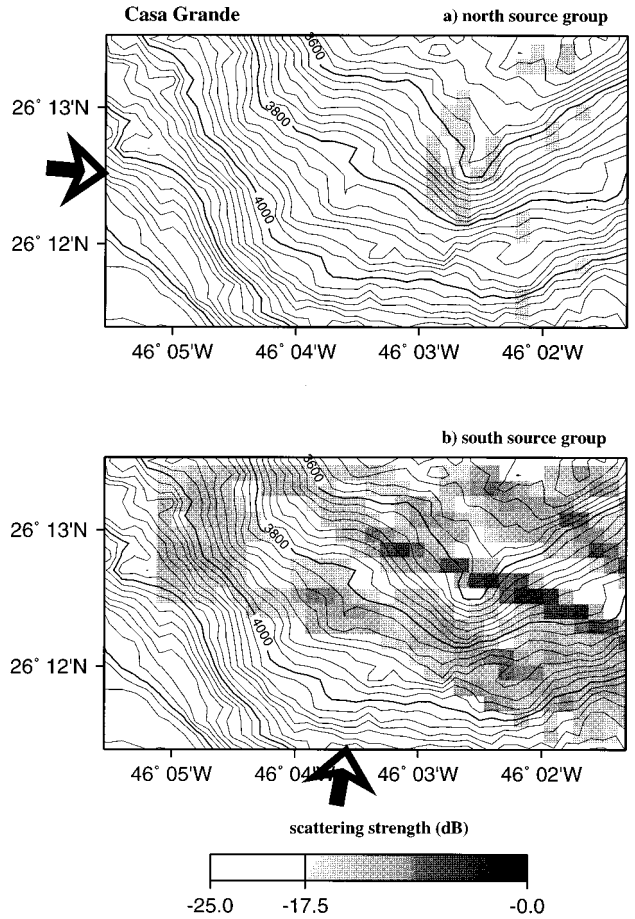


FIG. 16. Average (per grid cell) of stacked scattering strength of all intersecting beams mapped to Casa Grande. The gray scale shows the locations of just the highest values (above a threshold at −17.5 dB). The arrows show the general direction of insonification from each source group. (a) The data from the north source group shows a highlight near the top of the slope where a corner is formed. Strong scattering associated with such sharp changes in slope direction are likely to be due to edge diffractions. The majority of the slope is at an oblique angle to the main source azimuth and shows little response, similar to Fig. 14(a). (b) When insonified from the direction of the south source group, which is more perpendicular to the strike of the dip, the entire slope appears to have relatively high-scattering strength, similar to the response of the ridges in Fig. 14(c).

where seafloor dip is the dip that would be observed by an acoustic beam traveling horizontally (beam grazing angle $= 0°$). If scattering strength is more, or less, dependent on beam grazing angle than on seafloor dip, then making this adjustment is expected to change the slope of the regression line. Instead, what we see is that the slope of the regression lines are essentially unchanged. This suggests that scattering strength either varies similarly with changes in grazing angle or seafloor dip, or that we cannot resolve the difference, due either to the insensitivity of the scattering strength to changes in dip or to the low resolution of dip as measured by Hydrosweep data.

V. DISCUSSION AND CONCLUSIONS

Our analysis of the monostatic backscatter signal from IC and OC crust, in the region of site A, has led us to two seemingly contridactory conclusions. In map form, Figs. 15–17, there is good correspondence between high-scattering strength and source facing (''back-looking'') slopes with steep dip and the geometric patterns of high-scattering strength mimic the shape of the bottom feature. This result agrees with results reported by other investigators.[7,8,10] Contrary to this, the analysis based on crossplots of scattering strength versus true grazing angle and seafloor dip suggest that there is not a functional relationship between these parameters. The observed trend of increasing scattering strength with increasing seafloor dip or grazing angle is only about 0.1 dB/deg, is a weak linear correlation, and is insufficient, given the broad distribution in scattering strength at each angle, to distinguish low-dip areas from high-dip areas.

We believe that these observations are not contradictory, but rather show the effect of geologic scale on the analysis of seafloor scattering. The lack of a stronger numerical correlation may indicate that the actual scattering mechanisms are only weakly dependent on dip at the scale measured by Hydrosweep. The good map correspondence is mainly due to the fact that steeply dipping slopes are part of large high-standing features and are therefore more heavily insonified

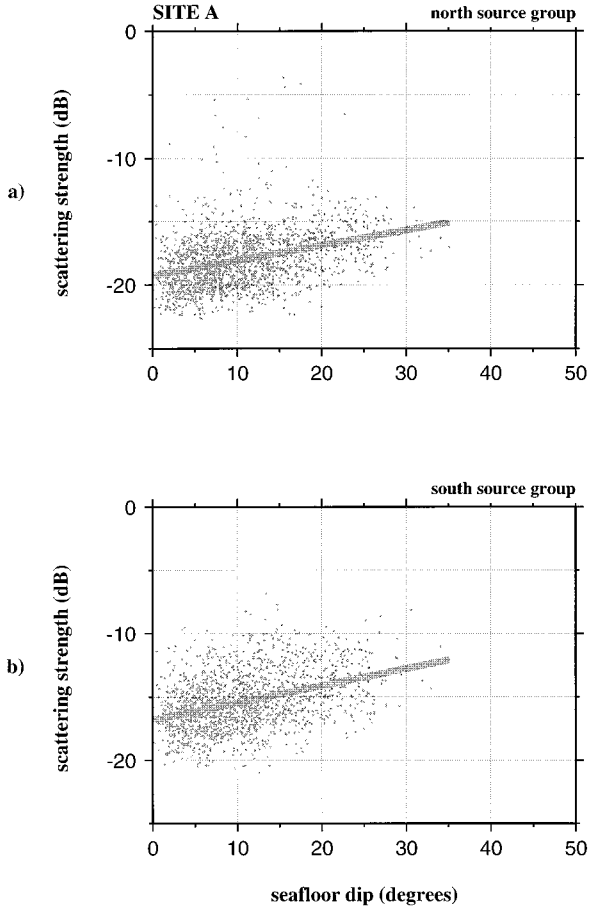


FIG. 17. Scattering strength of stacked intersecting beams, with the estimated effect of beam grazing angle removed, versus seafloor dip for site A. The linear fit shown in Fig. 8 is used to estimate the scattering strength due to beam grazing angle relative to a flat seafloor. The observed scattering strength is then reduced by this amount and plotted relative to the seafloor dip. This is the dip that would be observed by a horizontally propagation (grazing angle $=0°$) acoustic wave. After making this adjustment, the regression lines for the north source group (a) have slope of 0.12 ± 0.01 dB/deg and a correlation coefficient $r=0.35$ and for the south source group (b) have slope of 0.13 ± 0.01 dB/deg and $r=0.39$. The fact that these trends are the same as in Fig. 8 indicates that there is no distinction between the grazing angle and seafloor dip in terms of seafloor acoustic response.

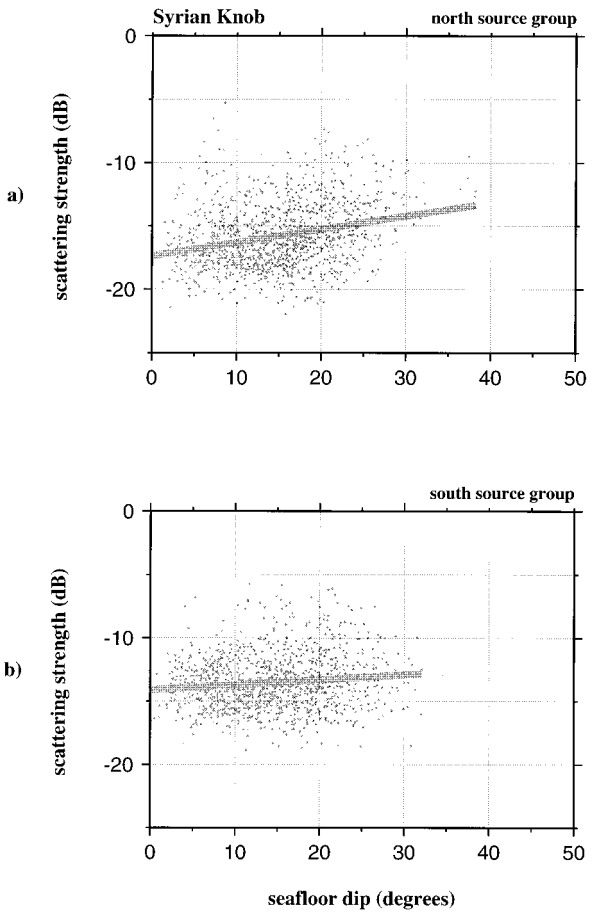


FIG. 18. Like Fig. 17 but for Syrian Knob. The linear fit shown in Fig. 9 is used to estimate and remove the scattering strength due to beam grazing angle relative to a flat seafloor. For the regression lines, in (a) slope $=0.10\pm0.01$ dB/deg and $r=0.27$ and in (b) slope $=-0.04\pm0.01$ dB/deg and $r=0.12$. These trends for seafloor dip are essentially the same as in those found for true grazing angle, shown in Fig. 9.

than deeper regions and are also likely to be more lightly sedimented. The correspondence of high-scattering strength to these features does not define a quantitative relationship between scattering strength and seafloor dip.

The observed variations in scattering strength are larger than can be explained by data error or seafloor dip. A significant part of the variation in scattering strength observed in the cross plots, is likely due to variations at acoustic wavelength scales (about 6 m) of seafloor or subseafloor properties. Even in the case of what may be considered smooth and flat seafloor, volume density and velocity heterogeneities at wavelength scales produce anomalous backscattered signals.[22,23]

For the data used in this analysis the peak acoustic wavelength was 6.5 m which is much smaller than the scale of resolution available from either the Hydrosweep data or the acoustic beam footprint. We did not, in this study, investigate the effects of small scale (less than 200 m) features. Instead, we tried to quantify and understand the effects of large scale features observed on IC and OC crust as a means of differentiating these crustal regions. The results of this study suggest that although scattering strength does in general increase with increasing seafloor dip, the trend is not strong enough to reliably predict seafloor dip from scattering strength. The difference in the mean dip of IC and OC crust is only a few degrees, and the overall distribution of dip and scattering strengths are so similar that no quantitative distinction between these areas, based on comparison of Hydrosweep data and backscatter data, can be made.

However, a qualitative analysis of the data based on mapping the scattering strength may be useful in distinguishing IC and OC crust. The correspondence between steeply dipping features and high-scattering strength can be used to delineate shapes of seafloor features. The long linear ridges characteristic of OC crust can be identified from backscattered data if at least two principle directions of insonification are used to acquire the data. We observed that when the high-standing ridge flanks are insonified at ''back-looking'' angles, they produce much higher average scattering strength. In contrast, blocky seafloor regions associated with

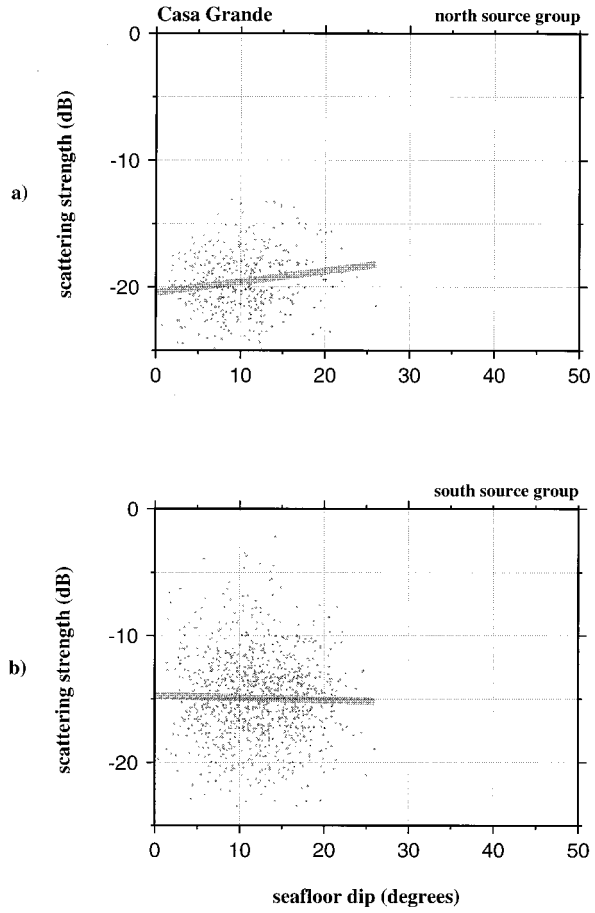


FIG. 19. Like Fig. 17 but for Casa Grande. The linear fit shown in Fig. 10 is used to estimate and remove the scattering strength due to beam grazing angle relative to a flat seafloor. For the regression lines, in (a) slope $=0.09\pm0.02$ dB/deg and $r=0.16$ and in (b) slope $=-0.02\pm0.02$ dB/deg and $r=-0.03$. These trends for seafloor dip are essentially the same as in those found for true grazing angle, shown in Fig. 10.

IC crust show high-scattering strength from different source directions since at least some portion of high-standing regions are insonified at ''back-looking'' angles.

It has been our goal to determine a method for distinguishing IC from OC crust using acoustic backscatter signals. In doing so, we have focused our attention on finding a relationship between seafloor dip and scattering strength at the scale of the resolution of Hydrosweep bathymetry data. Mapping data to the seafloor avoids making a quantitative judgement about any seafloor characteristic that might be controlling the level of backscatter. Contrary to this, plotting scattering strength versus seafloor dip is an explicit search for a quantitative relationship. We believe that there are quantitative relationships between scattering strength and seafloor properties, but conclude that we must look at wavelength scale features, i.e., much smaller than Hydrosweep resolution, to determine them.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Brian Tucholke for discussions on geological background and for internal review of this paper. Tom Bolmer has provided valuable advice and assistance on the computer applications. Dan Lizarralde provided encouraging and insightful discussions on the analysis. We would also like to express our appreciation for the support and review of this work provided by J.W. Caruthers at the Naval Research Lab, Stennis Space Center. This work was carried out under Office of Naval Research Grant Nos. N00014-93-1-1352 and N00014-90-J-1493. Woods Hole Contribution Number 9255.

APPENDIX: TRANSMISSION LOSS APPROXIMATION

In this Appendix we show that the transmission loss associated with geometric spreading and source spatial beam (radiation) pattern can be approximated by a shifted cylindrical spreading loss function.

The VLA far-field beam pattern is estimated by the response function of a vertical line source in the vicinity of a free surface (Ref. 4, p. 220). The equation for the far-field spatial beam pattern, $S(\theta)$, in terms of pressure for a harmonic (cw) source, is given by

$$S(\theta)=\frac{\sin[2\pi(v-V)(L/2)]}{2\pi(v-V)(L/2)}\exp[i2\pi(v-V)x_0] - \frac{\sin[2\pi(v+V)(L/2)]}{2\pi(v+V)(L/2)}\exp[-i2\pi(v+V)x_0], \tag{A1}$$

where x_0 is the depth from the free surface to the center of the line source which has length L and

$$v=f\frac{\sin\theta}{c}, \tag{A2}$$

$$V=f\frac{\sin\theta_s}{c}, \tag{A3}$$

where f is the source frequency, θ is the observation angle (grazing angle), θ_s is the source array downward steering angle, and c is the propagation velocity. For simplicity, uniform velocity across the array is assumed so that c is a constant. For the site A experiment $\theta_s=9°$, $L=20.6$ m, $x_0=184$ m, $c=1522$ m/s at source depth, and the frequency ranges from 200–255 Hz.

Unless the acoustic propagation environment is known very accurately, it is best to represent the source beam pattern by the ''incoherent'' pressure field, in which the phase information is neglected.[21] The discrete expression for the incoherent broadband source beam signal level $(BL(\theta))$ at a particular observation angle is

$$BL(\theta)=10\log_{10}\left(\Delta f\sum_f |S(\theta)|^2\right), \tag{A4}$$

where Δf is the frequency interval. For our calculation f ranged from 200 to 255 Hz and $\Delta f=1$ Hz. We normalize $S(\theta)$, prior to the summation over frequency, such that the total power output is equivalent to a point source that produces a pressure of $1\,\mu$Pa on the surface of a sphere 1 m in radius. This approach is consistent with defining the source level relative to the total energy emitted from a sphere, 1 m in radius, located at the center of the array. After the calcu-

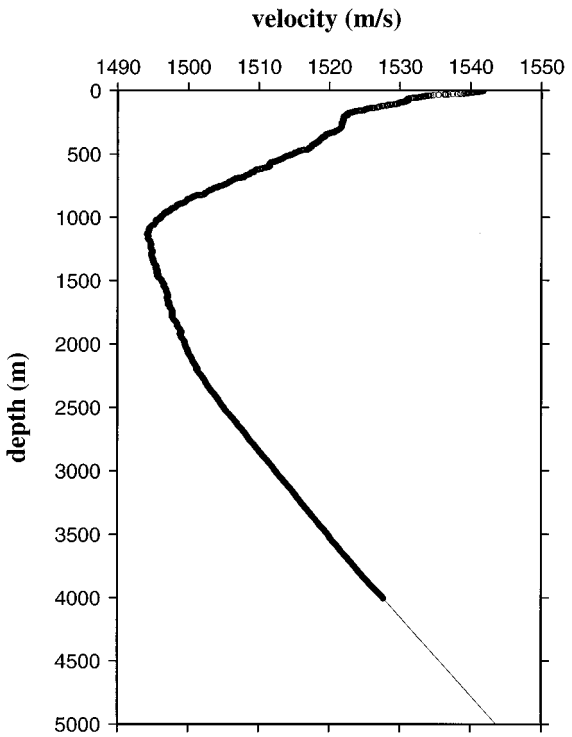


FIG. A1. Velocity-depth profile used in ray tracing. The data are from a CTD profile collected at 26°14.18′ N, 46°13.51′ W, on 7 Dec. 1993. Values below 4000 m (thin line) are extrapolated from the measured profile.

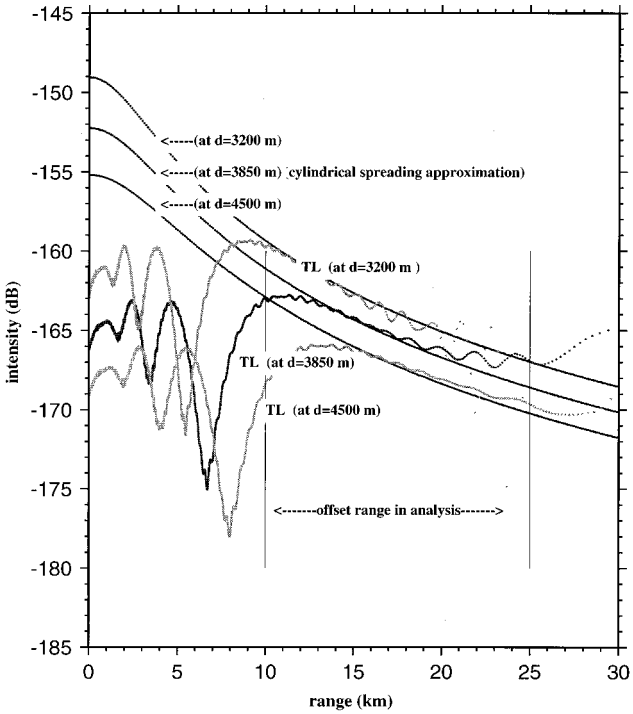


FIG. A2. Transmission loss predicted by two-way cylindrical spreading decay (shifted by $TL_0 = 0.00243d + 71.15$ dB, where d is the depth) compared to the transmission loss predicted by total energy decay (TL), calculated as the combination of ray divergence and the incoherent spatial beam pattern, for the depth range used in our analysis. The cylindrical spreading curves predict the energy loss between 12 and 22 km within ± 0.5 dB. Between 10 and 12 km the error is larger but only for deep seafloor areas. In the worst case, 4500 m at a range of 10 km, the error is about -4 dB. In the far ranges, the error is larger for the most shallow depths. In the worst case (3200 m at 25 km), the error is about 4 dB, however little data was observed in this range (see Fig. 7).

lation for the total incoherent pressure field, we again normalize the broadband beam level so that its peak level (at $\theta_s = 9°$) is 0 dB. This is consistent with the idea that the actual peak source level (SL = 230 dB) is measured on the central beam axis and that all other beam levels are measured relative to this.

The theoretical geometric spreading loss as a function of energy density is determined from ray divergence and combined with $BL(\theta)$ to find a net transmission loss correction term. Ray tracing through an ocean acoustic velocity model (Fig. A1) was used to estimate the true geometric energy decay due to wave field divergence. The energy density, or intensity, in a multilayered (velocity) ocean can be described in terms of horizontal range from the source, r, at a fixed depth, z, and the variation of the ray parameter, p, with range; that is, dp/dr. The ray parameter, for horizontally layered media, is defined by (Ref. 24, p. 92)

$$p = \frac{\sin\phi}{v}, \tag{A5}$$

where v is the velocity in any particular layer and ϕ is the propagation incidence angle relative to the vertical. We calculate dp/dr as a function of depth and range in the ray tracing. The decay in energy density can be shown to be given by (Ref. 24, p. 127),

$$\frac{I}{I_0} = \frac{1}{r} \frac{v_0^2 p}{\sqrt{1-(v_0 p)^2}\sqrt{1-(vp)^2}} \frac{dp}{dr}, \tag{A6}$$

where I_0 is an initial source intensity, v_0 is the velocity at the source, and v is the velocity at the depth for which we are calculating the loss. Transmission loss, based on ray divergence, expressed in dB, is then

$$\mathrm{TL_{rd}} = -10 \log_{10} \frac{I}{I_0}. \tag{A7}$$

To account for two-way propagation we multiply the decay expressed in Eq. A7 by a factor of 2.

In Fig. A2 we show the net observed intensity, $\mathrm{TL} = -(BL - 2\mathrm{TL_{rd}})$, predicted for signals returned from the range of depths used in our analysis. This is energy intensity, for two-way propagation, as a function of range for a reference source and is in absolute units (i.e., dB down from an axial level of 1 μPa at 1 m from the center of the array). Actual geometric propagation loss in the ocean is described by spherical spreading loss close to the source and cylindrical spreading loss in the far field.[21] For simplicity in processing the data, we fit the net intensity curves with transmission loss curves based solely on cylindrical spreading. The loss predicted by cylindrical spreading is

$$\mathrm{TL_{cyl}} = 2 \times (10 \log_{10} \sqrt{r^2 + d^2}) + \mathrm{TL_0}, \tag{A8}$$

where d is the seafloor depth and $\mathrm{TL_0}$ is a reference level used to match transmission loss within the offset range used

in our analysis. From 10 to 25 km, the cylindrical spreading curve has similar slope to the envelope of the sum of the beam pattern and ray divergence curve but is offset in absolute units. In Fig. A2 we plot $-\mathrm{TL}_{\mathrm{cyl}}$ for three depths, with

$$\mathrm{TL}_0 = 0.00243d + 71.15 \text{ dB}, \tag{A9}$$

which fits $-\mathrm{TL}_{\mathrm{cyl}}$ to $BL - 2\mathrm{TL}_{\mathrm{rd}}$ in this depth range. The cylindrical spreading curves predict the energy loss between 12 and 22 km within ± 0.5 dB. Between 10 and 12 km the error is larger but only for deep seafloor areas. In the worst case, 4500 m at a range of 10 km, the error is about -4 dB. At far ranges, the prediction error is larger for the most shallow scatterers; for example at the 25-km range and 3200 m depth, the error is as much as 4 dB. However, very little data used in our analysis have ranges greater than 20 km (see Fig. 7) or depths greater than 4000 m.

We conclude that the correction needed to account for source beam pattern and geometric transmission loss can be approximated by a cylindrical spreading loss as

$$\mathrm{TL} = -(BL - 2 \times \mathrm{TL}_{rd}) \approx \mathrm{TL}_{\mathrm{cyl}}^{\mathrm{src}} + \mathrm{TL}_{\mathrm{cyl}}^{\mathrm{rec}} + \mathrm{TL}_0, \tag{A10}$$

where

$$\mathrm{TL}_{\mathrm{cyl}}^{\mathrm{src}} = 10 \log_{10} r_{\mathrm{src}} \tag{A11}$$

and

$$\mathrm{TL}_{\mathrm{cyl}}^{\mathrm{rec}} = 10 \log_{10} r_{\mathrm{rec}} \tag{A12}$$

taking r_{src} as the propagation distance (slant range) from the source to a seafloor grid cell and r_{rec} as the propagation distance from the receiver to the grid cell. The net approximation to the equivalent target scattering strength is

$$\mathrm{TS} \approx \mathrm{RL}_s - \mathrm{SL} + \mathrm{TL}_{\mathrm{cyl}}^{\mathrm{src}} + \mathrm{TL}_{\mathrm{cyl}}^{\mathrm{rec}} + \mathrm{TL}_0 \tag{A13}$$

The surface scattering strength

$$S_s = \mathrm{TS} - 10 \log_{10} A \tag{A14}$$

is then approximated by

$$S_s \approx RL_s - \mathrm{SL} + \mathrm{TL}_{\mathrm{cyl}}^{\mathrm{src}} + \mathrm{TL}_{\mathrm{cyl}}^{\mathrm{rec}} + \mathrm{TL}_0 - 10 \log_{10} A. \tag{A15}$$

[1] B. E. Tucholke and J. Lin, ''A geological model for the structure of ridge segments in slow spreading ocean crust,'' J. Geophys. Res. **99** (B6), 11 937–11, 958 (1994).

[2] J. P. Severinghaus and K. C. MacDonald, ''High inside corners at ridge-transform intersections,'' Mar. Geophys. Res. **9**, 353–367 (1988).

[3] J. A. Karson and H. J. B. Dick, ''Tectonics of ridge-transform intersections at the Kane Fracture Zone,'' Mar. Geophys. Res. **6**, 51–98 (1983).

[4] Office of Naval Research, Acoustic Reverberation Special Research Program, Initial Report, Acoustics Experiment, R/V CORY CHOUEST, 5–26 July 1993.

[5] B. E. Tucholke, ''Geological and geophysical survey of the Acoustic Reverberation Corridor, Mid-Atlantic Ridge,'' in *Acoustic Reverberation Special Research Program Symposium* (Scripps Institution of Oceanography, La Jolla, CA, March 1991).

[6] M. C. Kleinrock, ''Capabilities of some systems used to survey the deep-sea floor,'' *CRC Handbook of Geophysical Exploration at Sea, Hard Mineral*, edited by Richard A. Geyer (CRC, Boca Raton, FL, 1992), 2nd ed., pp. 35–85.

[7] N. C. Makris, L. Z. Avelino, and R. Menis, ''Deterministic reverberation from ocean ridges,'' J. Acoust. Soc. Am. **97**, 3547–3574 (1995).

[8] J. R. Preston, T. Akal, and J. Berkson, ''Analysis of backscattering data in the Tyrrhenian Sea,'' J. Acoust. Soc. Am. **87**, 119–134 (1990).

[9] R. A. Stephen, P. R. Shaw, and J. W. Caruthers, ''Modeling deterministic features in the AR-SRP acoustic data,'' in *Acoustic Reverberation Special Research Program Symposium* (Scripps Institution of Oceanography, La Jolla, CA, December 1993).

[10] P. R. Shaw, D. K. Smith, R. A. Stephen, and J. W. Caruthers, ''Multiscale, multisensor analysis of seafloor topography in the ONR Atlantic Natural Laboratory,'' in *Acoustic Reverberation Special Research Program Symposium* (Scripps Institution of Oceanography, La Jolla, CA, December 1993).

[11] N. C. Makris, ''Imaging ocean-basin reverberation via inversion,'' J. Acoust. Soc. Am. **94**, 983–993 (1993).

[12] B. E. Tucholke, M. C. Kleinrock, and W. K. Stewart, ''Fine-scale geological and geophysical surveys at sites B′, C′, A and D in the Acoustic Reverberation Corridor,'' in *Acoustic Reverberation Special Research Program Symposium* (Scripps Institution of Oceanography, La Jolla, CA, December 1993).

[13] H. Schouten and R. S. White, ''Zero-offset fracture zones,'' Geology **8**, 175–179 (1980).

[14] J. O. A. Robertsson and A. Levander, ''A numerical study of seafloor scattering,'' J. Acoust. Soc. Am. **97**, 3532–3546 (1995).

[15] R. A. Stephen and M. E. Dougherty, ''Canonical seafloor models and the finite difference method for low-angle acoustic backscatter,'' in *Computational Acoustics*, edited by R. L. Lau and A. R. Robinson (Elsevier Science, Amsterdam, 1993), Vol. 1, pp. 227–246.

[16] A. Baggeroer (December 1994, personal communication).

[17] P. Elisseeff, ''ARSRP 93: Acoustic shape estimation of a linear towed array,'' Internal report, Dept. of Ocean Engineering, MIT (1995).

[18] 15-kHz Hydrosweep bathymetry data were collected during the 1992 R/V EWING Cruise 9208. The Principal Investigator B. E. Tucholke and Co-Principal Investigators, M. C. Kleinrock and J. Lin, were from Woods Hole Oceanographic Institution.

[19] N. R. Grindlay, P. J. Fox, and P. R. Vogt, ''Morphology and tectonics of the Mid-Atlantic Ridge (25°–27°30′S) from Sea Beam and magnetic data,'' J. Geophys. Res. **97**(B5), 6983–7010 (1992).

[20] R. J. Urick, *Principle of Underwater Sound* (McGraw-Hill, New York, 1983), 3rd ed.

[21] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1994).

[22] S. A. Swift and R. A. Stephen, ''The scattering of a low-angle pulse beam from seafloor volume heterogeneities,'' J. Acoust. Soc. Am. **96**, 991–1001 (1994).

[23] M. Gensane, ''Sea-bottom reverberation: The role of volume inhomogeneities of the sediment,'' in *Ocean Reverberation*, edited by D. D. Ellis, J. R. Preston, and H. G. Urban (Kluwer Academic, Dordrecht, 1993), pp. 59–64.

[24] K. Aki and P. G. Richards, *Quantitative Seismology, Theory and Methods* (Freeman, New York, 1980), Vol. 1.

Multifrequency acoustic measurement of concentration and grain size of suspended sand in water

Jan Schat
Institut für Elektrische Messtechnik und Grundlagen der Elektrotechnik (Professor Bethe), University of Braunschweig, Hans-Sommer-Str. 66, D-38106 Braunschweig, Germany

(Received 24 May 1994; revised 17 December 1995; accepted 22 June 1996)

The paper describes an ultrasonic backscatter device for measuring vertical profiles of the concentration of suspended sand in water in the presence of air bubbles. The use of six frequencies permits a simultaneous measurement of the concentration of suspended sand, the grain size, and the concentration of air bubbles. For measuring in the near field of a transducer, the sensitivity of an ideal piston-like transducer in the near field is calculated and compared with measurements. A simple, but fast and efficient inversion algorithm to convert the backscattered intensities into sediment concentration and grain size is described. Measurements of known suspensions with concentration and grain size are presented, they exhibit a low error (20% rms) over a grain radius from 40 to 300 μm. *In-situ* measurements of the concentration of suspended sediment in the surf zone are compared with the results of two optical backscattering sensors (OBS). © *1997 Acoustical Society of America.* [S0001-4966(97)02412-0]

PACS numbers: 43.30.Gv, 43.30.Ma, 43.30.Pc [JHM]

INTRODUCTION

The measurement of concentration of suspended sediment in seawater is an important problem in marine geology. The main objectives for measurement of suspended sediment concentration are the validation of models of the erosion and resuspension, and providing data for assessment of erosion and deposition rates. In usual applications it is required to measure the concentration at different levels, typically in the range 0–1 m above the seabed, with a high temporal (approximately 0.1 s) and spatial (1 cm) resolution.

While for measurements in the laboratory, scales and Coulter–Counters permit a precise determination of the concentration and particle size, for *in-situ* measurement mainly optical and acoustical methods are applied. In the last 10 years, acoustic systems have won recognition.[1–5] In these systems, a downlooking transducer situated about 1 m above the bottom emits short (approximately 10 μs) ultrasonic pulses. The signal is backscattered by sand grains at different levels and received by the same transducer (monostatic configuration), then amplified, filtered, and range-gated. After corrections for attenuation, spherical spreading, and the backscattering properties of the suspended sediment, the concentrations at different discrete levels (range bins) are obtained.

One remaining problem is the presence of air bubbles in water. At standard measuring frequencies in the MHz range, the backscattered signal by air bubbles near the surf zone equals to about 0.1–1 g/l of sediment, and when using a single frequency, it is indistinguishable from sand. Also the sensitivity to a homogenous suspension in the near field of a transducer is given only by approximate formulae, which has led some authors to measure only in the far field.

The *in-situ* measurement of the grain size and sediment concentration was solved by Crawford *et al.*[5–7] by using a three-frequency system. For an additional measurement of the concentration of air bubbles, however, more frequencies are needed. This paper describes a system named ASAP (acoustic sand and air bubble sensitive profiler) that uses six frequencies and a standard inversion algorithm in order to reduce the error of sand concentration measurements in the presence of air bubbles.

I. BACKSCATTER THEORY

For obtaining the backscatter intensity $I(f,r)$ (r is the distance between range bin and transducer, f is the measuring frequency) of a range bin, the square of the output voltage $V(f,r)$ of the amplifier following the transducer must be corrected with the overall system sensitivity $S(f)$, the absorption due to water, the absorption due to suspended matter, and the range dependence due to near-field effects or spherical spreading $E(r,f)$, respectively,

$$I(f,r) := \frac{V^2(f,r)}{S(f)E(r,f)} \exp\left(-4\alpha_0 r - 4\int_{r'=0}^{r} \alpha_s \, dr' \right), \tag{1}$$

where α_0 is the coefficient of attenuation due to water and α_s is the attenuation due to suspended matter.

The backscatter intensity is a function of the particles in the range bin. It is caused by suspended sand of the concentration C_S, the grain radius a_s, and the backscatter cross section σ_s, as well as air bubbles with the concentration C_a, the bubble radius a_a and the backscatter cross section σ_a:

$$I(f,r) = K\left\{ C_s(r)\left\langle \frac{\sigma_s(f)}{a_s^3} \right\rangle + C_a(r)\left\langle \frac{\sigma_a(f)}{a_a^3} \right\rangle \right\}, \tag{2}$$

where the $\langle\ \rangle$ represents averaging over the grain radius and the bubble radius distribution, respectively, and K is a device-dependent factor.

 0001-4966/97/101(1)/209/9/$10.00

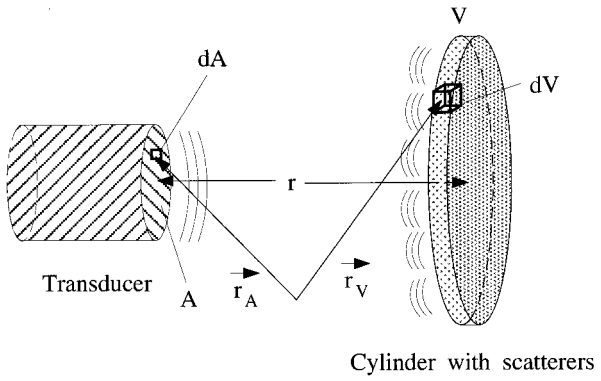


FIG. 1. Geometry for calculation of near-field sensitivity.

The parameter in Eq. (1) that is the most difficult to calculate is the near-field sensitivity $E(r)$. It only depends on the features of the transducer and will be discussed in the following section.

II. NEAR-FIELD SENSITIVITY

The near field of a piston-like transducer (a_t is the radius of the active area) extends to about $r_n := \pi a_t^2/\lambda$. The sensitivity $E(r)$ to a homogenous suspension of scatterers in the near field cannot be described by the r^{-2} term valid in the far field, but must be calculated numerically. Since this sensitivity is difficult to measure or to calculate, in some backscatter devices the transducers are placed so high above the bottom, that the relevant range bins near the bottom are in the far field of the transducer.[5–8] This leads, however, to longer paths between transducer and these range bins, and thus to higher losses due to spherical spreading and attenuation along the path. Also, the sand concentrations in the range bins near the transducer, that must be taken into account for the attenuation, cannot be calculated.

For calculating the sensitivity in the near field, an ideal piston-like transducer is assumed. The total pressure amplitude $p(\mathbf{r}_V)$ at a volume element dV given by $\mathbf{r}_V$ (Fig. 1) can be calculated:

$$p(\mathbf{r}_V) := \frac{p(r_0) r_0}{\pi a_t^2} \left| \int_A \frac{1}{|\mathbf{r}_V - \mathbf{r}_A|} \exp(-ik|\mathbf{r}_V - \mathbf{r}_A|) dA \right|, \tag{3}$$

where $p(r_0)$ is the on-axis pressure amplitude at the reference distance r_0. A part of this pressure is backscattered by the randomly dispersed scatterers in dV and causes an incident pressure $dp_i(\mathbf{r}_A)$ at the surface of the transducer:

$$\frac{dp_i(\mathbf{r}_A)}{dV} = \frac{d(\sqrt{\sigma_b})}{dV} \frac{p(\mathbf{r}_V)}{\sqrt{4\pi}} \frac{1}{|\mathbf{r}_V - \mathbf{r}_A|} \exp(-ik|\mathbf{r}_V - \mathbf{r}_A|), \tag{4}$$

where σ_b is the backscattering cross section of the scatterers in dV. The output signal of a transducer only depends on the pressure amplitude integrated over the active area, i.e., on the force amplitude dF_i/dV applied on the active area. It can be obtained by

$$\begin{aligned} \frac{dF_i}{dV} &= \int_A \frac{dp_i(\mathbf{r}_A)}{dV} dA = \frac{d(\sqrt{\sigma_b})}{dV} \frac{p(\mathbf{r}_V)}{\sqrt{4\pi}} \left| \int_A \frac{1}{|\mathbf{r}_V - \mathbf{r}_A|} \right. \\ &\quad \left. \times \exp(-ik|\mathbf{r}_V - \mathbf{r}_A|) dA \right| \\ &= \frac{d(\sqrt{\sigma_b})}{dV} \frac{1}{\sqrt{4\pi}} \frac{p(r_0) r_0}{\pi a_t^2} \left| \int_A \frac{1}{|\mathbf{r}_V - \mathbf{r}_A|} \right. \\ &\quad \left. \times \exp(-ik|\mathbf{r}_V - \mathbf{r}_A|) dA \right|^2. \end{aligned} \tag{5}$$

This force amplitude corresponds to an incident time-averaged power P_i with

$$\frac{d\overline{P_i}}{dV} = \frac{Z}{2} \frac{d(p_i^2)}{dV} = \frac{Z}{2} \frac{1}{\pi a_t^2} \frac{d(F_i^2)}{dV}, \tag{6}$$

where Z is the acoustic impedance of water. Consider now the effect not only of the volume element dV, but of all volume elements in the distance r to the transducer. Since the acoustical waves from all volume elements arrive incoherently, the incident time-averaged powers dP_i, not the forces dF_i, must be integrated over all volume elements dV in the range bin V to obtain the total incident power on the active surface:

$$\begin{aligned} \overline{P_i}(r) &= \int_V \frac{d\overline{P_i}(r)}{dV} dV \\ &= \frac{Z}{2\pi a_t^2} \left(\frac{d(\sqrt{\sigma_b})}{dV} \frac{1}{\sqrt{4\pi}} \frac{p(r_0) r_0}{\pi a_t^2} \right)^2 \\ &\quad \times \int_V \left| \int_A \frac{1}{|\mathbf{r}_V - \mathbf{r}_A|} \exp(-ik|\mathbf{r}_V - \mathbf{r}_A|) dA \right|^4 dV. \end{aligned} \tag{7}$$

For $r \gg r_n$, this formula describes the well-known spherical spreading in the far field. From Eq. (7), the constants can be eliminated by dividing $P_i(r)$ by $P_i(r_x)$, where r_x is an arbitrarily chosen reference distance. Then, the relative and dimensionless sensitivity $E(z)$ is obtained [note that these constants are device-dependent and thus contained in the system sensitivity $S(f)$]:

$$E(z) := \frac{\overline{P_i}(r)}{\overline{P_i}(r_x)}. \tag{8}$$

Evaluating this expression numerically for different r, a_t and λ shows, that for $a_t/\lambda \gg 1$, $E(r, a_t, \lambda)$ depends only on $z := r/r_n$. For transducers with $a_t/\lambda > 4$, which is commonly the case for the high-frequency transducers used for measuring sediment suspensions, the following approximation has less than 6% maximum error [referred to Eq. (8)] at distances $r > 0.06 r_n$:

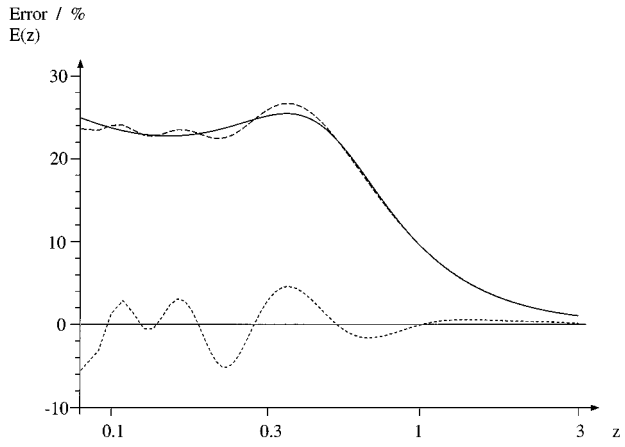


FIG. 2. $a_t/\lambda=12$: ——— $E(z)$ from Eq. (8), --- approximated $E(z)$ from Eq. (9) and .... relative error of approximated $E(z)$.

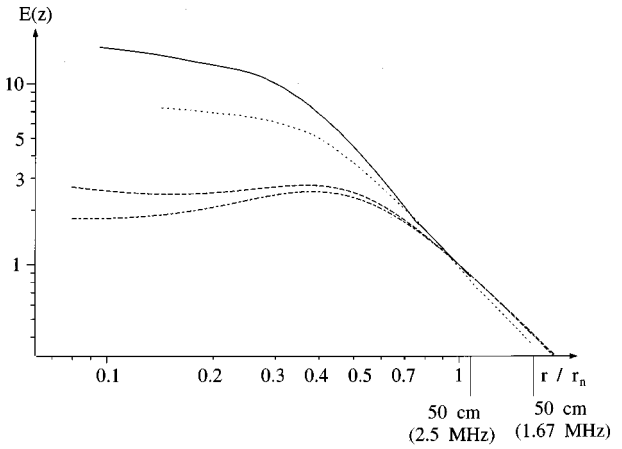


FIG. 4. Near-field sensitivity of the 2-MHz transducer, ---: Eq. (9), –·–·: Eq. (10), ...: 1.67 MHz, ———: 2.5 MHz.

$$E(z)=\frac{1}{1+a_1z+a_2z^2+a_3z\,\exp(-4z)+a_4z^2\,\exp(-5z)},$$

$$a_1=-0.012,\quad a_2=+5.154,$$

$$a_3=+26.6,\quad a_4=-80.113. \tag{9}$$

For $z\gg1$ (far field), it exhibits the expected $1/z^2$ behavior due to spherical spreading. Figure 2 shows the calculated [Eq. (8)] and the approximated [Eq. (9)] $E(z)$ as well as the relative error. Downing *et al.* proposed[9] the following formula as an approximation to the result of comparable calculations:

$$E(z)=\left(\frac{1}{z}\,\frac{1+1.35z+(2.5z)^{3.2}}{1.35z+(2.5z)^{3.2}}\right)^2. \tag{10}$$

This formula gives very similar results to Eq. (9). The two theoretical formulae are presented in the following together with measured data.

$E(z)$ can be measured simply using a dilute homogeneous suspension, in this case Eq. (1) and Eq. (2) indicate that $E(r)$ is proportional to $V^2(r)\,\exp(-4\alpha_0 r)$. To measure the sensitivity in the near field, a suspension of polystyrene spheres with a specific weight ($\gamma\approx1.035$ g/cm^3) nearly equal to that of water was prepared. In contrast to sand grains, these plastic spheres lead to a very homogeneous suspension. The concentration was about 0.1 g/l, so that absorption or multiscattering can be excluded as sources of error. The range measured was $r=5$–40 cm for all transducers. Figures 3, 4, and 5 show the responses for the 1-, 2-, and 6-MHz transducers, respectively, as well as the response predicted by Eqs. (9) and (10). These responses were normalized to $E(z)=1$ at $r=r_n$, i.e., at $z=1$. Equations (9) and (10) give almost the same result, but the difference of these theoretical formulae to the measured data is comparatively high, especially for the 2-MHz, but also for the 6-MHz transducer. This is very likely due to the fact that transducers cannot be considered as ideal baffled pistons, because the effective (''acoustic'') radius in some transducers is up to 27% smaller than the geometrical radius.[8] Figure 6 shows, as an example, the measured response of the 2-MHz transducer (that has the highest deviation from the theoretical response of all three transducers) as well as Eqs. (9) and (10), with an effective radius taken as 80% (1.67 MHz) and 70% (2.5 MHz) of the geometrical radius, respectively. These values were chosen because the manufacturer states that the effective radius is 75% to 80% of the geometrical radius at the

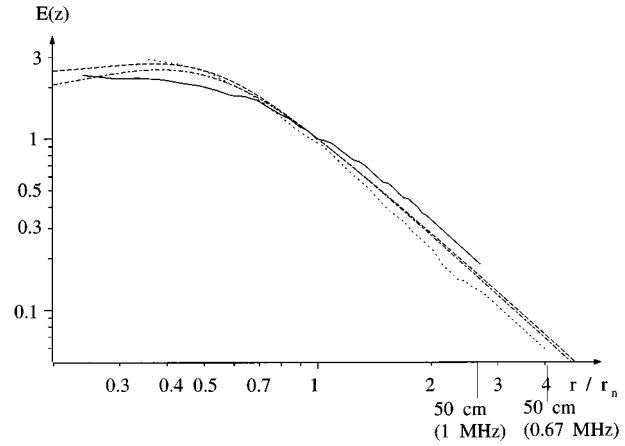


FIG. 3. Near-field sensitivity of the 1-MHz transducer, ---: Eq. (9), –·–·: Eq. (10), ...: 0.67 MHz, ———: 1 MHz.

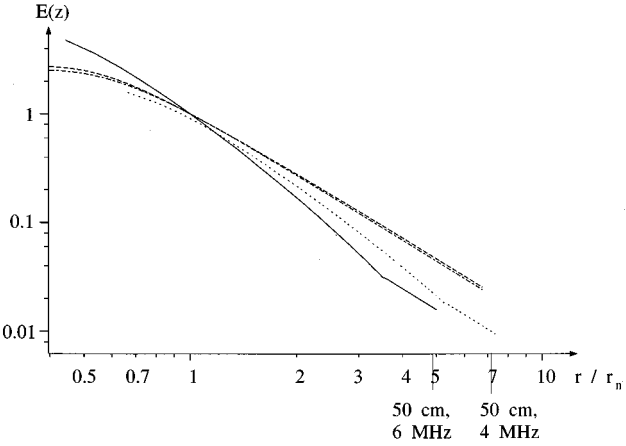


FIG. 5. Near-field sensitivity of the 6-MHz transducer, ---: Eq. (9), –·–·: Eq. (10), ...: 4 MHz, ———: 6 MHz.

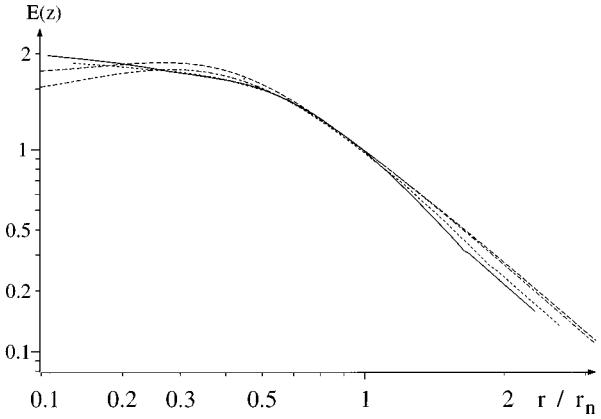


FIG. 6. Near-field sensitivity of the 2-MHz transducer, ---: Eq. (9), –·–·: Eq. (10), ...: 1.67 MHz (effective radius=80% of geometrical radius), ———: 2.5 MHz (effective radius=70% of geometrical radius).

nominal frequency of 2 MHz. The deviation of the measured data from the theoretical response is much lower than with an effective radius equal to the geometrical radius assumed, as in Fig. 4.

Therefore, it seems that exact values for the near-field sensitivity of a given transducer should be obtained by measurement of a homogeneous suspension with this transducer, rather than by calculation. This was also done with the system presented here.

III. FREQUENCY RESPONSE OF AIR BUBBLES

When measuring near the surf zone, air bubbles in the water are the largest source of error, and the most difficult to deal with. Unlike other sources of error, it cannot be reduced by averaging.

The following two experiments show, that the frequency response of the backscattering cross section $\sigma_a(f)$ of ''natural'' air bubbles (generated by a water jet or in the surf zone) in different situations is relatively constant. This is in contrast to suspended sand of different grain size, that shows a strongly varying frequency response depending on the grain size.

(1) In the laboratory, a water jet was injected into a water-filled tank to generate air bubbles, and the backscattered intensities were measured for each frequency after

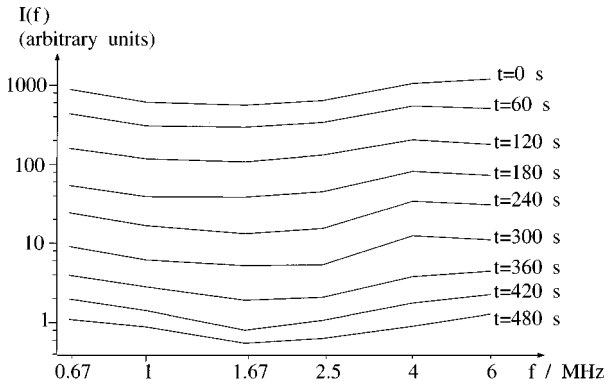


FIG. 7. Frequency response of the corrected backscattered intensity of dissoluting air bubbles in water.

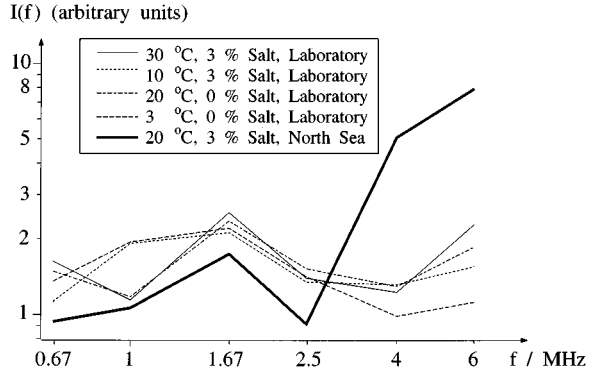


FIG. 8. Frequency response of the corrected backscattered intensity from air bubbles in water of different temperature and salt content in laboratory, and in the surf zone.

fixed intervals of time. Figure 7 shows the backscattered intensities as a function of the frequency, 0,60,...,480 s after the injection was stopped. It can be seen that the backscattered intensity decreases with time nearly constantly for all frequencies, for a concentration range of about 1:1000. Since small bubbles dissolve faster than large, the size distribution of the bubbles may tend to larger bubbles with time, so that a change of the frequency response of the intensity could have been expected, but was not seen. With salt water (3% NaCl), similar results are obtained, except that the bubbles dissolve more slowly.

(2) Figure 8 shows the backscattered intensity of air bubbles in laboratory, generated as before with a water jet, at different temperatures and at 0% and 3% salt content, respectively. Also shown is the backscattered intensity of a mixture of a large concentration of air bubbles and a small sediment concentration in the surf zone. The deviations of the backscattered intensity and thus of the backscattering cross section at different frequencies in the laboratory measurements—some 30% to 50%—are rather low compared with the absolute ratios of a factor of 100 and more between the backscattered intensities of suspended sediment of different radii (Fig. 9, calculated values using formulae presented in Refs. 8 and 10). With the measurement of air bubbles in the surf zone, the suspended sediment is presum-

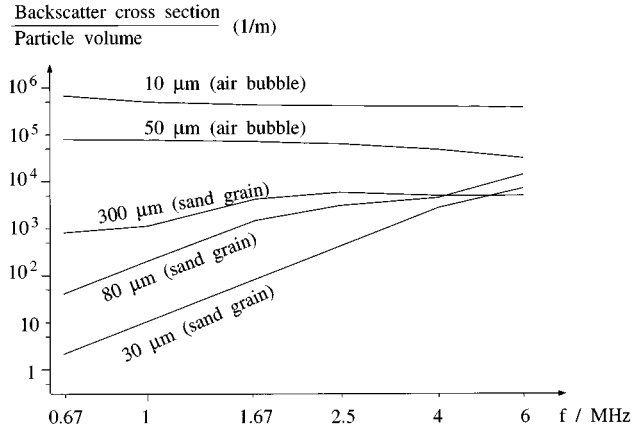


FIG. 9. Backscattering cross section (calculated) referred to particle volume of air bubbles and suspended sediment.

ably responsible for the higher backscattered intensity at 4 and 6 MHz.

The results of these two experiments confirm, that in the MHz range, the frequency response of the backscattering cross section of air bubbles in laboratory, and air bubbles in the surf zone, is approximately constant.

The backscatter cross sections at the different frequencies, $\sigma_a(f_0)\cdots\sigma_a(f_5)$ can thus be written as

$$\begin{aligned}&\sigma_a(f_0)=K_0C_a,\\&\sigma_a(f_1)=K_1C_a,\\&\cdots\\&\sigma_a(f_5)=K_5C_a,\\&K_0\approx K_1\approx\cdots\approx K_5,\end{aligned}\tag{11}$$

where $K_0\cdots K_5$ are constants and C_a is the concentration of air bubbles. This is characteristic for nonresonant air bubbles, so that in the following the air bubbles are assumed to be nonresonant.

Also acoustic *in-situ* measurements[11,12] show that the number of air bubbles per 1-μm radius step decreases below a bubble radius of approximately 30 μm. Therefore, there will be only few bubbles having the resonant radius at 0.67 to 6 MHz, namely 0.54 to 5 μm. Even though their acoustic cross section is about 10^4 times higher then their geometrical cross section most backscattered energy comes from nonresonant air bubbles.

For calculating the concentration of air bubbles from the given intensities, however, the radius distribution must be known. In the device described here, a_a is assumed to be an arbitrarily chosen value of 100 μm. Thus air bubble concentrations calculated with this device can only be compared among each other, they are not to be taken as absolute values. This is meaningless in this application, however, since only the sediment concentration is interesting and the air bubble concentration is only a by-product.

IV. NEAREST-NEIGHBOR ALGORITHM

The inversion of the signals—i.e., the transformation of the received intensities into concentration of sediment, concentration of air bubbles, and grain size—is done by a nearest-neighbor algorithm (NNA). This method of estimating is widely used in multichannel inversion.[13] For the calculation, the vector $\mathbf{I}:=(I(f_0,r),\ldots,I(f_5,r))$, denoted for simplicity as $\mathbf{I}(r)$, consists of the corrected backscattered intensities [Eq. (1)] of the six frequencies. Remember that the values $I(f,r)$ are a function of the scatterers in the considered range bin only.

Since natural sediments approximately conform to a log-normal size distribution,[5] the suspended sand is described by three parameters: Concentration, mean grain size, and standard deviation of the grain size. In air bubbles, however, the measurements described above show, that in the MHz range, the frequency response of the backscattering cross section of air bubbles in the surf zone is not different from those in laboratory, and it does not depend on the temperature and the salt content of the water. Therefore, the backscatter cross section of ''natural'' air bubbles can be described by their concentration and an arbitrarily chosen radius alone.

Thus every range bin of water containing air bubbles and suspended sediment has four unknown parameters:

(1) Concentration of sediment C_s,
(2) concentration of air bubbles C_a,
(3) average grain size $\overline{a_s}$, and
(4) standard deviation σ of the grain size.

They are combined in a parameter set $\underline{P}$ with $\underline{P}:=(C_s,C_a,\overline{a_s},\sigma)$.

In an off-line procedure, a large number of reference vectors $\mathbf{I}_{\text{ref}}(\underline{P}_k)$ are calculated [Eq. (2)] from the literature values for the backscattering cross sections of suspended sediment[8] and nonresonant air bubbles,[10] for different parameter sets $\underline{P}_k$, k is the number of the set. So, to each $\underline{P}_k$ set a $\mathbf{I}_{\text{ref}}(\underline{P}_k)$ vector is directly associated. Each of the parameters has typically 1–50 values, e.g., $\overline{a_s}=25,30,35\cdots500$ μm.

Each reference vector thus stands for a mixture of suspended sediment and air bubbles, and of course there are also reference vectors for pure suspended sediment and pure air bubbles. To determine the parameters of a measured vector $\mathbf{I}_{\text{meas}}$, from all reference vectors the reference vector $\mathbf{I}_{\text{ref}}(\underline{P}_k)$ is found, for which the Euclidean distance to $\mathbf{I}_{\text{meas}}$, namely $|\mathbf{I}_{\text{meas}}-\mathbf{I}_{\text{ref}}(\underline{P}_k)|$, is minimum. The parameter set $\underline{P}_k$ is then taken to be valid also for $\mathbf{I}_{\text{meas}}$ with a negligible error.

A disadvantage of this form of the NNA is the large amount of calculation time and storage capacity needed, since for each possible combination of sediment concentration, air bubble concentration, grain size, and standard deviation a reference vector must be generated, stored, and compared. To overcome this problem, a modified NNA was used and is presented here.

The normalized reference vectors $\mathbf{I}_{\text{ref,norm}}(\underline{P}_k)$ are defined by

$$\mathbf{I}_{\text{ref,norm}}(\underline{P}_k):=\frac{\mathbf{I}_{\text{ref}}(\underline{P}_k)}{|\mathbf{I}_{\text{ref}}(\underline{P}_k)|}.\tag{12}$$

Likewise, the normalized measured vector is defined by

$$\mathbf{I}_{\text{meas,norm}}:=\frac{\mathbf{I}_{\text{meas}}}{|\mathbf{I}_{\text{meas}}|}.\tag{13}$$

After $\mathbf{I}_{\text{meas}}$ is measured and normalized to $\mathbf{I}_{\text{meas,norm}}$, the ''nearest neighbor'' (defined by minimum Euclidean distance) $\mathbf{I}_{\text{ref,norm}}(\underline{P}_k)$ is found. Then can the parameters belonging to $\mathbf{I}_{\text{meas}}$, namely $C_{s,\text{meas}}$, $C_{a,\text{meas}}$, $\overline{a_{s,\text{meas}}}$, and σ_{meas}, be calculated from $\underline{P}_k=(C_{s,k},C_{a,k},\overline{a_{s,k}},\sigma_k)$:

$$\begin{aligned}&C_{s,\text{meas}}=\frac{|\mathbf{I}_{\text{meas}}|}{|\mathbf{I}_{\text{ref}}(\underline{P}_k)|}\,C_{s,k},\quad C_{a,\text{meas}}=\frac{|\mathbf{I}_{\text{meas}}|}{|\mathbf{I}_{\text{ref}}(\underline{P}_k)|}\,C_{a,k},\\&\overline{a_{s,\text{meas}}}=\overline{a_{s,k}},\qquad\qquad\sigma_{\text{meas}}=\sigma_k.\end{aligned}\tag{14}$$

By this normalization, the number of reference vectors needed is largely reduced, e.g., for air bubbles of every concentration only one single normalized reference vector is needed. A single normalized reference vector accounts for a given ratio of sediment concentration to air bubble concentration, independent from the absolute concentration. The ab-

TABLE I. Frequencies used.

Measuring frequency/MHz	0.67	1	1.67	2.5	4	6
Transducer's nominal frequency/MHz	1		2		6	
Radius of transducer/mm	10		10		3	
Half-beamwidth ($7^0\ a_t/\lambda)/^0$	1.54	1	0.6	0.42	0.88	0.58
Near-field length ($\pi a_t^2/\lambda$)/mm	140	210	349	523	75	113
Type (manufacturer: Karl Deutsch, Germany)	TS 20 WB 1		TS 20 WB 2		TS 6 WB 6	

solute concentration, however, is contained in $|\mathbf{I}_{\text{meas}}|$ and $|\mathbf{I}_{\text{ref}}(\underline{P}_k)|$, respectively.

An advantage of the NNA over other inversion algorithms is, that constraints can easily be applied, especially useful in this application is the constraint of a minimum and / or maximum average grain size $\overline{a}_{s,\min}$ and $\overline{a}_{s,\max}$. The minimum and the maximum grain size of the bottom sediment is known in most cases, so that also the grain size of the suspended matter must be within these limits. In this case, the NNA compares $\mathbf{I}_{\text{meas,norm}}$ only with those normalized reference vectors $\mathbf{I}_{\text{ref,norm}}(\underline{P}_k)$ for which $\overline{a}_{s,\min} \leqslant \overline{a}_{s,k} \leqslant \overline{a}_{s,\max}$ is true.

For the system described here, eleven ratios of sediment concentration / air bubble concentration (namely 10:0, 9:1, 8:2, 7:3, 6:4, 5:5, 4:6, 3:7, 2:8, 1:9, 0:10), 30 grain radii from 25 to 500 μm, and one standard deviation ($\sigma=1.2$) are used, which results in 301 reference vectors ($11\cdot30\cdot1=330$ could be expected, but note that the normalized reference vectors for pure air bubbles are identical for all grain radii). The time to calculate the parameters for one $\mathbf{I}_{\text{meas}}$ vector is as little as 4 ms on a Pentium-90 PC. The average error due to the remaining distance between $\mathbf{I}_{\text{meas,norm}}$ and $\mathbf{I}_{\text{ref,norm}}(\underline{P}_k)$ was calculated in detailed simulations. It is much smaller than other sources of error like configurational noise (the variation in backscatter intensity due to the arrangement of the scatterers within the range bin). Thus the number of 301 normalized reference vectors can be considered as sufficient, provided that σ is constant.

V. HARDWARE OF THE BACKSCATTER DEVICE

As described before, the water–sand–air bubble suspension in a range bin in the surf zone has four unknown parameters. These could be calculated from four known values, i.e., the backscattered intensities of four measuring frequencies. In the device described here, six frequencies are used instead of the bare minimum of four. The two extra frequencies serve to cater for a larger range of measurable grain sizes, and to improve the accuracy of the measurement. Table I shows the used frequencies and the parameters of the transducers. Each of the three transducers is driven with two frequencies.

Three pulse generators create pulses of the frequencies f_0 to f_5. To minimize errors due to fluctuations of the sediment concentration, the pulses are emitted in a $f_0|f_1|f_2|f_3|f_4|f_5\cdots f_0|f_1|f_2|f_3|f_4|f_5$ sequence instead of a $f_0|f_0|f_0|f_0\cdots f_5|f_5|f_5|f_5$ sequence. Three separate power amplifiers drive the ultrasonic transducers. The received signal from the transducers passes a preamplifier with a gain of 14 dB (0.67/1 MHz), 20 dB (1.67/2.5 MHz), and 40 dB (4/6 MHz), respectively. The different gains are adapted to the different sensitivities of the transducers. A narrow-band filter ($Q=6-60$ dependent on the frequency) reduces the noise, then the signal is selected by a multiplexer, and heterodyned down to an intermediate frequency of 200 kHz. The signal is full-wave rectified and smoothed by a tapped delay line that has a signal response equal to an integration of the last 14 μs, the time corresponding to the length of one range bin. A 12-bit A/D-converter finally digitizes the signal. These values are stored by microcontroller (Motorola MC68332, 16 MHz, 16-bit data width). They are time averaged after typically 50 pulses to the values $V(f,r)$ [Ref. Eq. (1)], which are transferred digitally at a rate of 80 kbit/s via a cable to the interface and the PC (Fig. 10). The circuits dynamic range of 60 dB permits the measurement of sediment concentrations from 10 mg/l to 20 g/l at any grain size.

To avoid interference of the preamplifiers, care is taken that all clock frequencies for the pulse generator, the A/D-converter, etc. are above the highest signal frequency, 6 MHz. Furthermore, the PC is completely galvanically isolated from the rest of the system by means of opto-couplers in order to avoid ground loops, and interference by spikes from the PC's ground.

VI. CALIBRATION

To obtain absolute values of the overall system sensitivity $S(f)$ for each frequency, a calibration setup similar to the one described by Sheng was used. As it is described extensively in his work,[8,14] it is only depicted briefly here. The setup uses a steel wire as a standard target, the radius is $a_w=0.1$ mm corresponding to a ka_w from 0.3 to 2.8. The length of the wire (40 cm) is much higher than the footprint of the transducers lobes (max. 6 cm) at the distance from transducer to wire ($r=50$ cm), so the backscattering cross section of the wire and then $S(f)$ can be calculated. These values for the system sensitivities were finally reduced by 10%–30% depending on the frequency, to fit the laboratory measurements of suspended sand.

VII. LABORATORY MEASUREMENTS

To determine the accuracy of the ASAP, a setup was built to generate a homogeneous suspension (Fig. 11): sand is injected from a container at a constant rate of 0.1–0.3 g/s, depending on the grain size, into a vertical, water-filled tube. It is mixed by an impeller, and distributes to a homogeneous suspension while sinking. During each measurement, different samples of 250 ml were taken, the sand was allowed to settle down and was dried and weighed. The root mean square deviation of the concentration of such samples taken

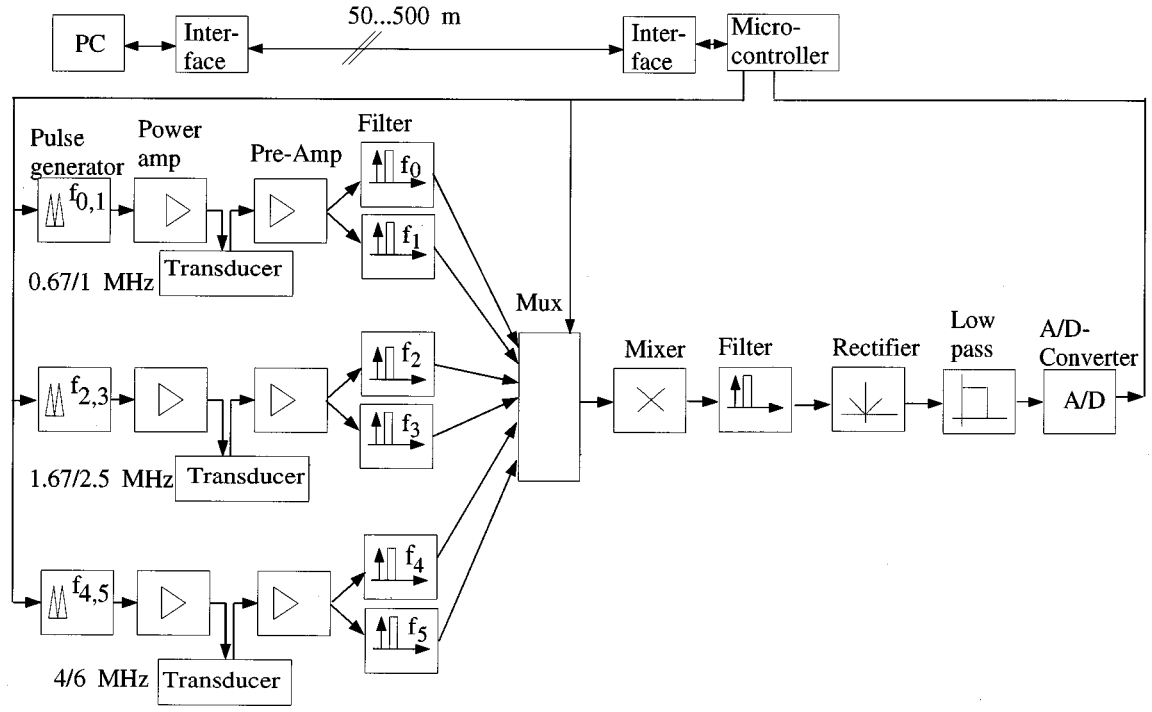


FIG. 10. Hardware of the ASAP.

at different times from the same suspension was lower then 5%. The concentration C_S ranged from 0.4 to 2 g/l, depending on the grain size and thus on the sinking velocity. At these concentrations, attenuation and multiple scattering can be neglected. Figure 12 shows the concentration measured by the ASAP versus the true concentration that is known from the samples, the horizontal line shows the desired behavior. The rms (root mean square) error is about 20%, with individual values up to $\pm 30\%$. Figure 13 shows the grain size measured by the ASAP versus true grain size. Here, the diagonal line shows the desired behavior. The rms deviation is about 15%.

It should be noted, however, that these relatively small errors are in some degree due to the benign laboratory conditions: low fluctuation of the sand concentration, absence of air bubbles, and a spatially homogeneous sand concentration. The error when measuring in the surf zone is higher, but the following comparison with optical sensors indicates that it does normally not exceed a factor of 2.

VIII. FIELD RESULTS

With the ASAP, a number of *in-situ* measurements were carried out near the island of Sylt, North Sea, Germany. The transducers were positioned in the surf zone, 35 cm above the ground. For comparison, two OBS (optical backscattering sensors, manufacturer D&A Instruments Co.) were mounted 10 and 25 cm below the transducer. The OBS were calibrated to the average grain diameter of the ground sediment, 325 μm. Also the wave height was monitored with a

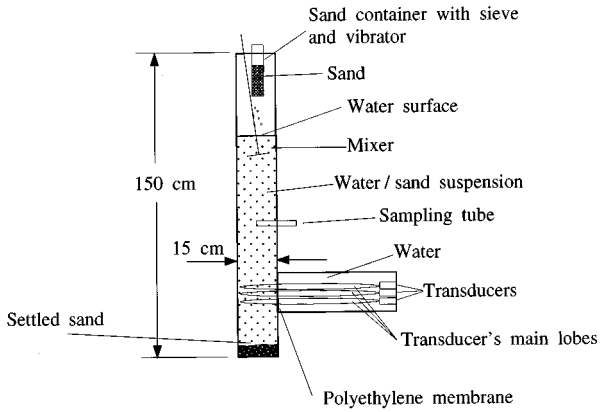


FIG. 11. Setup for generating homogeneous suspensions.

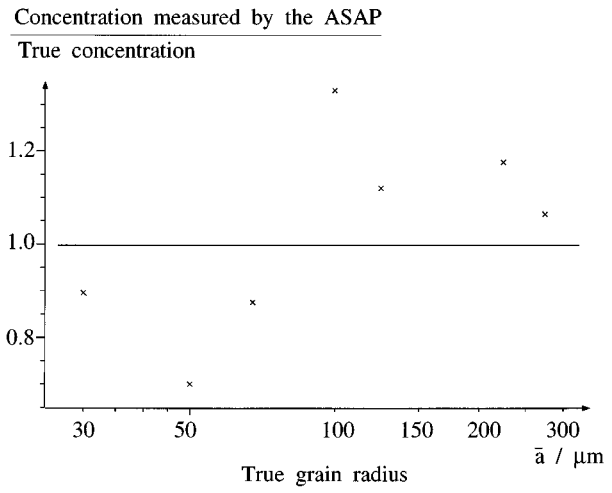


FIG. 12. Ratio of concentration measured by ASAP to true concentration versus grain radius.

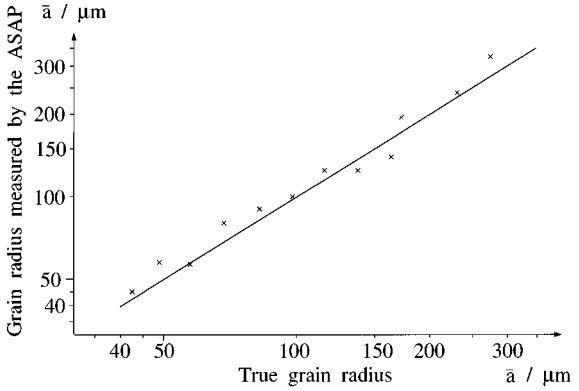


FIG. 13. Grain radius measured by ASAP versus true grain radius.

pressure gauge. Figure 14 shows the result of a 38-s measurement with some strong resuspension events, as well as the wave height above the bottom.

Figure 15 shows the concentration of the measurement from Fig. 14 at the height 10 cm above the bottom, as well as the calculated concentrations from the OBS signal mounted at the same height. Since the measured volumes of OBS and ASAP had a distance of some 20 cm, high short-time discrepancies occur, and the peaks of the suspension clouds have a time difference of some seconds.

For this measurement, Fig. 16 shows the time-averaged sand concentration versus height above bottom. The exponential behavior, which is also indicated by some theories of suspension of sediment in the surf zone,[15] can clearly be

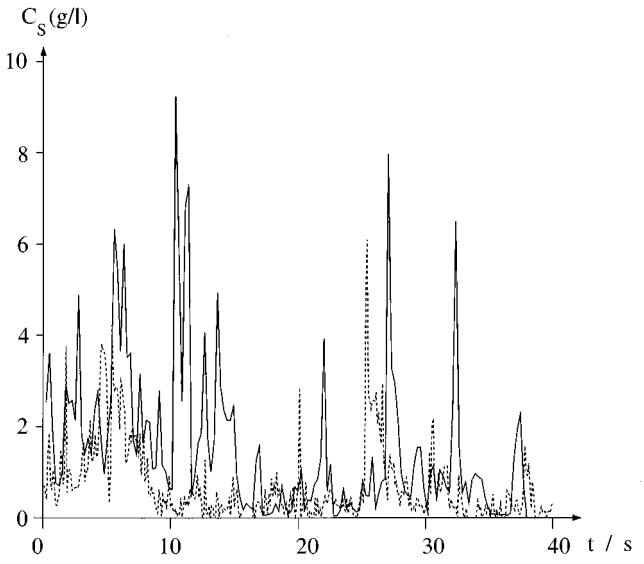


FIG. 15. Comparison of the concentrations measured by OBS (----) and ASAP (———).

seen. The two crosses mark the time-averaged concentrations of the two OBS. They are only about half as high as the ASAP results.

IX. SUMMARY AND CONCLUSIONS

Due to the different frequency response of suspended sand grains and air bubbles, the use of six frequencies per-

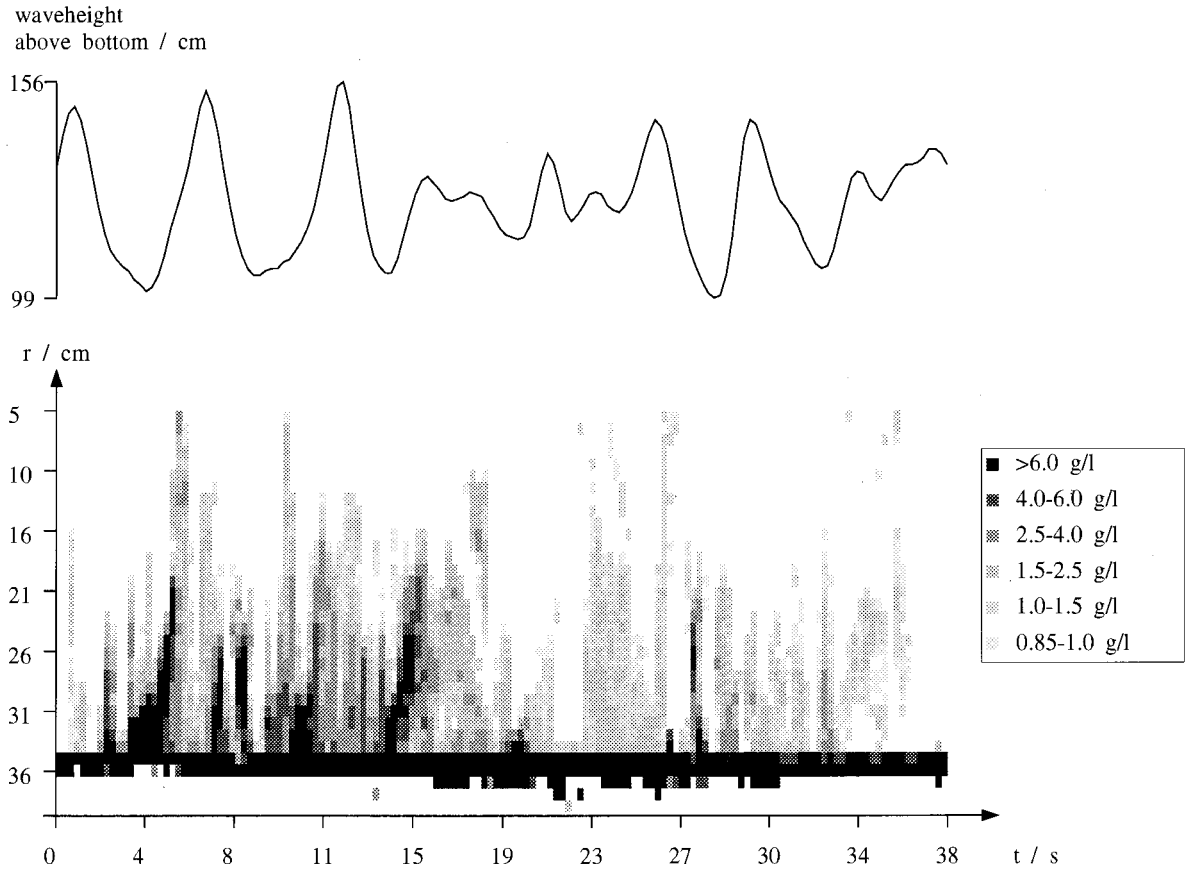


FIG. 14. Measurement with ASAP in the surf zone near Sylt.

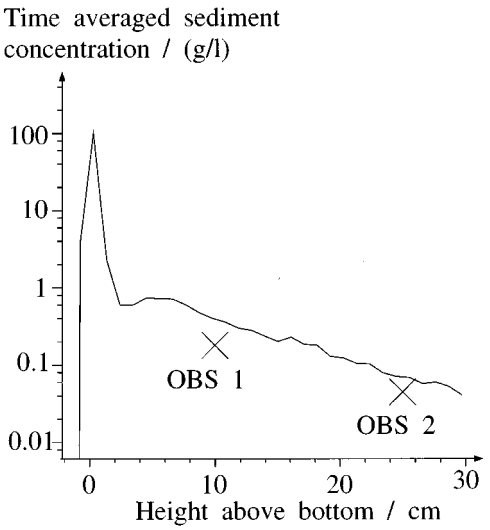


FIG. 16. Time-averaged sediment concentration versus height above bottom.

mits a precise measurement of concentration and grain size of suspended sediment in the presence of air bubbles. The part of the backscattered intensity due to air bubbles can be determined.

The near-field length extends to up to 500 mm, dependent on the transducer. Therefore, the transducer sensitivity to a homogeneous suspension was both calculated and measured. While the difference between the computed formula and a corresponding formula found in literature is quite small, higher discrepancies between the computed formula and results of measurements were observed, very likely due to the fact that the transducers are not ideal baffled pistons. The errors thus can be greatly reduced by assuming an ''acoustic'' radius smaller than the geometric radius.

During dissolution, air bubbles in fresh water and in salt water in laboratory exhibit almost the same frequency response. An alike frequency response is also seen in the surf zone. Air bubbles therefore can be described by their concentration alone and can be discriminated from suspended sediment, their concentrations can be separately calculated.

As a standard inversion algorithm, a modified nearest neighbor algorithm is used with good accuracy and execution time.

With sand of different grain sizes, the rms error of the concentration and grain size estimation in laboratory is about 20%. For *in-situ* measurements, higher errors occur due to air bubbles (whose influence can numerically be compensated, but due to measuring errors a residual error remains), absorption of ultrasound and multiple scattering. A comparison with two (also error-prone) optical backscattering sensors showed deviations of about a factor of 2.

Finally, it seems that the use of six frequencies seems to be a step toward higher accuracy in acoustic concentration profilers. The necessary amount of hardware and software is only a little higher than for a three-transducer, three-frequency system.

ACKNOWLEDGMENTS

The project was supported by the Deutsche Forschungsgemeinschaft (DFG) (Az. Be 1110/5-2). The ''Leichtweiss-Institut für Wasserbau,'' University of Braunschweig—mainly Professor Raudkivi, Dr. Dette, Hr. Thiel, and Dipl.-Ing. Peters—contributed support, discussions, and the use of the Sylt measuring facilities.

[1] C. Libicki and K. W. Bedford, ''The interpretation and evaluation of a 3-MHz acoustic backscatter device for measuring benthic boundary layer sediment dynamics,'' J. Acoust. Soc. Am. **85**, 1501–1511 (1989).
[2] P. D. Thorne, C. E. Vincent, P. J. Hardcastle, S. Rehman, and N. Pearson, ''Measuring suspended sediment concentrations using acoustic backscatter devices,'' Mar. Geol. **98**, 7–16 (1991).
[3] C. E. Vincent, D. M. Hanes, and A. J. Bowen, ''Acoustic measurements of suspended sand on the shoreface and the control of concentration by bed roughness,'' Mar. Geol. **96**, 1–18 (1991).
[4] F. R. Hess and K. W. Bedford, ''Acoustic backscatter system (ABSS); the instrument and some preliminary results,'' Mar. Geol. **66**, 357–379 (1985).
[5] A. E. Hay and J. Sheng, ''Vertical Profiles of Suspended Sand Concentration and Size from Multifrequency Acoustic Backscatter,'' J. Geophys. Res. **97** (C10), 15661–15677 (1992).
[6] A. E. Hay, ''Sound scattering from a particle-laden, turbulent jet,'' J. Acoust. Soc. Am. **90**, 2055–2074 (1991).
[7] A. M. Crawford and A. E. Hay, ''Determining suspended sand size and concentration from multifrequency acoustic backscatter,'' J. Acoust. Soc. Am. **94**, 3312–3324 (1993).
[8] J. Sheng, ''Remote Determination of Suspended Sediment Size and Concentration by Multi-frequency Acoustic Backscatter,'' Ph.D. thesis, Dept. of Physics, Memorial University of Newfoundland (1991).
[9] A. Downing, P. D. Thorne, and C. E. Vincent, ''Backscattering from a suspension in the near field of a piston transducer,'' J. Acoust. Soc. Am. **97**, 1614–1620 (1995).
[10] J. Wu, ''Bubble populations and spectra in near-surface ocean: summary and review of field measurements,'' J. Geophys. Res. **86**(C1), 457–463 (1981).
[11] S. Vagle and D. M. Farmer, ''The measurement of Bubble-Size Distributions by Acoustical Backscatter,'' J. Atmos. Ocean. Technol. **9**, 630–644 (1992).
[12] J. Wu, ''Bubbles in the Near-Surface Ocean: Their Various Structures,'' J. Phys. Oceanogr. **24**(9), 1955–1965 (1994).
[13] S. Kotz and N. L. Johnson, *Encyclopedia of Statistical Sciences* (Wiley, New York, 1986).
[14] J. Sheng and A. E. Hay, ''Spherical wave backscatter from straight cylinders: Thin wire standard targets,'' J. Acoust. Soc. Am. **94**, 2756–2766 (1993).
[15] R. Dette, ''Ein verlinfachtes Verfahren zur Ermittlung der Suspensionsfracht ausserhalb der Brandung,'' HANSA No. **23/24**, 1432–1438 (1991).

Acoustic characterization of an adult bubble injected into a fully developed turbulent flow field

Ali R. Kolaini and Alexei G. Goumilevski
National Center for Physical Acoustics, University of Mississippi, University, Mississippi 38677

(Received 28 February 1995; revised 31 December 1995; accepted 22 June 1996)

The acoustical characteristics of bubbles injected into a fully developed turbulent flow field are studied. By injecting an ''adult'' bubble into a flow, generated by a submerged axisymmetric horizontal water jet, the acoustic reexcitation of the bubble with and without breakup may occur in the shear-induced flow region. Bubbles of various sizes were introduced into jets of various speeds by means of interchangeable hypodermic needles. Results of the role of the turbulent flow characteristics in determining the acoustic bubble response are discussed. The characterizations of both the acoustical and the dynamics of the bubbles encountering a turbulent flow field depend upon the estimated integral and microlength scales, the corresponding Reynolds numbers, and the critical Weber numbers for both bubble distortion and breakup. A critical Weber number was estimated to be 0.52 for an ''adult'' bubble acoustic reexcitation without breakup and 1.10 for bubble breakup. A simple model is given to account for the bubble acoustic radiation. The Rayleigh–Plesset equation was altered by incorporating the turbulent pressure fluctuation of the flow as a driving force of the bubble. The turbulent energy spectrum obtained experimentally was used to predict the monopole acoustic emission by the bubble. Some conclusions are given about the nature of the turbulent flow field and conditions under which the bubble acoustic reexcitation may occur. [S0001-4966(97)05812-8]

PACS numbers: 43.30.Lz, 43.30.Nb [JHM]

INTRODUCTION

Much work has been done in the past on the dynamics of a single air bubble in water subject to various forms of excitation. However, the mechanisms of excitation of a single bubble, by a turbulent pressure field are not yet fully understood. In general, the pressure spectrum of a turbulent flow is relatively broad, and there is the possibility that large changes in bubble volume and shape may arise from the high-frequency pressure fluctuations surrounding the bubble.

Some of the earliest studies of the noise in the ocean were published in 1948 by Knudsen *et al.*[1] Since then, much progress has been made toward the understanding of ocean ambient noise. We now know that many sources contribute to ocean ambient noise. Some of these bubble-related mechanisms are splash or rain noise (Prosperetti *et al.*[2] and Pumphery *et al.*[3]), hail and snow (Crum *et al.*[4]), breaking waves (Medwin and Daniel,[5] Farmers and Ding,[6] Loewen and Melville,[7] Kolaini and Crum[8]), and capillary waves (Longuet-Higgins,[9] Kolaini *et al.*[10]). The noise from these sources is all bubble-related.

In the ocean, bubbles can be deformed by turbulent flow fields created by breaking waves, motion of marine vehicles, and Langmuir circulation. An important scientific question is what role the turbulent pressure fluctuation plays on monopole acoustic radiation of bubbles. The linear theory predicts that, the most efficient mechanism that a bubble has for the emission of sound into the far field is through volume pulsations where the higher modes (surface modes) of oscillations are volume preserving and thus radiate inefficiently. Following Lamb,[11] the shape oscillations of a gas bubble of equilibrium radius R_0 in an unbound liquid of density ρ, in which the radial displacement is described by a spherical harmonic of degree n, the radian frequency ω_n is given by

$$\omega_n^2=(n-1)(n+1)(n+2)(T/\rho R_0^3), \tag{1}$$

where T denotes the surface tension, and n is an integer number. On the other hand there is a purely radial oscillation whose radian frequency is given by the well-known equation,[12]

$$\omega_0^2=\frac{3\gamma P_\infty}{\rho R_0^2}+\frac{2(3\gamma-1)T}{\rho R_0^3}, \tag{2}$$

where P_∞ is the equilibrium gas pressure and γ is the ratio of the gas specific heats and equal to 1.4 for adiabatic processes. Longuet–Higgins[13] showed that there is a resonance, at second order, between nth normal mode, with frequency ω_n, being equals to $1/2\omega_0$, the frequency of the breathing mode. He incorporated the damping effect and found the stimulation of the breathing mode arises mainly during the first few cycles of oscillation, and is not due to any long-term exchange of energy between the modes. In a series of papers, Longuet-Higgins[14,15] pointed out due to the peculiarity of the damping characteristics of the bubble, the sound emissions from the initial distortion would appear to have approximately the frequency of the breathing mode rather than twice the frequency of a distortion mode. Longuet-Higgins[13] predicted that the distortion modes could contribute to the far-field sound radiation by producing a monopole radiation of sound at a second-order of magnitude. Ffowcs-Williams and Guo[16] used the method of multiple scales to examine the problem. They acknowledged that a transfer of energy from the surface mode to the volume mode occurs at the exact resonance, and that a periodic exchange will occur near resonance. However, they believe that the amount of energy exchanged is so small that no significant sound radiation occurs. The acoustic reexcitation of a bubble by turbulent flow

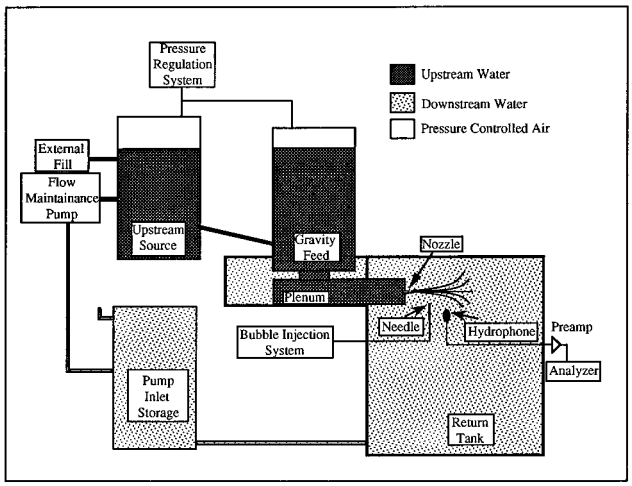


FIG. 1. Schematic drawing of the experimental arrangement (not drawn to scale).

fields was ruled out by Crighton and Ffowcs-Williams.[17] They showed that the effective turbulent length scales are much smaller than the bubble radius to cause coherent forcing over the bubble entire surface.

More recently Mao *et al.*[18] have shown experimentally that the nonlinear coupling between the surface and volume modes of an oscillating bubble is possible at low-ambient pressures (i.e., a few cm Hg) where the condition $\omega_0=2\omega_n$ was met. However, they were unable to detect the surface modes which can be in resonance with the volume mode at atmospheric pressure.

The highly chaotic motion of the turbulent flow field in the ocean left behind by breakers may distort and deform the ''adult'' bubbles and subsequently radiate sound. Recently, the effect of the bubble relative motion on its acoustic emissions in a flow field have been studied.[19,20] It has been shown that the nonstationary motion of a bubble in a quiescent liquid initiates its volume oscillations.[20]

This study was designed to investigate experimentally and theoretically the role of the turbulent flow field on the acoustical behavior of an ''adult'' bubble. In this study, we have chosen a submerged water jet as a host turbulent flow field where its fluid dynamics have been extensively investigated and reported in the literature.

I. EXPERIMENTAL PROCEDURES

The experimental setup to examine the sound production mechanisms in a turbulent flow containing bubbles consisted of a Plexiglas® laboratory tank with dimensions of 2.1 m×2.1 m×1.8 m (width, length, and height). A simple system that enables us to introduce gas bubbles into a turbulent region was developed and positioned in the trough section of the tank attached to it (Fig. 1). A bubble was injected into the fully developed region of the turbulent jet produced by a submerged nozzle at $(x/D)\sim 10$, where x is the distance along the jet and D is the nozzle diameter. The injected bubble was acoustically ''inactive'' by the time it entered the turbulent portion of the flow. The turbulent jet was produced by the nozzle that expels water from a plenary volume into a return tank. The pressure difference between the plenum and the return thank was maintained by a gravity feed tank mounted above the plenum. A steady-state flow was achieved using a flow maintenance pump that transferred water from the return tank to the gravity feed tank. The pump inlet storage tank and the upstream source tank serve to isolate the return tank and the plenum region from sound produced by the pump. This isolation reduced the background noise in the return tank where acoustic signals were acquired. From the plenum interchangeable nozzles were used to create turbulent jets of various intensities. The jet speed at the nozzle's exit ranged from 1.03 to 3.0 m/s and were accurately controlled by a Cole–Palmer flow meter. Three needle sizes: 0.19, 0.39, and 1.03 mm were used to create various size bubbles. Air was supplied to the needles utilizing a Cole–Palmer 3418 syringe pump. The pump speed was selected to produce a bubble per second. The needle was positioned on a holder and was placed below the jet boundary. A B&K 8103 hydrophone suspended in the tank from an aluminum I-beam was used to receive the acoustic emissions of the bubble. The B&K hydrophone high-passed at 200 Hz using a Stewart VBF8 filter was connected to a digitizing LeCroy 9400A oscilloscope via a B&K type 2635 charge amplifier and was digitized with sampling rate of 40 k/s. The waveforms were then downloaded to a Macintosh II computer for storage and further analysis.

The knowledge of the flow field is of paramount importance for an understanding of the sound mechanisms in a turbulent flow. Before injecting the bubble into the flow field, the distribution of mean velocity, mean shear, and turbulent intensities were measured utilizing calibrated hot wire anemometry. The concurrent bubble deformation and the aforementioned turbulent flow parameters are being measured using stereoscopic particle tracking velocimetry. This technique would provide the nearly instantaneous spatial velocity field measurements.[21]

II. RESULTS

A. Acoustic emissions

Using high-speed cinematography we were able to correlate the bubble dynamics in the turbulent field and its acoustic emissions. A Photec IV rotating prism camera with speed of 1000 frames/s was used to film the turbulent flow field containing an adult bubble; at the same time, the sounds made by the bubble were detected by a hydrophone, amplified, and displayed on an oscilloscope. The oscilloscope screen was photographed directly onto the movie film via an auxiliary lens on the camera. The film recorded a continuous oscilloscope trace and also a series of images of the flow. The flow region was brightly lit through the back of the tank using studio lights. Figure 2(A) shows a series of high-speed photographs of a turbulent jet with an exit velocity of 2.2 m/s ($\mathrm{Re}=U_jD/\nu\approx 7300$) where a bubble of 1.1 mm in radius was injected into the fully developed region of the flow ($x/D\sim 10$). The bubble was formed by forcing air through a hypodermic needle (not shown) and was allowed to rise through still water, before encountering a horizontal turbulent jet flowing from right to left. Frames (a)–(e) show the

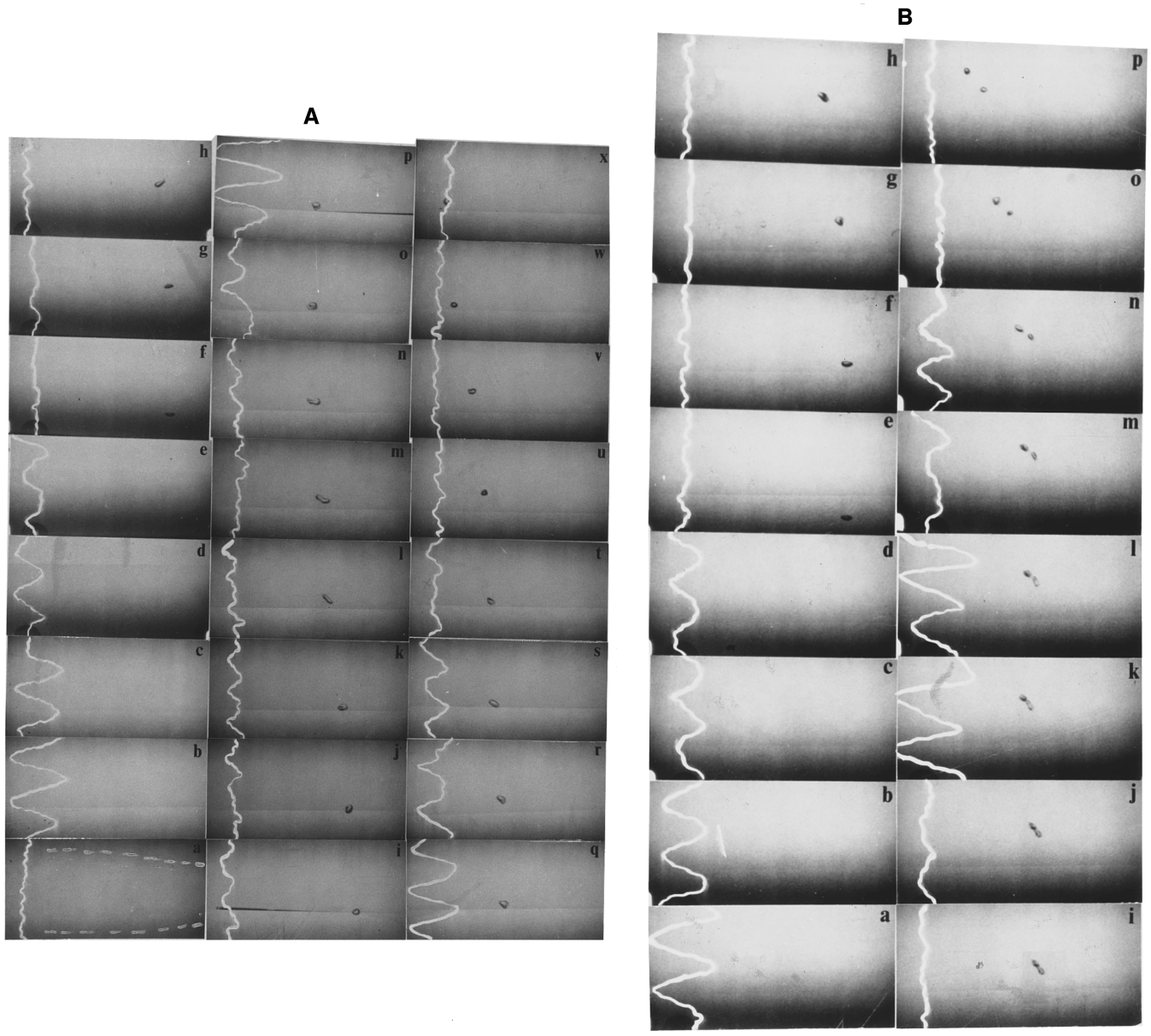


FIG. 2. (A) Photographs of a sequence of frames from high-speed movie depicting the bubble encountering with a horizontal turbulent jet. The original film was made at a speed of 1000 frames/s. The bubble was formed at a nozzle (not shown in the image) and then allowed to rise through still water. The thin line on the left-hand side of the images is an oscilloscope trace showing the radiated acoustic pressure generated by the bubble. In frames (a–e), the bubble detaches from the nozzle and oscillates at its resonance frequency of approximately 3.1 kHz. The bubble is floating upward and encountering the turbulent jet which is flowing from left to right. In frames (g–n), the bubble experiences the influence of the jet and deforms. The reexcitation of the adult bubble occurs at frames (o–s) and 3.1-kHz noise is emitted without bubble breakup. Time intervals for each frame are as follows: (a–e) each increment 1 ms, (f) 76 ms, (g) 95 ms, (u) 97 ms, (i) 105 ms, (j) 111 ms, (k) 116 ms, (l) 124 ms, (m) 127 ms, (n) 129 ms, (o–s) each increment 1 ms, (t) 138 ms, (u) 143 ms, (v) 148 ms, and (w) 155 ms. The jet velocity at the nozzle is 2.2 m/s and nozzle inner diameter is 3.35 mm. The white marks depicting the oscilloscope traces were high lighted for better visualization. The jet boundary is identified by dashed lines in image a. The viewing area is (55 cm×30 cm). (B) The same as part (A) except jet velocity has increased to 2.5 m/s at the nozzle exit. Frames (a–d) show the pinch-off noise (3.1 kHz). The bubble encounters the turbulent jet [frames (e)–(o)] and breaks into two smaller bubbles each with frequencies of 3.5 and 4.5 kHz [frames (k–n)]. Time intervals for each frames are as follows: (a–d) each increment 1 ms, (e) 80 ms, (f) 99 ms, (g) 117 ms, (h) 121 ms, (i–n) 125–130 ms, (o) 135 ms, and (p) 140 ms.

acoustic emissions from bubble pinching off the needle. Upon entering the jet (frame f, 76 ms), the bubble experiences deformation due to the shear rate of the jet velocity. The deformation continues with no significant acoustic radiation until frame (o). This frame and the subsequent ones (o–t) show reexcitation and sound radiation from the bubble without breakup. The ''adult'' bubble acoustic reexcitation is due to the turbulent pressure fluctuations acting on the surface of the bubble. In part C of this section we will discuss a simple theory that predicts the acoustic radiation of the ''adult'' bubble upon encountering the highly shear-induced region of the flow. Figure 2(B) shows a typical example of a bubble (1.1 mm) encountering a turbulent jet with an exit velocity of 2.5 m/s. The pinch-off noise is shown in frames (a–d). The bubble enters the jet at frame e (80 ms). In later frames the bubble is distorted into a long filament. This filament breaks into two bubbles and produces sound (frames k–n). Two frequencies were emitted: one of approximately 2.5 kHz and the other 15 kHz.

These photographs illustrate that the turbulent jet can

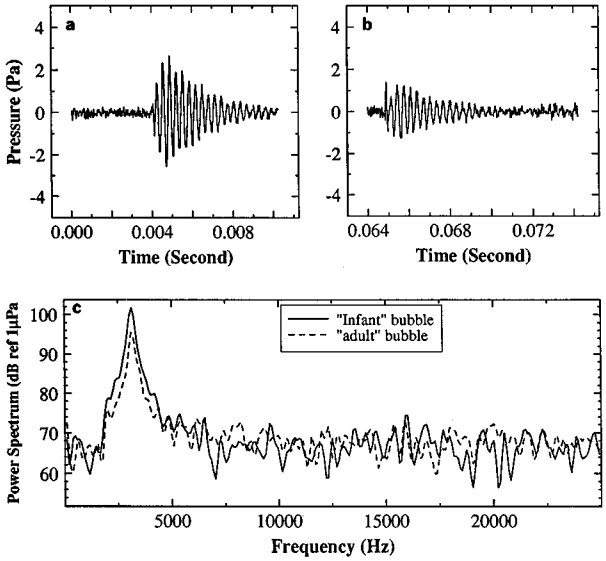


FIG. 3. The pressure time traces of a bubble of 1.1 mm in radius entering a turbulent jet with an exit velocity of 2.2 m/s. (a) shows the pinch-off noise, (b) is the signal when the bubble encounters turbulent fluctuations in the jet, and (c) is the corresponding power spectra.

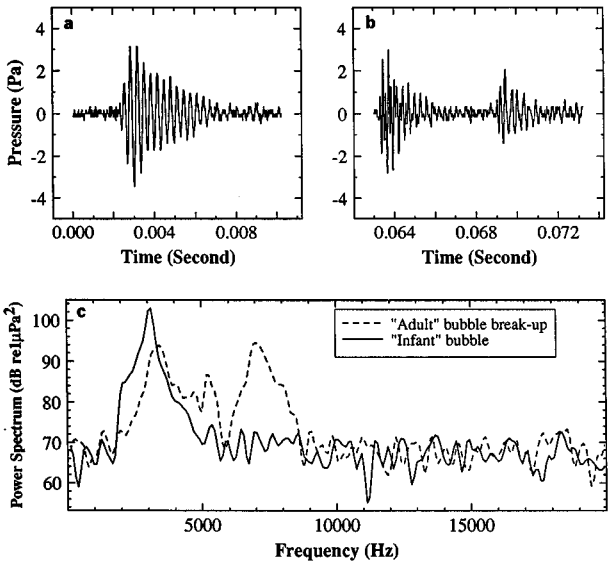


FIG. 5. The same as Fig. 3, except this time the bubble has broken into three smaller bubbles with resonance frequencies of 3.5, 5.2, and 16 kHz. The jet velocity is about 2.5 m/s.

indeed lead to ''adult'' bubble acoustic reradiation and break-up. These mechanisms are further illustrated in Fig. 3 where a bubble of 1.1 mm in radius was allowed to encounter with a turbulent jet produced by a nozzle of 3.35 mm in diameter with an exit velocity of 2.2 m/s. Figure 3(a) shows the acoustic emissions from bubble detaching the needle followed by the sound pressure of the bubble encountering the jet [Fig. 3(b)]. The power spectra of these two signals [Fig. 3(c)] show the monopole acoustic excitation of the ''adult'' bubble caused by turbulent fluctuations acting on the bubble

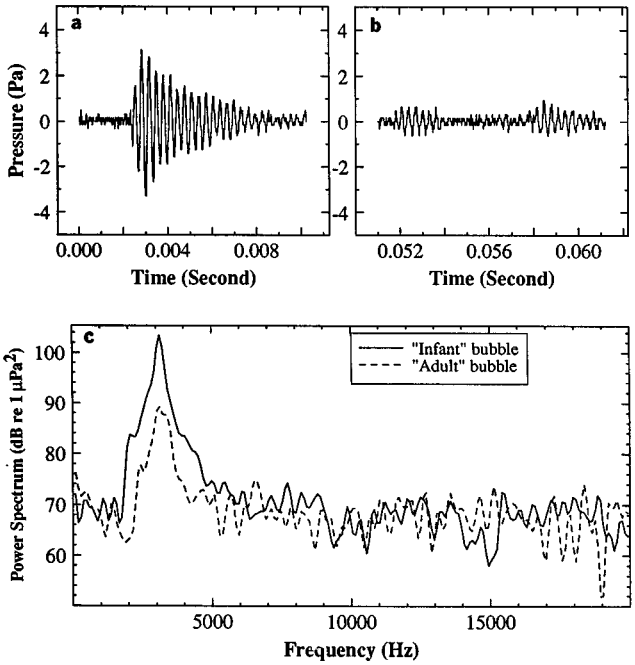


FIG. 4. The same as Fig. 3, except the multiple exchange of energy has occurred between the turbulent flow field and the bubble.

wall. The exchange of energy between turbulent flow field and the bubble dynamics which leads to acoustic emissions could occur periodically if the bubble frequency coincide with the resonance frequency of eddies [see Fig. 4(a)–(c)]. Finally, Fig. 5(a)–(c) illustrates an example of the bubble encountering a turbulent jet with the exit velocity of 2.5 m/s where the adult bubble of 1.1 mm breaks into at least three smaller bubbles with the resonance frequencies of 3.4, 5.2, and 9.4 kHz. The noise resulted from bubble splitting was hypothesized to be related to an instability rather than to the motion driven by turbulent fluctuations.[19] The high-speed photographs shown in Fig. 2(a) and (b) and pressure-time traces shown in Figs. 3–5, serve to illustrate the mechanisms in which turbulent pressure fluctuation may cause a bubble to breakup and/or reradiate sound without breakup.

The bubble breakup and/or deformation and the subsequent acoustic radiation depend upon the size of the bubble, jet velocity, frequency and intensity of the velocity fluctuations. It is the instantaneous vortical structures and intermittent nature of the flow that causes bubbles to be reexcited. Figure 6(a)–(f) illustrates histograms of 100 events each where various bubble radii were released from three different needles that encountered various jet velocities. These events are characterized into three regions; (1) bubble breakup, (2) bubble acoustic reradiation, and (3) no acoustic radiation. These regions are labeled by 1, 2, 3, on the horizontal axes, respectively. The histograms shown in Fig. 6(a)–(c) are for the same size nozzle (3.35 mm in diameter) with jet speeds of 1.88, 2.13, 2.38 m/s, respectively. In each histogram three needle sizes were used. As the bubble size increases, the percentage of bubble breaking up was increased. The bubble break up was also enhanced by increasing the jet velocities. The histograms shown in Fig. 6(d)–(f), are for a fixed bubble size with nozzle sizes being 2.28, 3.35, 3.81 mm, respectively. In these histograms the jet velocities were varied form 1.03 to 2.38 m/s. Majority of the 1.1-mm bubbles shown in

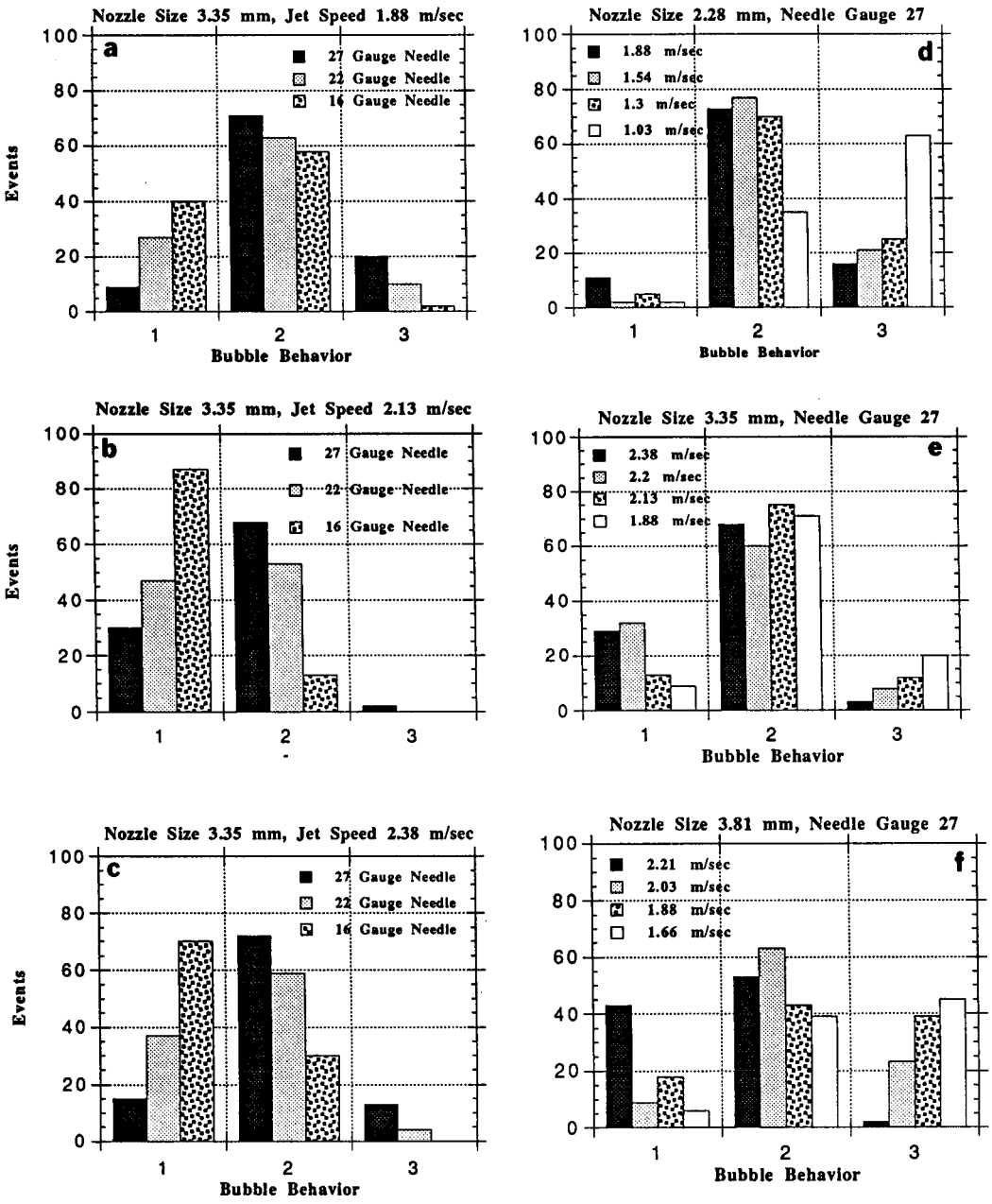


FIG. 6. (a)–(c) Histograms of each with 100 events of the acoustical behavior of bubbles with nozzle diameter of 3.35 mm encountering various jet velocities and bubble sizes, and (d–f) histograms of bubbles with a 1.1-mm size encountering various jet velocities and nozzle diameters.

Fig. 6(d)–(f) were reexcited indicating the importance of the acoustic reradiation of bubbles encountering the turbulent flow field. These histograms illustrate the change in bubble acoustical behavior as the jet speeds, bubble radii, and nozzle sizes were altered.

B. Bubble deformation

Prediction of the bubble deformation and breakup in a turbulent flow field is a formidable problem, primarily because of the inherent theoretical and experimental difficulties in treating turbulent two-phase flows. The detailed quantitative characterization of bubble deformation due to the interaction with high Reynolds number flow fields is an ongoing research. In this section we discuss the effect of the overall velocity field on bubble deformation and subsequent acoustic emissions utilizing the hot wire anemometry (HWA) and particle tracking velocimetry (PTV) velocity field measurement techniques. The HWA technique provides long-time temporal averaging of the fluctuating velocity components while the PTV provides the instantaneous spatial information of the flow.

Before injecting bubbles into the turbulent flow field, the distribution of mean velocity, mean shear, turbulent intensities, and length scales of the water jet were measured at a distances of $x/D \sim 10$, 20, 30, and 40 by means of 2-D hot film anemometers. At 10 diameters from the nozzle, the flow

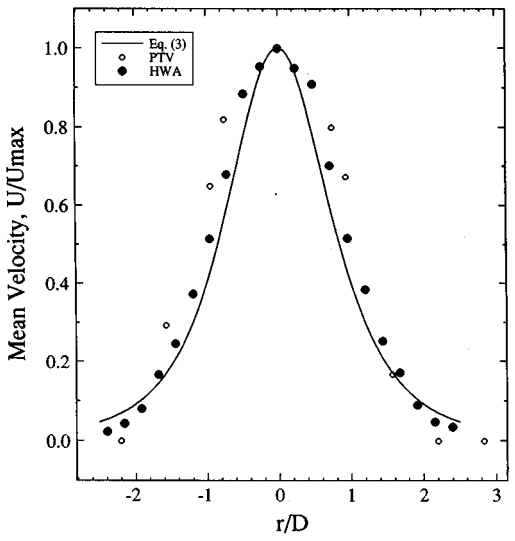


FIG. 7. The horizontal mean velocity distribution of the turbulent jet measured at $x/D\sim10$ using the HFA and PTV techniques and are compared with theory.

is in a fully developed turbulent region. The mean horizontal velocity distribution in the jet is shown in Fig. 7. Görtler[22] developed a theory for round submerged jets with velocity profile of

$$\frac{U}{U_{\max}}=\left(1+\frac{\xi^2}{4}\right)^{-2}, \tag{3}$$

where $\xi=15.2r/x$, and $U_{\max}=7.4(J/\rho)^{1/2}/x$. The term J is the jet momentum, ξ is the similarity variable, $U_{\max}$ is the maximum velocity at the center of the jet, and x is the distance from the nozzle. A comparison of the mean horizontal velocity measurement shows a very good agreement with Eq.

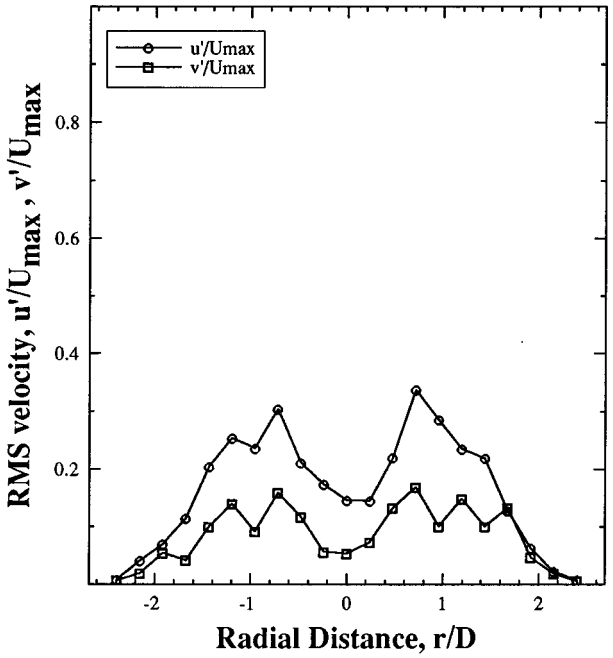


FIG. 8. The turbulent fluctuating velocities $\sqrt{u'^2}$ and $\sqrt{v'^2}$ measured at $x/D\sim10$ from the jet.

(3). The radial distribution of turbulent fluctuating velocities are shown in Fig. 8. The turbulent intensity has a maximum value at a region where shear stress is maximum. This region of the flow field plays an important role in the adult bubble deformation and subsequent acoustic radiation. In a fully developed turbulent flow field the lateral Taylor microscale can be estimated to be[23]

$$\lambda_g\cong\sqrt{2\overline{u'^2}\Big/\overline{\left(\frac{\partial u'}{\partial r}\right)^2}}. \tag{4}$$

The value of λ_g measured at $x/D\sim10$ and $r\cong D$ is about 0.23 cm. The Reynolds turbulence number, $\mathrm{Re}_{\lambda_g}=\sqrt{u'^2}\lambda_g/\nu$, based on the root-mean-square turbulent velocity and lateral microscale, λ_g, is equal to about 700. The Kolmogorov microscale η is therefore given by

$$\eta=\lambda_g(15)^{-1/4}(\mathrm{Re}_\lambda)^{-1/2}, \tag{5}$$

and equals to 4.4×10^{-3} cm. The life-time of the turbulent Kolmogorov eddies can be estimated to be around 147 μs using $\tau_\eta=\eta/\sqrt{u'^2}$. The lifetime of the Taylor microscale, $\tau_{\lambda_g}=\lambda_g/\sqrt{u'^2}$ is estimated to be about 5.1 ms. The lifetime of the Taylor microscale is sufficient enough to excite a bubble with radius of 1.1 mm. The eddies with length scales in the range of the Taylor microscale may be responsible for bubble deformation and monopole acoustic radiation. The turbulent vortices, in general, have three-dimensional structures and are characterized by their core diameter and the vortex diameter. The Taylor microscale, which is responsible for the turbulent shear, is characterized by the vortex diameter. The analysis of splitting of drops and bubbles due to the capillary and dynamic pressures were made by Hinze.[23] The fate of a bubble depends on two nondimensional ratios, the Weber number defined as

$$\mathrm{We}=\frac{\rho\overline{v^2}}{T/2R_0}, \tag{6}$$

and ratio of length scales $2R_0/\eta$, where T is the surface tension in dyn/cm. For our case $2R_0\gg\eta$, and for high Reynolds numbers the wave numbers of the energetic eddies and the dissipation range are widely separated. Under these circumstances, Kolmogorov's universal equilibrium theory postulates an inertial subrange at the lower end of the equilibrium range where energy transfer through the spectrum is independent of viscosity and, therefore,

$$\overline{v^2}=2.0(2\epsilon R_0)^{2/3}, \tag{7}$$

where ε is the energy dissipation per unit mass and time. Under these conditions Eq. (6) will have a new form,

$$(\mathrm{We})_{\mathrm{critical}}=6.35\left(\frac{\rho}{T}\,\epsilon^{2/3}R_0^{5/3}\right). \tag{8}$$

Since the small-scale structure of turbulence at a large Reynolds number is always approximately isotropic (Fig. 8), therefore,

$$\epsilon=15\nu\,\frac{\overline{u'^2}}{\lambda_g^2}, \tag{9}$$

where λ_g is defined by Eq. (4). To obtain the lowest Weber number at which bubble deformation and subsequent acoustic radiation will occur, we use the maximum Reynolds shear stress $\rho\overline{u'^2}$ in a region where the monopole acoustic radiation of the bubble may have occurred. The resulting critical Weber number is about 0.52. The critical Weber number at which bubble breakup will occur is about 1.10 and is slightly smaller than the value of 1.20 predicted by Sevik and Park.[24] It should be noted that these two Weber numbers, 0.51 and 1.10, were calculated based upon the values of the high shear-induced region of the jet.

C. Theoretical considerations

In this section a simple model is developed to account for the monopole acoustic radiation of a spherical ''adult'' bubble encountering the turbulent flow field. To account for the bubble acoustic reradiation and the turbulent velocity coupling several assumptions are made. These assumptions are, the bubble surface displacement is small compared to its equilibrium radius, the turbulent scale in the inertial subrange ($-5/3$ power) is larger than bubble dimensions, and the flow field around the bubble is locally incompressible and irrotational. The shear-induced bubble deformation increases as Reynolds number becomes larger than 400 with increasing Weber number.[25] We will further assume the bubble size is small enough so that its equilibrium shape can be considered as a sphere. The second assumption can be justified by considering values of the turbulent scale L (turbulent integral or Taylor microscale) and the bubble radius R_0. For bubbles with diameters larger than L, vorticities apparently disturb their surface locally, causing their shape to oscillate and possibly breaking it into filaments. The third assumption can be justified for bubbles with dimensions much larger than the viscous boundary layer thickness, $\sqrt{\nu/\omega}$, formed on the bubble's surface (see Ref. 26).

Relative velocity of the bubble's center of mass can be obtained by considering the Newton's equations,[27]

$$\frac{dV}{dt}=2\left(\frac{dU}{dt}-g\right)-\omega^*\left(1+\frac{3\,\mathrm{Re}}{16}\right)^{1/2}V, \tag{10}$$

where U is the fluid velocity, $\omega^*=9\nu R_0^{-2}$ is the characteristic frequency of the viscous boundary layer around the bubble, and $\mathrm{Re}=2VR_0/\nu$. Equation (10) states that the bubble experiences relative acceleration due to the pressure gradient that a fluid element would experience in the absence of the bubble, buoyancy force, and Stokes drag force modified for large Reynolds numbers. The bubble experiences pressure forces that are three times larger than the pressure acting on the fluid element alone. For relatively slow fluid velocities changing periodically with time, the Eq. (10) has an analytical solution of the form,

$$V(t)=V_0e^{-\omega^*t}+\frac{2U_0\omega}{\omega-i\omega^*}\,(e^{-i\omega t}-e^{-\omega^*t})-\frac{2g}{\omega^*}\,(1-e^{-\omega^*t}). \tag{11}$$

For low-frequency velocity fluctuations ($\omega\ll\omega^*$), the bubble moves with fluid in the horizontal direction while rising up with the ascent velocity. On the other hand, in a flow with high-frequency fluctuations ($\omega\gg\omega^*$), the second term in the right-hand side of Eq. (10) is small compare to first term with gravity being zero. Under these conditions the bubble slip velocity is twice the horizontal velocity of the host fluid. In a turbulent flow with broadband frequency energy spectrum, the bubble's relative velocity may be twice the sum of the velocity of the turbulent fluctuations and the rising velocity of the bubble. These conclusions can be considered as qualitative only because the assumption of the small Reynolds number is for an ideal case. The Reynolds number based on the bubble rising velocity is about 400 for a bubble with a radius of 1.1 mm. For larger Reynolds numbers, Eq. (10) is not linear and the superposition principle may not be appropriate to apply.

We consider oscillation of the bubble embedded in a liquid and follow Longuet–Higgins approach.[13–15] For simplicity we choose spherical coordinate system associated with the bubble center. The velocity of the liquid can be described by the velocity potential,

$$\tilde{\Phi}=-V(t)r\left(1+\frac{R_0^3}{2r^3}\right)+\Phi, \tag{12}$$

where the first term represents the velocity potential of a sphere moving with relative velocity $V(t)$ given by Eq. (10) and the second term represents potential function associated with the bubble acoustic radiation. The general procedure to derive equations for bubble oscillation is to satisfy the kinematics and dynamic boundary conditions. For relatively small amplitude bubble oscillations ζ, these equations can be linearized,

$$\frac{\partial\zeta}{\partial t}=\frac{\partial\Phi}{\partial r}\quad\text{and}\quad P_b-P_f=\frac{T}{R_0}\left(2-2\frac{\zeta}{R_0}-\frac{\nabla_s^2(\zeta)}{R_0}\right), \tag{13}$$

where $\nabla_s^2=(1/\sin\theta)(\partial/\partial\theta)\sin\theta(\partial/\partial\theta)$. Keeping the second-order terms in Eq. (13) reveal a possible mechanism in which volume pulsation could be coupled to the surface mode,[13–15] however this coupling has been reported to be a weak acoustic radiator.[16]

If we assume the bubble is expanding adiabatically and apply the Bernoulli integral equation for the pressure outside the bubble we obtain,

$$\frac{d^2\zeta_n}{dt^2}+\omega_n^2\zeta_n=\frac{(n+1)V^2}{16R_0}\int_0^{\pi}P_n(\cos(\theta))\times\sin(\theta)(5-9\cos^2\theta)\,d\theta, \tag{14}$$

where ζ_n is the spherical amplitude of bubble oscillations, ω_n is the bubble natural frequency ($n=0$) and surface mode frequencies ($n\geqslant 2$) and $P_n(\cos(\theta))$ is the Legendre polynomials. The driving force for the bubble oscillation given by Eq. (14) is the dynamic pressure which contributes to the bubble volume and second surface mode excitations. We will further consider the most energetic monopole radiation of the bubble. Introducing the total damping term, δ which is consisted of the dissipation due to the viscosity of the liquid,

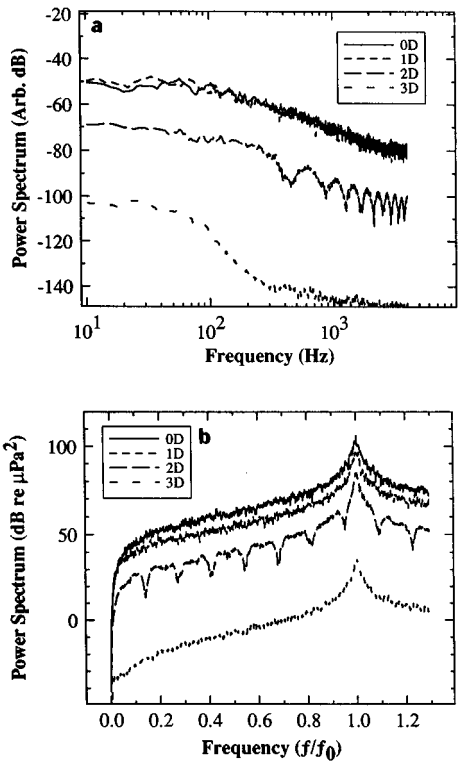


FIG. 9. (a) The turbulent fluctuation power spectra measured at 0D, 1D, 2D, and 3D for a jet velocity of 2.2 m/s at 10D from the jet, and (b) the estimated acoustic radiated energy of the bubble with a size of 1.1 mm in radius encountering the jet with turbulent spectra shown in part (a) as inputs.

thermal, and acoustic radiation the Rayleigh–Plesset[28] equation can be modified to give the nondimensional displacement of the bubble surface ($\zeta=\zeta_0$),

$$\frac{d^2\zeta}{dt^2}+2\delta\frac{d\zeta}{dt}+\omega_0^2\zeta=\frac{1}{4R_0}V(t)^2, \tag{15}$$

where $V(t)$ is given by Eq. (10). The general linearized solution of the Eq. (15) is

$$R_0\zeta=Ce^{-\delta t}e^{i\sqrt{\omega_0^2-\delta^2}t}+\frac{1}{4}\int_{-\infty}^{+\infty}G(t-t')V^2(t')dt', \tag{16}$$

where $G(t)=\theta(t)/\sqrt{\omega_0^2-\delta^2}e^{-\delta t}\sin(\sqrt{\omega_0^2-\delta^2})t$, is the Green function and $\theta(t)$ is the Heaviside function. We assume the bubble is at rest at time $t=0$. Taking the Fourier transformation of the Eq. (16), we obtain,

$$4R_0\hat{F}(\zeta)=\frac{\hat{F}(V^2)}{\omega_0^2-\omega^2+2i\delta\omega}, \tag{17}$$

where denominator is the Fourier transformation of the Green's function. The sound-pressure radiation of the bubble at a distance r from it due to its volume pulsation is related to the acceleration of the bubble wall and is given by

$$p'=\rho R_0^2\ddot{\zeta}/r=-\rho\omega^2R_0^2\zeta/r, \tag{18}$$

where ω is the frequency of the spectral component of the dynamic pressure. To compute the bubble acoustic response to the turbulent pressure fluctuations we assume that the flow field is consisted of ensemble of vortices raging from the jet dimensions to Kolmogorov dissipation scale. We further assume that the velocity $V(t)$ is twice the horizontal turbulent velocity fluctuations u' if we disregard the bubble ascent velocity in space. Under these assumptions, the sound pressure spectrum of the deformed bubble, $S_{p'}(\omega)$ may be related to the spectral density of the turbulent fluctuations, $S_t(\omega)$,

$$S_{p'}(\omega)=\rho\frac{R_0}{r}\frac{S_t(\omega)}{\sqrt{(1-\omega_0^2/\omega^2)^2+4(\delta^2/\omega^2)}}, \tag{19}$$

where $S_{p'}(\omega)$ is the spectral density of the pressure amplitude induced by the bubble at the zero-order mode and is given by

$$S_{P'}(\omega)=\left|\int_{-\infty}^{+\infty}P'(t)e^{-i\omega t}\,dt\right|. \tag{20}$$

From Eqs. (14) and (19) it follows that a bubble can be reexcited by turbulent flow field when it encounters vortices with the same frequencies as the volume mode or the second surface mode oscillations of the bubble. However, the surface modes of oscillation are shown to be very weak and they do not contribute to the far-field sound radiation. The turbulent energy spectrum of the jet is broadband and covers large vortices down to the Kolmogorov dissipation eddies. One might expect that in a turbulent flow field there are always eddies with the same frequencies as the bubble volume pulsation, however these eddies may not be strong enough to excite the bubble. On the other hand, laminar flows have narrow-band spectrum and would not excite the bubble at all. This has been confirmed by our experimental observations. The intermittent nature of the turbulent flow field forces the bubble acoustic emissions to be stochastic. The acoustic reradiation of the bubble encountering the turbulent flow is strongly dependent upon the size and frequency of the vorticities contained in the fluid. What size vorticities are responsible for bubble acoustic radiation is the subject of the ongoing investigation.

The energy spectrum of the turbulent fluctuations was calculated at points distributed evenly across the jet at ten nozzle diameters downstream. The turbulent intensity is stronger at the vicinity of the center of the jet and becomes weaker as we get closer to the jet periphery. The power spectra of turbulent energy based upon horizontal and vertical velocity components at points located at the center of the jet and at one, two, and three nozzle diameters from the center are shown in Fig. 9(a). These energy spectra reflect the general trend of the velocity fluctuations and are decreasing with an increase in the radial distance. The drop in energy levels at higher frequencies are more pronounced at the jet periphery. This confirms that the flow at boundary of the jet has an intermittent structure and could be in transition from turbulence to laminar (jet's boundary corresponds to about 2.5D). Observations show that the bubble is ''acoustically active'' in the regions of the high-velocity fluctuations where it encounters violent vortices. The computed sound-pressure spectra of the bubble with turbulent energies of Fig. 9(a) as inputs are depicted in Fig. 9(b). Each curve represents the acoustic behavior of the bubble with size of 1.1 mm

trapped in various positions in the jet. These spectra were computed at a distance of 10 cm from the center of the bubble with total dissipation coefficient of $2\delta=0.014\omega$ (see Ref. 29). These predictions show that there exist a strong exchange of energy between the bubble and the turbulent flow when the frequency of eddies becomes equal to the natural frequency of the bubble. The inclusion of the third velocity component, assuming the turbulence intensity is isotropic, does not change the simulated acoustic response significantly. The different background noise levels in Fig. 9(b) correspond to various velocity fluctuations at radial distances of 0, 1, 2, and 3 nozzle diameters. In the case of bubble rising up through the jet the background noise in Fig. 9(b) would be the same.

In this paper, we have shown new evidence, both theoretically and experimentally, that the turbulent flow may reexcite adult bubbles and contribute to the underwater ambient noise in the ocean in the absence of the newly created bubbles. This may have cumulative effect in their contribution to the ambient noise.

III. CONCLUSIONS

We have examined the evidence of the ''adult'' bubble acoustic reradiation by a turbulent jet. This phenomena may occur in the ocean when breaking waves leave a chaotic highly turbulent flow field containing adult bubbles under several meters from the surface of the water. It is possible that the turbulent flow field may contribute to the underwater ambient noise in the absence of ''infant'' bubbles and may have cumulative effect in their contribution to the ambient noise.

It has been shown that an adult bubble injected into a fully developed turbulent flow field experiences a severe deformation. This deformation leads to bubble oscillations. The scaling analysis shows eddies with length scales in the range of the Taylor microscales may be responsible for bubble deformation and monopole acoustic radiation. A simple theory was given to account for the bubble excitation of the bubble. The Rayleigh–Plasset equation was modified by incorporating the turbulent pressure fluctuation of the flow as a driving force of the bubble. These predictions show there exist a strong exchange of energy between the bubble and the turbulent flow when the frequency of the turbulent pressure fluctuation, or twice the frequency of the velocity pulsations, becomes equal to the natural frequency of the bubble. The turbulent energy spectrum obtained experimentally was used to predict the monopole acoustic emission by the bubble.

The reexcitation phenomenon depends on the critical We number. In the case of an adult bubble acoustic reradiation, the critical number estimated to be around 0.52 while it is 1.1 for a bubble breakup.

ACKNOWLEDGMENTS

The author would like to thank Lawrence A. Crum, and Andrea Prosperetti for their valuable comments and enthusiasm, Jeffrey Kolodziejczak for the original design of the experimental setup, K. Markiewicz and Bogdan Marinescu for data collection and acquisition, and Sumon Sinha for his help in flow measurements techniques. This work is supported by the Office of Naval Research and the David Taylor Research Center.

[1] S. J. Knudsen, R. S. Afford, and J. W. Emling, ''Underwater ambient noise,'' J. Mar. Res. **7**, 410–429 (1948).
[2] A. Prosperetti, H. C. Pumphrey, and L. A. Crum, ''The underwater noise of rain,'' J. Geophys. Res. **94**, 3255–3259 (1989).
[3] H. C. Pumphery, L. A. Crum, and L. Bjørnø, ''Underwater sound produced by individual drop impacts and rainfall,'' J. Acoust. Soc. Am. **85**, 1518–1526 (1989).
[4] L. A. Crum, R. A. Roy, and A. Prosperetti, ''The underwater sounds of precipitation,'' Nav. Res. Rev., 2–12 (1992).
[5] H. Medwin and A. C. Daniel, ''Acoustical measurements of bubble prediction by spilling breakers,'' J. Acoust. Soc. Am. **88**, 408–412 (1990).
[6] D. M. Farmers, and L. Ding, ''Coherent acoustical radiation from breaking waves,'' J. Acoust. Soc. Am. **92**, 397–402 (1992).
[7] M. R. Loewen and W. K. Melville, ''An experimental investigation of the collective oscillations of bubble plumes entrained by breaking waves,'' J. Acoust. Soc. Am. **95**, 1329–1343 (1994).
[8] A. R. Kolaini, and L. A. Crum, ''Observations of underwater sound from laboratory breaking waves and the implications concerning ambient noise in the ocean,'' J. Acoust. Soc. Am. **96**, 1755–1765 (1994).
[9] S. M. Longuet-Higgins, ''Bubble noise mechanisms—A review,'' *Natural Physical Sources of Underwater Sound* (Kluwer, Dordrecht, The Netherlands, 1993).
[10] A. R. Kolaini, L. A. Crum, and R. A. Roy, ''Bubble production by capillary-gravity waves,'' J. Acoust. Soc. Am. **95**, 1913–1921 (1994).
[11] H. Lamb, *Hydrodynamics* (Dover, New York, 1945).
[12] M. Minnaert, ''On musical air-bubbles and the sounds of running water,'' Philos. Mag. **16**, 235–248 (1933).
[13] S. M. Longuet-Higgins, ''Monopole emission of sound by asymmetric bubble oscillation, Parts I and II,'' J. Fluid Mech. **201**, 525–565 (1989).
[14] S. M. Longuet-Higgins, ''Resonance in nonlinear bubble oscillations,'' J. Fluid Mech. **224**, 531–549 (1991).
[15] S. M. Longuet-Higgins, ''Nonlinear damping of bubble oscillations by resonant interaction,'' J. Acoust. Soc. Am. **91**, 1414–1422 (1992).
[16] J. E. Ffowcs-Williams, and Y. P. Guo, ''On resonant nonlinear bubble oscillations,'' J. Fluid Mech. **224**, 507–529 (1991).
[17] D. G. Crighton and J. E. Ffowcs-Williams, ''Sound generated by turbulent two-phase flow,'' J. Fluid Mech. **36**, 585–603 (1969).
[18] Y. Mao, L. A. Crum, and R. A. Roy, ''Nonlinear coupling between the surface and volume modes of an oscillating bubble,'' J. Acoust. Soc. Am. **98**, 2764–2771 (1995).
[19] W. K. Frizell, J. Cheng, and R. E. Arndt, ''Characteristics of bubble spilling noise,'' ASME Cavitation and Polyphase Flow Forum, 1987.
[20] A. V. Chalov, ''Acoustic emission upon the non stationary motion of a bubble in a liquid,'' Acoust. Phys. **40** (3), 386–388 (1994) (translated from Akus. Zh.).
[21] A. R. Kolaini, S. Sinha, and V. Rajendran, ''Interaction of bubbles with a turbulent flow: particle tracking and flow field characterization,'' The Proceedings of the 20th Symposium on Naval Hydrodynamics, 199–208, Santa Barbara, CA, 1994.
[22] H. Görtler, ''Berechnung von Aufgaber der freien Turbulenz auf Gund eines Neuen Naherungsansatzes,'' Z. Angew. Math. Mech. **22**, 244–254 (1942).
[23] J. O. Hinze, AIChE. J. **1**, 289–295 (1955).
[24] M. Sevik, and S. H. Park, ''The splitting of drops and bubbles by turbulent fluid flow,'' ASME J. Fluids Eng. **58**, 53–60 (1973).
[25] S. Takagi, A. Prosperetti, and Y. Matsumoto, ''Drag coefficient of a gas bubble in an axisymmetric shear flow,'' Phys. Fluids A **6**(9), 3186–3188 (1994).
[26] L. D. Landau and E. M. Lifshitz *Fluid Mechanics* (Pergamon, New York, 1959).
[27] G. R. Ruetsch and E. Meiburg, ''On the motion of small spherical bubbles in two-dimensional vortical flows,'' Phys. Fluids A **5**(10), 2326–2341 (1993).
[28] T. G. Leighton, *The Acoustics of Bubble* (Academic, New York, 1994).
[29] M. Strasberg, ''Gas bubbles as sources of sound in liquids,'' J. Acoust. Soc. Am. **28**, 20–26 (1956).

Wave reflection from a transitional layer between the seawater and the bottom

Rafael Carbó
Instituto de Acústica (C.S.I.C.), Serrano 144, 28006 Madrid, Spain

(Received 20 September 1995; revised 6 March 1996; accepted 22 June 1996)

This paper presents a theoretical study of sound wave propagation through media whose properties change along one of the coordinate rectangular axes, and remain constant in the plane orthogonal to this axis. The uppermost sea bottom layer could be modeled as such an inhomogeneous medium where the density, sound velocity, and absorption of acoustical energy varies along one axis. The results were applied to obtain the reflection coefficient of the water-bottom surface considering that there is a transitional inhomogeneous layer with density, velocity, and absorption gradient, and without discontinuity with the bottom. It was concluded that the reflection coefficient of the bottom surface shows a frequency variation, dependent on the physical property gradients in the uppermost sea bottom layer. © *1997 Acoustical Society of America.* [S0001-4966(97)05012-1]

PACS numbers: 43.30.Ma [JHM]

INTRODUCTION

The mineral and biological particles suspended in the sea near the bottom form a layer where the physical properties change with the depth. This is a layer whose porosity, density, grain size, etc., show a gradient in one rectangular axis, while remaining constant in the plane orthogonal to this axis.

The waves propagating through such a medium are affected by the gradients of density, sound velocity, and acoustical energy absorption. Brekhovskikh[1] deals with wave reflection from an inhomogeneous layer of the simplest form; in previous papers we studied the reflection coefficient[2,3] and the impulsive response[4] of a layer with exponential density gradient, and the reflectivity layer with arbitrary profiles of density and sound velocity.[5]

Analytical solutions of the Helmholtz equation are derived by Robins[6] for the reflection of plane acoustic waves from a layer of varying density between two homogeneous media. The solutions show that the reflection coefficient is dependent on the dimensionless vertical wave number and on the shape of the density profile in the intermediate layer. The results presented in that paper illustrate the dependence of the reflection coefficient on wave number and on the structure of the layer. In the limit of very low frequency, the reflection coefficient tends to the Rayleigh coefficient appropriate to the overall density change between upper and lower media, while, in the high-frequency limit, the reflection is given by the Rayleigh coefficient for the step change only.

In another paper, Robins[7] considers the transmission of an acoustic plane wave through a horizontally stratified fluid layer whose density and sound speed both vary continuously with depth.

In this paper a review of plane-wave reflection by a layer with density and sound velocity gradient is given, and we introduce the effect of a sound absorption gradient in the reflection coefficient. In this way the Biot[8] theory of sound wave propagation in a fluid-saturated porous solid, adapted by Hovem[9] to characterize the viscous attenuation of sound in suspensions and high porosity marine sediments, is considered.

I. MODELING THE GRADIENT MEDIUM

The following numerical model uses an inhomogeneous viscous fluid whose density varies with the depth, and whose sound speed and absorption vary with the depth and the frequency.

A. Theoretical model

In a gradient medium where the porosity $\beta(z)$ varies through the z axis, the density is written as

$$\rho(z)=\rho_f\beta(z)+\rho_s[1-\beta(z)], \quad \beta(z)=V_f/V, \tag{1}$$

where V_f is the volume of the fluid, V is the total volume, ρ_f is the density of the fluid, and ρ_s is the density of the solid grains.

A model of sound propagation through such a material is characterized by the sound wave velocity and the absorption coefficient. At low frequencies the sound velocity could be expressed by the Wood equation[10] applied to a suspension (such as mineral particles in water), or to any medium lacking rigidity:

$$c(z)=1/[\chi(z)\rho(z)]^{1/2}, \tag{2}$$

where χ is the total compressibility expressed by

$$\chi(z)=\beta(z)\chi_f+[1-\beta(z)]\chi_s, \tag{3}$$

where χ_f is the compressibility of the fluid and χ_s is the compressibility of the solid grains.

At high frequency, Biot predicts an increased velocity caused by the decrease in the effective density due to relative solid–fluid movement. The Biot model describes the wave propagation by two coupled differential equations giving as the solution two dilatational waves. Energy losses in the skeletal frame of the sediment, as well as viscous losses in the pore water, result in an overall attenuation which is a nonlinear function of frequency.

In a very soft sediment when both the shear and the bulk moduli of the frame are nearly zero, Hovem reduces the Biot solutions to one wave with a complex wave number which can be expressed, in the whole frequency range, by

$$(\mathbf{K}/\omega)^2=(\rho\chi)[(\rho_f^2/\rho)-\rho_c+i\,\eta F(\zeta)/\omega\kappa]/$$
$$[(2\rho_f-\rho-\rho_c)+i\,\eta F(\zeta)/\omega\kappa]. \quad (4)$$

According to Stoll,[11] κ is the absolute permeability of the medium, and ρ_c is an apparent mass density accounting for the fact that not all of the suspending material moves in the direction of the macroscopic pressure:

$$\rho_c=\epsilon\rho_f/\beta, \quad (5)$$

because of the complex geometry of the pore spaces in real sediments. In general, ϵ is slightly greater than one. Stoll[12] considers that $\epsilon=3$ is an adequate value when the pores are in all possible orientations.

The function $F(\zeta)$ is a function of frequency given by complex Kelvin functions, which accounts for the deviation from Poiseuille flow at high frequencies. The argument of this function is

$$\zeta=a_p(\omega\rho_f/\eta)^{1/2}. \quad (6)$$

Here, a_p is a pore size parameter and η is the fluid viscosity. For spherical grains Hovem and Ingram[13] derive the relation

$$a_p=\beta d_m/[3(1-\beta)], \quad (7)$$

where d_m is the sediment grain diameter. For well-sorted sand, the permeability appears to agree well with the Kozeny–Carman equation,[14] which gives κ as

$$\kappa=(\beta/5)(d_m/6)^2[\beta/(1-\beta)]^2. \quad (8)$$

The real part of the complex wave number **K** determines the sound wave velocity and the imaginary part the sound absorption coefficient.

Therefore, in an unconsolidated sediment such as the uppermost layer, when the shear and the bulk moduli of the frame can be neglected, and the physical parameters of the grain sediment material (ρ_s,χ_s) and fluid of the pores (ρ_f,χ_f,η) are given, the sound velocity and absorption coefficient are dependent on the frequency, porosity, and grain size. When the uppermost sea bottom layer shows a gradient of porosity and/or grain size in the vertical axis, the measurements at a given frequency of the sound velocity and the absorption coefficient along the length of the core samples of the uppermost bottom layer may be related with its porosity and the grain size gradients.

B. Laboratory measurements

Present knowledge of marine sediments has been mostly obtained through geophysical methods, using acoustics as a tool, and through samples collected and analyzed in laboratories. Different measurement and analysis techniques using data collected from different types of marine sediments in the vicinity of La Spezia, Italy, have been tested in a research program.[15]

Laboratory methods are available to measure compressional wave velocity and energy absorption in sediment samples removed from the bottom and brought to the surface through the process of coring. Core samples were collected carefully by divers and brought to the laboratory for the measurement of porosity, grain size diameter, compressional wave velocity, and acoustical energy absorption along the length of the core samples at 1-cm intervals.

Compressional wave velocity and attenuation were determined using an Underwater System Inc. (model USI-103) transducer–receiver head. A Tektronix PG 501 pulse generator, Tektronix FG 504 function generator, Krohn-Hite 3100R band pass filter, and Hewlett-Packard 1743A dual-time interval oscilloscope were substituted for the electronics and oscilloscope usually employed with the USI-103 velocimeter. These substitutions increased the resolution of compressional wave velocity measurements and allowed accurate measurements of received voltages required for attenuation measurements.

The transducer was driven with a 400-kHz, 20-V peak-to-peak sine wave triggered for 25-μs duration every 10 ms. The received signal was filtered (1 kHz–1 MHz high cutoff and low cutoff) prior to making time delay and received voltage measurements. Compressional wave velocity was calculated from the difference between time delay measurements made on a core filled with distilled water and time delay measurements made on the sediment cores. All velocities were corrected to 23 °C, 3.5% salinity, and 1 atm. Attenuation measurements were calculated as 20 log of the ratio of received voltage through distilled water versus received voltage through the sediment.

Two different marine sediments have been selected:

(1) The area situated at the eastern entrance of the Portovenere channel in the Gulf of La Spezia. Side-scan sonar records indicate a flat bottom, sub-bottom profiles taken along the measurement base indicate four distinct reflectors, the shallowest of which is about 1.5-ms distance from the bottom. The material of this bottom consists of clay (called in this paper ''soft bottom'').

Measurements in the core samples along the first 16 cm of the bottom give almost constant and very fine grain size distribution between 1 and 2 μm, while the porosity decreases quickly in the two first cm. This gradient of porosity is produced by a thin layer of suspended particles in water (turbidity). The acoustical parameters, compressional wave velocity, and energy absorption, measured at a frequency of 400 kHz, also give gradients in the first layer, negative in velocity and positive in absorption.

(2) Another area is located near Lerici in the Gulf of La Spezia, where the sea bed is covered with silty sand (called in this paper ''sand bottom''). Side-scan records indicate a regular bottom with some outcrops. Over the range in which the measurements were conducted the sea floor was uniform. The measurements in the core samples show a different behavior: a lower porosity and a grain size of 50 μm. The acoustical values measured along the cores taken from this sediment give positive gradients of compressional wave velocity and absorption energy losses.

The collected data in both areas are shown in Figs. 1 and 2.

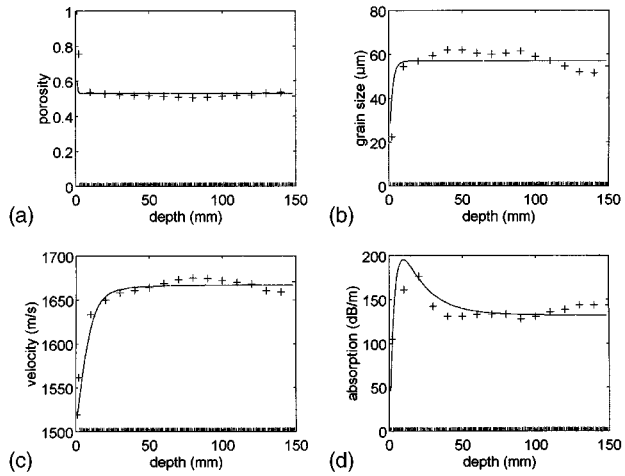


FIG. 1. Experimental data and fitted theoretical curves of (a) porosity, (b) grain size, (c) compressional wave velocity, and (d) sound absorption in Lerici area (''sand bottom'').

C. Numerical simulation

Models of soft sediments with high porosity and small grain sizes, and sandy sediments with low porosity and large grain sizes will be considered. Two bottom models are proposed using expression (4) and choosing the physical parameters summarized in Table I. Equations were devised for the variation of porosity and grain size with depth in each sediment, to give good fits to the measured values of sound speed and absorption along the length of the core samples, at the frequency of 400 kHz.

In the case of soft sediment functional fits in porosity and grain size given by

$$\begin{aligned} \beta &= 0.72 + (1 - 0.72)\exp(-50z), \\ d_m &= [1.05 - 0.03\ \exp(-300z)] \times 10^{-6}\ \text{m} \end{aligned} \tag{9}$$

are in good agreement with the experimental data.

In the sandy sediment the corresponding functional fits for the variation in porosity and grain size are given by

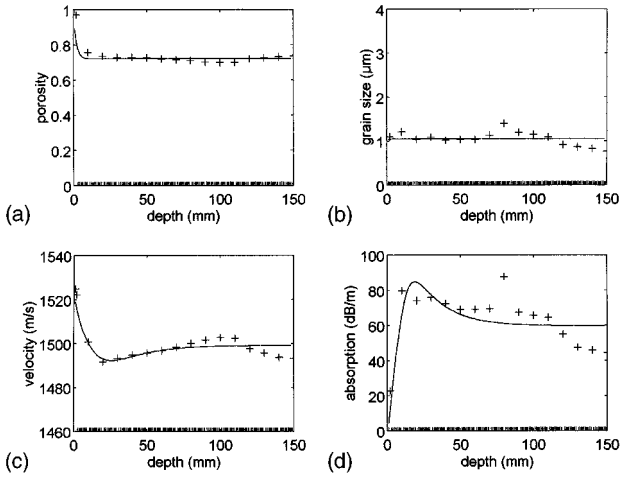


FIG. 2. Experimental data and fitted theoretical curves of (a) porosity, (b) grain size, (c) compressional wave velocity, and (d) sound absorption in Portovenere area (''soft bottom'').

TABLE I. Values of the fixed parameters used in the model calculations.

Parameter	Soft and sand sediment
Water density	$\rho_f = 1000$ kg/m^3
Water compressibility	$\chi_f = 4.3\times10^{-10}$ Pa^{-1}
Water viscosity	$\eta = 0.001$ Pa s
Grain density	$\rho_s = 2650$ kg/m^3
Grain compressibility	$\chi_s = 2.78\times10^{-11}$ Pa^{-1}
Structure constant	$\epsilon = 3$
Frequency	$f = 400$ kHz

$$\begin{aligned} \beta &= 0.53 + (1 - 0.53)\exp(-200z), \\ d_m &= [57 - 47\ \exp(-50z)] \times 10^{-6}\ \text{m}. \end{aligned} \tag{10}$$

The continuous lines in Figs. 1 and 2 shows the computed curves of porosity and grain size as a function of the depth according to these laws, and the computed variation of sound velocity and absorption with the depth at the frequency of 400 kHz, calculated from Eq. (4).

This model of the transitional layer assumes the anelasticity, therefore the sound velocity and absorption loss are functions of the frequency. Figure 3 shows 3D representations of the velocity and absorption coefficient as a function of frequency and depth in the ''soft bottom'' and ''sand bottom.''

II. ACOUSTIC FIELD IN AN INHOMOGENEOUS MEDIUM

The fundamental equations of the acoustic field in an inhomogeneous medium have the form[16]

$$\begin{aligned} \frac{\partial p}{\partial t} &= -\rho c^2\ \text{div}\ \mathbf{v}, \\ \frac{\partial \mathbf{v}}{\partial t} &= -\frac{1}{\rho}\ \text{grad}\ p, \end{aligned} \tag{11}$$

where p is the acoustic pressure, $\mathbf{v}$ is the particle velocity in the wave, and c the velocity of sound propagation. In the general case ρ and c are functions of position. The first of these equations is the equation of continuity and the second

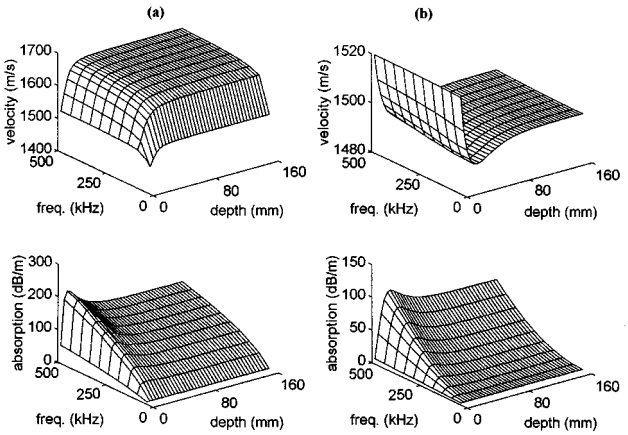


FIG. 3. Frequency and depth dependence of the sound velocity and absorption coefficient in the transitional layer. (a) ''Sand bottom;'' (b) ''soft bottom.''

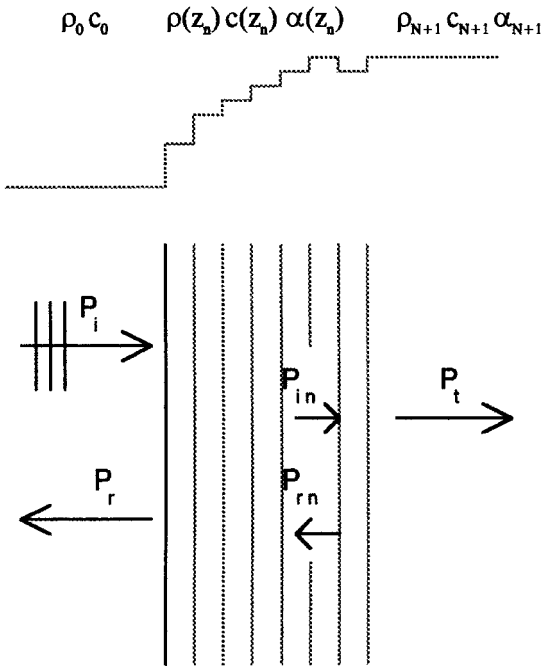


FIG. 4. Reflection of plane waves normally incident on a transitional inhomogeneous layer.

the Euler equation. Assuming harmonic time dependence, $\partial/\partial t = -i\omega$, and eliminating $\mathbf{v}$ from Eqs. (11), we obtain

$$\rho \ \mathrm{div}[(1/\rho)\mathrm{grad}\ p] + k^2 p = 0, \tag{12}$$

$$\nabla^2 p + k^2 p - (1/p)\mathrm{grad}\ \rho\ \mathrm{grad}\ p = 0. \tag{13}$$

Introducing in place of p a new potential ϕ, defined by

$$\phi = p/(\rho^{1/2}), \tag{14}$$

after some transformations, we obtain the wave equation for ϕ

$$\nabla^2 \phi + \mathbf{K}(z)^2 \phi = 0, \tag{15}$$

where

$$\mathbf{K}(z)^2 = \mathbf{k}(z)^2 + (1/(2\rho(z))\nabla^2 \rho(z) - (3/4)[(1/\rho(z))\mathrm{grad}\ \rho(z)]^2. \tag{16}$$

This wave equation can be integrated in several cases,[1–3] but a general solution of wave equation (15) is not possible. Nevertheless we solve the problem by imagining a layer consisting of a large number of thin, plane parallel, homogeneous layers in contact with one another. In passing from layer to layer, the properties of the medium change discontinuously. However, letting the thickness of the layers approach zero while their number approaches infinity, we obtain a layered-inhomogeneous medium with continuously varying parameters.

A. Reflection by an inhomogeneous layer

The plane-wave reflection coefficient plays an important role in the interpretation of acoustic data in modern marine seismology. This coefficient is based on the partitioning of energy at the interface between water and sediment, and the result is the classical Rayleigh reflection coefficient, which relates the amplitudes of homogeneous incident wave and homogeneous reflected wave.

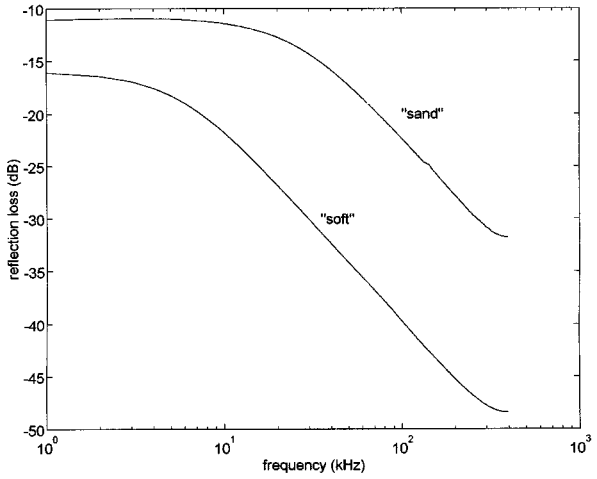


FIG. 5. Reflection loss by a transitional inhomogeneous layer: soft bottom and sand bottom.

Let us consider a one-dimensional inhomogeneous layer in the z-axis direction, with a thickness d, sandwiched between two homogeneous and semi-infinite media: the water, with acoustic impedance $\rho_0 c_0$, and the bottom with equal impedance to the layer at $z=d$. From the water a plane-wave sound beam at normal incidence strikes the upper boundary surface of the layer (see Fig. 4).

Now we can get the pressure of the incident and reflected waves in the water

$$\mathbf{P}_i = \mathbf{A}_0 \exp[i(\omega t - k_0 z)], \quad \mathbf{P}_r = \mathbf{B}_0 \exp[i(\omega t + k_0 z)], \tag{17}$$

and the velocity of the particles,

$$\mathbf{v}_i = (1/Z_0)\mathbf{A}_0 \exp[i(\omega t - k_0 z)], \quad \mathbf{v}_r = (-1/Z_0)\mathbf{B}_0 \exp[i(\omega t + k_0 z)], \tag{18}$$

where $Z_0 = \rho_0 c_0$ is the acoustic impedance, k the wave number in the water, and $\mathbf{A}_0$, $\mathbf{B}_0$ the complex amplitudes of these waves.

Assuming that a plane wave is propagated through the inhomogeneous layer divided into N homogeneous layers in contact with thickness $\Delta z (\Delta z = d/N)$, the pressure and the particle velocity in the nth layer will be

$$\mathbf{P}_{\mathrm{in}} = \mathbf{A}_n \exp[i(\omega t - \mathbf{k}_n z)], \quad \mathbf{P}_{\mathrm{rn}} = \mathbf{B}_n \exp[i(\omega t + \mathbf{k}_n z)], \tag{19}$$

$$\mathbf{v}_{\mathrm{in}} = (1/\mathbf{Z}_n)\mathbf{A}_n \exp[i(\omega t - \mathbf{k}_n z)], \quad \mathbf{v}_{\mathrm{rn}} = (-1/\mathbf{Z}_n)\mathbf{B}_n \exp[i(\omega t + \mathbf{k}_n z)], \tag{20}$$

where $\mathbf{Z}_n = \rho(z_n)c(z_n)/(1 + i\alpha_n/k_n)$ is the complex impedance of the nth layer and $\mathbf{k}_n = k_n + i\alpha_n$ is the complex wave number, where the imaginary part is the absorption coefficient of the nth layer. Since the absorption coefficient is dependent on the frequency, the complex impedance is also dependent on the frequency.

The N layers division of the inhomogeneous layer considered above lead to a time relationship between the pres-

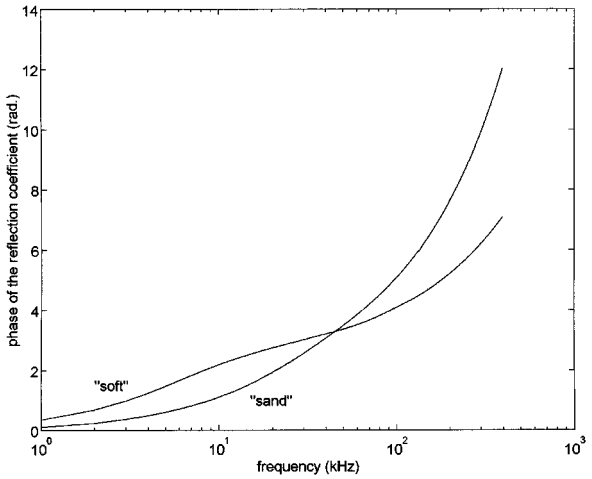


FIG. 6. Phase of the reflection coefficient as a function of frequency by a transitional layer: ''soft'' bottom and ''sand'' bottom.

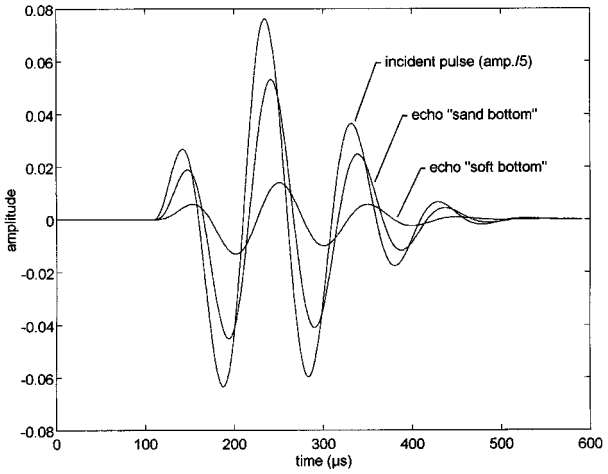


FIG. 7. Incident acoustic pulse (amplitude/5) and echoes from a ''sand'' bottom transitional layer and a ''soft'' bottom transitional layer.

sure of the wave reflected in layer $n+1$, $\mathbf{P}_{\mathrm{rn}}$, and the pressure of the wave reflected in the first layer, $\mathbf{P}_r$, given by the phase

$$\Phi_n = \sum_{m=0}^{n} \mathbf{k}_m \Delta z.$$

Finally, in the bottom space, opposite to the water space, only a transmitted wave travels in the positive z axis direction, with pressure and particle velocity

$$\begin{aligned} &\mathbf{P}_t = \mathbf{A}_t \exp[i(\omega t - \mathbf{k}_{N+1} z)], \\ &\mathbf{v}_t = (1/\mathbf{Z}_{N+1})\mathbf{A}_t \exp[i(\omega t + \mathbf{k}_{N+1} z)]. \end{aligned} \tag{21}$$

According to Stoll[17] the complex amplitudes may be found by considering the following boundary conditions at each interface: (1) continuity of normal traction; (2) continuity of shear traction; (3) continuity of fluid movement in and out of the skeletal frame in a direction normal to the interface; (4) continuity of pore pressure. In an unconsolidated sediment, such as the uppermost bottom layer, the shear and bulk moduli of the frame can be neglected. Therefore, the boundary conditions can be reduced to the continuity of the acoustic pressure and the particle velocities on both sides of the surfaces. These boundary conditions must be satisfied at all times for all points on the $N+1$ boundaries between the layers. Applying these conditions we obtain a pair of equations at each nth boundary given by the matrix expression

$$\begin{pmatrix} \mathbf{A}_n \\ \mathbf{B}_n \end{pmatrix} = \mathbf{T}_n \begin{pmatrix} \mathbf{A}_{n+1} \\ \mathbf{B}_{n+1} \end{pmatrix}, \tag{22}$$

where $\mathbf{T}_n$ is the matrix

$$\mathbf{T}_n = \begin{pmatrix} (\mathbf{Z}_{n+1}+\mathbf{Z}_n)\exp[-i(\mathbf{k}_{n+1}-\mathbf{k}_n)z_n] & (\mathbf{Z}_{n+1}-\mathbf{Z}_n)\exp[i(\mathbf{k}_{n+1}+\mathbf{k}_n)z_n] \\ (\mathbf{Z}_{n+1}-\mathbf{Z}_n)\exp[-i(\mathbf{k}_{n+1}+\mathbf{k}_n)z_n] & (\mathbf{Z}_{n+1}+\mathbf{Z}_n)\exp[i(\mathbf{k}_{n+1}-\mathbf{k}_n)z_n] \end{pmatrix}. \tag{23}$$

By applying Eq. (23) N times we obtain the relation between the amplitude of the pressure on both sides of the inhomogeneous layer

$$\begin{pmatrix} \mathbf{A}_0 \\ \mathbf{B}_0 \end{pmatrix} = \mathbf{T}_0 \quad \mathbf{T}_1 \quad \ldots \quad \mathbf{T}_N \begin{pmatrix} \mathbf{A}_t \\ 0 \end{pmatrix}. \tag{24}$$

The complex reflection coefficient of the inhomogeneous layer will be determined by

$$\mathbf{R} = \mathbf{A}_0 / \mathbf{B}_0. \tag{25}$$

It is very important to choose carefully the number of layers when dividing the inhomogeneous layer into N homogeneous layers. In order to check the approximate solution we compare the reflection coefficient with the value obtained for twice the number of layers, and we consider the solution to be correct when the difference in the result is lower than 1%. To achieve a good approximation at high frequencies (400 kHz) it was necessary take a thickness of layers of 0.5 mm along the transitional layer ($d=160$ mm).

According to this consideration we apply the approximate method to two examples studied, ''soft'' and ''sand'' sediment, in order to calculate the frequency dependence of the reflection coefficient. Figure 5 shows this variation; the amplitude of the reflection coefficient decreases strongly and the phase increases with the frequency in both sediments. The reflection loss in the soft bottom rises from 16 dB at 1 kHz to 40 dB at 100 kHz, while in the sand bottom it rises from 11 dB at 1 kHz to 22 dB at 100 kHz.

The phase of the reflected wave changes on reflection from the transitional layer. The phase increases with the fre-

quency of the wave in both kinds of bottom. Figure 6 shows how at low frequencies, the phase of the reflection coefficient is lower in the sand bottom than in the soft bottom, while at high frequencies the phase is higher in the sand bottom.

In order to compare the incident pulse with the echo from a bottom transitional layer, the impulsive response of the modeled bottom has been calculated by FFT of the reflection coefficient. An analogue signal of a sonar pulse of 100-μs length at 10 kHz was also introduced and convolved with the impulse response of both a soft and sand bottom transitional layer, obtaining the echoes shown in Fig. 7. A different reduction of intensity is observed, 28 times in the soft bottom and 7 times in the sand bottom, and the shape of the echoes also is clearly modified.

III. CONCLUSIONS

Biological and mineral particles suspended in the sea near the bottom in a coastal region, where there are marine currents, produce high optical turbidity. The acoustic wave propagation in such a medium is strongly modified at high frequencies. A transitional layer is formed between the clear bottom and the sedimentary bottom with continuous variation of its physical characteristics.

The effect of the transitional layer is to reduce the bottom echo amplitude at higher frequencies and cause the center frequency of the reflected signal to be lowered. Therefore, calculation of the acoustical impedance of the bottom by an inversion method based on measurement of the amplitude of the echo pulse at only one frequency may be erroneous, even if the bottom surface is a plane.

The pulses emitted by echo sounders are usually narrow bandwidth pulses, so the echo has very reduced frequency information. Measurement of the reflection coefficient of the bottom surface layer with two or more central frequency pulses is very convenient to deduce its frequency dependence. To obtain the acoustical impedance of the bottom covered by a transitional layer it is necessary to measure the value of the reflection coefficient at very low frequency. For instance, for the bottoms studied in this paper this measurement is required at frequencies bellow 2 kHz.

ACKNOWLEDGMENTS

The author wishes to thank Dr. M. Richardson of NORDA, Mississippi, for the provision of acoustical data, and to extend special thanks to Dr. T. Akal of SACLANTCEN, La Spezia, Italy, for his assistance in the experimental cruise.

[1] L. M. Brekhovskikh, *Waves in Layered Media* (Academic, New York, 1980).

[2] C. Ranz-Guerra and R. Carbó-Fite, ''Impulsive response of sediment layers with variable density gradient,'' in *Acoustics and the Sea-Bed* (Bath University, Bath, 1983), pp. 51–60.

[3] R. Carbó-Fite, M. Camarasa-Rius, C. Ranz-Guerra, and A. Soler-Navarro, ''Modelo acústico de fondo marino: cálculo del coeficiente de reflexión complejo, $R(w)$, para un fondo semi-infinito y una capa ambos con gradiente de densidad,'' Anal. Física B **72**, 108–114 (1983).

[4] R. Carbó-Fite, M. Camarasa-Rius, C. Ranz-Guerra, and A. Soler-Navarro, ''Estudio experimental de la propagación de un impulso de corta duración en un medio inhomogéneo,'' Anal. Física B **81**, 61–67 (1985).

[5] P. Cobo, C. Ranz, and R. Carbó, ''Reflectivity of a layer with arbitrary profiles of density and sound velocity,'' in *Progress in Underwater Acoustics* (Plenum, New York, 1986), pp. 311–317

[6] A. J. Robins, ''Reflection of plane acoustic waves from a layer of varying density,'' J. Acoust. Soc. Am. **87**, 1546–1552 (1990).

[7] A. J. Robins, ''Reflection of a plane wave from a fluid layer with continuously varying density and sound speed,'' J. Acoust. Soc. Am. **89**, 1686–1696 (1991).

[8] M. Biot, ''Theory of propagation of elastic waves in fluid-saturated porous solid. I Low-frequency range, and II Higher-frequency range,'' J. Acoust. Soc. Am. **26**, 168–191 (1956).

[9] J. M. Hovem, ''Viscous attenuation of sound in suspensions and high porosity marine sediments,'' J. Acoust. Soc. Am. **67**, 1559–1563 (1980).

[10] A. W. Wood, *A Textbook of Sound* (Bell, London, 1957).

[11] R. D. Stoll, ''Acoustic waves in marine sediments,'' in *Ocean Seismo-Acoustics* (Plenum, New York, 1986), pp. 417–434.

[12] R. D. Stoll and G. M. Bryan, ''Wave attenuation in saturated sediments,'' J. Acoust. Soc. Am. **47**, 1440–1447 (1970).

[13] J. M. Hovem and G. D. Ingram, ''Viscous attenuation of sound in saturated sand,'' J. Acoust. Soc. Am. **66**, **1807–1812 (1979).**

[14] P. C. Carman, *Flow of Gases through Porous Media* (Academic, New York, 1956).

[15] T. Akal, H. Schmidt, and P. Curzi, ''The use of love waves to determine the geoacoustic properties of marine sediments,'' in *Ocean Seismo-Acoustics* (Plenum, New York, 1986), pp. 841–852.

[16] L. Landau and E. Lifshitz, *Mechanics of Continuous Media* (Gostckhizdat, Moscow, 1944).

[17] R. D. Stoll and T. K. Kan, ''Reflection of acoustic waves at a water-sediment interface,'' J. Acoust. Soc. Am. **70**, 149–156 (1981).

Internal friction and boundary conditions in lossy fluid seabeds

Grant B. Deane[a)]
Marine Physical Laboratory, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093-0238

(Received 9 March 1994; accepted for publication 13 September 1996)

There are two distinct mechanisms associated with compressional wave absorption in lossy media, internal relaxation and internal friction. For the special case of propagation in an homogeneous, unbounded medium, both mechanisms can be modeled by adopting the convention of a complex sound speed and are, in this sense, equivalent. For the more realistic case of propagation in a stratified medium, the convention of complex sound speed does not give a correct description for losses which modify the linearized equation of motion, such as internal friction. In the presence of boundaries, internal friction can be modeled by the introduction of a complex quiescent density in addition to complex sound speed. Propagation models which use complex sound speed only in the presence of boundaries make the tacit assumption that seafloor losses are caused by internal relaxations only. A solution is developed for propagation in a lossy Pekeris channel where absorption in the lower fluid is caused by internal friction. The example that has been considered yields a sound level 3 dB less than the standard description over a 50-km path. © *1997 Acoustical Society of America.* [S0001-4966(97)04401-9]

PACS numbers: 43.30.Ma [MBP]

INTRODUCTION

There are two types of physical process associated with sound absorption in granular materials, internal relaxation and internal friction. Internal relaxations represent the acoustic energy lost exciting a thermodynamic state intrinsic to the material. The chemical relaxation associated with magnesium sulfate in sea water and thermal conduction are two examples of internal relaxation mechanisms. Internal frictions represent losses associated with the motion of fluid particles and are associated with the diffusion of momentum. Viscosity and Darcy's law for fluid flow through a cylindrical tube are two examples of internal frictions.

Despite the collection of a large quantity of experimental data,[1] the cause of compressional wave absorption in marine sediments, many of which are granular in nature, remains a subject of debate.[2] There is evidence that the absorption mechanism in dry, porous materials is intergranular friction[3,4] whereas saturated, porous materials appear to exhibit two loss mechanisms, intergranular friction and viscosity of the pore fluid.[5,6] Along with experimental studies, detailed analytical models have been developed for the geoacoustic properties of saturated, porous materials[7–9] and marine sediments.[6,10] These models describe propagation and absorption in infinite (unbounded) materials. Because of the complexity of these analytical models and the uncertainty in the experimental data surrounding absorption mechanisms, underwater acousticans usually adopt a phenomenological description for losses.

The widely adopted convention to account for compressional wave absorption in marine sediments is to make the sound speed complex.[11,12] The convention is used for propagation in both homogeneous media without boundaries, and for problems which involve boundaries between media. However, models which use complex sound speed only in the presence of boundaries make the tacit assumption that compressional wave absorption is caused by internal relaxations, and do not allow for the possibility of internal friction. This observation, which is justified in Sec. I, is a consequence of how the loss mechanisms appear in the linearized acoustic equations. Internal relaxations are modeled by modifying the lossless equation of state, whereas internal frictions modify the equation of motion.[13]

An example of a problem where the difference between internal relaxation and friction is important is propagation in a Pekeris wave guide with a lossy bottom. Kornhauser and Raney[14] solved this problem assuming the absorption in the lower fluid to be representable by a complex sound speed, and thus caused by an internal relaxation. The authors based their analysis on a dispersion relation derived by Morse[15] while considering sound propagation in granular media. We reconsider this problem with the assumption that the compressional wave absorption in the sea bed is caused by an internal friction mechanism which modifies the equation of motion.

Section I outlines how losses due to internal relaxation and friction are introduced into the lossless, linear acoustic equations. In Sec. II, we derive a solution for propagation in a Pekeris wave guide assuming a friction loss mechanism, and compare the results with those obtained assuming internal relaxation. Section III gives a brief account of how relaxation and frictional losses might be distinguished from acoustic measurements and the concluding remarks are given in Sec. IV.

[a)]On leave at Naval Research Laboratory, Washington, DC 20375.

I. THEORY

A. The equations of linear acoustics for a lossless fluid

The time domain form of the linear acoustic equations governing propagation through a lossless fluid are (see, for example, Pierce,[16] p.15)

$$\rho \nabla \cdot \tilde{\mathbf{v}} + \frac{\partial \tilde{\rho}'}{\partial t} = 0, \tag{1}$$

$$\nabla \tilde{p} + \rho \frac{\partial \tilde{\mathbf{v}}}{\partial t} = 0, \tag{2}$$

$$\tilde{p} = (K_s/\rho)\tilde{\rho}', \tag{3}$$

where K_s is the adiabatic bulk modulus, ρ is the quiescent fluid density and $\tilde{\rho}'$, $\tilde{p}$, and $\tilde{\mathbf{v}}$, respectively, are the time-dependent fluctuations in medium density, pressure, and velocity. We have adopted the notation that a tilde over a symbol denotes a time-domain variable.

Equation (1) arises from the conservation of mass, Eq. (2) is the linearized form of Euler's equation of motion for a fluid, and Eq. (3) is the equation of state. To derive the lossless wave equation, we eliminate $\tilde{\rho}'$ from Eq. (1) using Eq. (3), differentiate the resulting expression with respect to time and substitute the result into Eq. (2) to obtain

$$\rho \nabla \cdot \frac{1}{\rho} \nabla \tilde{p} - \frac{\rho}{K_s} \frac{\partial^2 \tilde{p}}{\partial t^2} = 0. \tag{4}$$

A more careful derivation of Eq. (4) has been given by Bergmann.[17]

B. Internal relaxation losses

The standard way to model losses due to an internal relaxation mechanism is to introduce a convolution relation between the pressure and density perturbations in Eq. (3) of the form

$$\tilde{p} = \frac{K_s}{\rho} \int_{-\infty}^{t} \tilde{h}(t-\tau)\tilde{\rho}'(\tau)\, d\tau, \tag{5}$$

where $\tilde{h}$ is an impulse response function relating the pressure to an impulsive density change. To ensure that the material being modeled is causal, we require $\tilde{h}(t)=0$ for $t<0$. The tacit assumption underlying Eq. (5) is that the material under consideration is a linear, time-invariant system. No particular physical mechanism needs to be (or indeed usually is) ascribed to the impulse response function, except to assume that it corresponds to an internal relaxation mechanism and therefore modifies the equation of state. The Fourier transform of the impulse response function, $h(w)$, is assumed to be of the form

$$h(w) = 1 - i\epsilon, \tag{6}$$

over some frequency range $\omega_0 < \omega < \omega_1$, where $\epsilon \ll 1$. The experimental observation that sound absorption increases linearly with frequency leads to the requirement that ϵ be independent of frequency.

The convention of complex sound speed arises from combining h and K_s into a single complex constant K_s. The internal relaxation mechanism is now conveniently accounted for by defining the complex sound speed to be

$$\overline{c} = \sqrt{K_s/\rho}. \tag{7}$$

This definition is no more than a convention which allows us to account for an internal relaxation absorption mechanism.

C. Internal friction losses

Internal friction can be modeled by introducing a dissipative drag force into the equation of motion. The drag force is expressed as a convolution between velocity or gradient in velocity and an impulse and an impulse response function $\tilde{g}$:

$$\overline{\mathbf{F}}_d = \rho \int_{-\infty}^{t} \tilde{g}(t-\tau) L(\tilde{\mathbf{v}}(\tau))\, d\tau, \tag{8}$$

where L is a spatial differential operator determined by the exact nature of the internal friction. Equation (8) is sufficiently general to model several known frictional mechanisms.

1. Viscosity

Let the spatial differential operator $L=\nabla^2$ and the impulse response function $\tilde{g}=4/3\nu\delta(t)$, where ν is the kinematic viscosity and δ is the Dirac delta function. Then the drag force is the longitudinal part of the viscous stress given by[18]

$$\tilde{\mathbf{F}}_{\text{viscosity}} = \tfrac{4}{3}\rho\nu\nabla^2\tilde{\mathbf{v}}. \tag{9}$$

Equation (8) is accurate provided that the bulk viscosity of the material is negligible compared to the kinematic viscosity. When incorporated into the wave equation, viscosity yields an absorption which depends on the square of frequency at low frequencies and the square root of frequency at high frequencies, a result which is inconsistent with the observation that absorption scales linearly with frequency in marine sediments.[1]

2. Darcy's law

Let L be the identity operator and $\tilde{g}=-\alpha_0\delta(t)$ where α_0 is a constant. Then

$$\tilde{\mathbf{F}}_d = -\rho\alpha_0\tilde{v}, \tag{10}$$

which is Darcy's law for fluid flow through a cylindrical tube. When incorporated into the wave equation, Darcy's law yields an absorption that is independent of frequency, and so is also inconsistent with observation. The failure of Darcy's law and viscosity to yield the observed dependence on frequency does not rule out an internal friction loss mechanism contributing to (or dominating) compressional wave absorption. As is done for internal relaxations, we will adopt a phenomenological approach and assume an unspecified friction mechanism modeled by Eq. (8).

D. The equations of linear acoustics for a lossy fluid

The acoustic equations for a fluid with relaxation and frictional losses are obtained by adding the drag force of Eq. (8) to Eq. (2), and adopting the modified equation of state. The resulting three equations are

$$\rho \nabla \cdot \tilde{\mathbf{v}} + \frac{\partial \tilde{\rho}'}{\partial t} = 0, \tag{11}$$

$$\nabla \tilde{p} + \rho \frac{\partial \tilde{\mathbf{v}}}{\partial t} + \rho \int_{-\infty}^{t} \tilde{g}(t-\tau) L(\tilde{\mathbf{v}}(\tau)) d\tau = 0, \tag{12}$$

$$\tilde{p} = \frac{K_s}{\rho} \int_{-\infty}^{t} \tilde{h}(t-\tau) \tilde{\rho}'(\tau) d\tau. \tag{13}$$

These equations are similar to those used by Verweij[13] to model losses. Verweij's formulation is slightly different in that Eqs. (11) and (13) have been combined to eliminate density perturbations, and inertia and compliance memory functions are used in preference to impulse response functions. In his analysis, Verweij considered absorption in an unbounded medium, whereas we are concerned with material discontinuities and boundary conditions.

Further progress cannot be made with Eqs. (11)–(13) in the time domain unless specific mechanisms are adopted for the impulse response functions $\tilde{h}$ and $\tilde{g}$. We can, however, perform a pseudo-harmonic analysis and Fourier transform the equations with the intention of employing experimentally determined forms for the transformed impulse response functions. The transformed equations are

$$\rho \nabla \cdot \mathbf{v} - i\omega \rho' = 0, \tag{14}$$

$$\nabla p - i\omega \rho \left(1 + i \frac{g(\omega)}{\omega} L \right) \mathbf{v} = 0, \tag{15}$$

$$p = (K_s/\rho) h(\omega) \rho', \tag{16}$$

where $\mathbf{v}$, p, and ρ, respectively, are the Fourier transformed variables $\tilde{\mathbf{v}}$, $\tilde{p}$, and $\tilde{\rho}'$. Equations (14), (15), and (16) are the basis of our analysis of acoustic absorption in a lossy fluid.

At this point we are obliged to make a particular choice for L, and for the sake of simplicity set the operator to unity. This choice models a frictional loss which depends directly on the acoustic velocity field, and in this sense is related to Darcy's law [see Eq. (10)]. It is not precisely Darcy's law because of the linear dependence of absorption on frequency introduced later. Solving (14)–(16) for p results in the Helmholtz equation for the lossy material

$$\rho \nabla \cdot \left(\frac{1}{\rho \left(1 + i \frac{g(\omega)}{\omega} \right)} \nabla p \right) + \frac{\rho \omega^2}{K_s h(\omega)} p = 0. \tag{17}$$

If we now make the substitutions

$$\overline{\rho} = \rho \left(1 + i \frac{g(\omega)}{\omega} \right) \tag{18}$$

and

$$\overline{c} = \sqrt{\frac{K_s h(\omega)}{\rho (1 + i[g(\omega)/\omega])}}, \tag{19}$$

we obtain our final form for the reduced wave equation:

$$\overline{\rho} \nabla \cdot \left(\frac{1}{\overline{\rho}} \nabla p \right) + \frac{\omega^2}{\overline{c}^2} p = 0. \tag{20a}$$

Examining Eqs. (18)–(20a) we see that introducing a loss term into the equation of motion has resulted in the convention of both complex quiescent density and complex sound speed. The medium density is, of course, actually real. The idea of complex density is, like complex sound speed, simply a convention adopted to simplify algebraic manipulations for the chosen form of frictional loss. For the special case of Darcy's drag law [where $\tilde{g} = -\alpha_0 \delta(t)$ and α_0 is a constant], Eq. (18) is equivalent to the expression given by Morse and Ingard (Ref. 18, p. 254, equation 6.2.25) in connection with the analysis of acoustic propagation in porous materials. Morse and Ingard call $\overline{\rho}$ ''effective'' density.

Inside each layer of a horizontally stratified medium, quiescent density is constant and Eq. (20a) simplifies to the standard form for the wave equation:

$$\nabla^2 p + (\omega^2/\overline{c}^2) p = 0. \tag{20b}$$

For homogeneous materials, only complex sound speed appears in the wave equation. However, if frictional losses are present complex density reenters the problem through the boundary conditions which must be satisfied at each layer interface. For fluid layers, both the pressure across the boundary and the particle velocity normal to the boundary must be continuous. These conditions can be expressed in terms of the velocity potential, defined by $-\nabla \psi = \mathbf{v}$. The requirement that the component of velocity normal to the boundary be continuous can be expressed as

$$\frac{\partial \psi_1}{\partial z} = \frac{\partial \psi_2}{\partial z}, \tag{20c}$$

where the equality holds at the boundary. From Eq. (15)

$$p = -i\omega \rho \left(1 + i \frac{g}{\omega} \right) \psi \tag{20d}$$

and hence

$$\overline{\rho_1} \psi_1 = \overline{\rho_2} \psi_2, \tag{20e}$$

where the subscripts 1 and 2, respectively, denote variables in the upper and lower layers. The densities in Eq. (20e) are complex because of the frictional loss term which modifies the equation of motion, and this is the crux of the matter. Because frictional losses are associated with the diffusion of momentum, they manifest as a drag force and yield different boundary conditions from relaxation losses. Thus the choice of description for the absorption mechanism is important for problems where boundaries appear, such as stratified ocean models.

E. Values for complex density

To illustrate the effect of assuming a frictional loss rather than an internal relaxation, we will set $h(\omega) = 1$ (no internal relaxation), and the friction impulse response function to

$$g(\omega) = \alpha |\omega|, \quad \omega_0 < |\omega| < \omega_1, \tag{21}$$

where α is a real, dimensionless constant and ω_0 and ω_1, respectively, are the lower and upper limits for the observed linear dependence of absorption on frequency. Taking the

modulus of ω ensures that $g(-\omega)=g^*(\omega)$, where * denotes complex conjugate, which satisfies the requirement that the time-domain drag force in Eq. (8) be real. Equation (21) for $g(\omega)$ results in absorption proportional to frequency over the range ω_0 to ω_1 and is therefore consistent with the available experimental evidence. An absorption proportional to frequency implies that the quality factor Q of the material is constant, which cannot be true for all frequencies without violating the principle of causality.[2] For this reason, we have limited the range for the linear dependence of g on frequency.

Substituting Eq. (21) into Eq. (18) yields a complex density of the form

$$\overline{\rho}=\rho\left(1+i\,\frac{g(\omega)}{\omega}\right)=\rho(1+i\alpha),\quad \omega_0<\omega<\omega_1. \tag{21a}$$

As the factor $(1+i\alpha)$ appears frequently, we will adopt the notation

$$\overline{\rho}=\rho\beta,\quad \text{where}\quad \beta=1+i\alpha. \tag{22}$$

With $h(\omega)=1$, the complex sound speed is given by

$$\overline{c}=c\beta^{-1/2},\quad \text{where}\quad c=\sqrt{K_s/\rho}. \tag{23}$$

Consider a one dimensional plane wave $p(x)=\exp(ikx)$ propagating through an unbounded, homogenous medium which exhibits an internal friction as described above. It follows from Eqs. (20b), (23), and (22) that the pressure wave decays at a rate $\exp(-\alpha_e x)$ where

$$\alpha_e=f\,\frac{\sqrt{2}\pi}{c}\,((1+\alpha^2)^{1/2}-1)^{1/2} \tag{23a}$$

is the absorption expressed in nepers per meter and α is the dimensionless constant introduced in Eq. (21). For small values of α, a Taylor series expansion of Eq. (23a) yields the approximate expression

$$\alpha_e\approx(\pi\alpha/c)f. \tag{23b}$$

For $\alpha\ll 1$, Eq. (23b) provides the relationship between absorption expressed in nepers per meter and the constant α associated with the frictional mechanism. For a constant α, the absorption scales linearly with frequency.

The relationship between small α and the medium absorption Ω expressed in dB per wavelength is

$$\Omega=2\pi\Delta\alpha, \tag{24}$$

where Δ is a constant defined by $\Delta=20\log_{10}(e)\cong 8.686$.

From Eqs. (22) and (23), the complex density scales as β and the complex sound speed scales as $\beta^{-1/2}$. In Fig. 1 we have plotted the imaginary components of β and $\beta^{-1/2}$ which respectively correspond to the imaginary parts of $\overline{\rho}/\rho$ and $\overline{c}/c$. The horizontal axis is logarithmic in Ω, and covers seabed absorption in the range 0.01 to 1 dB/wavelength. Over this range the real component of the scaled complex sound speed is unity to a very good approximation. The real component of the scaled complex density is always unity. For small values of seabed absorption the imaginary part of the density is approximately twice the magnitude of the imaginary part of the sound speed, and has the opposite sign.

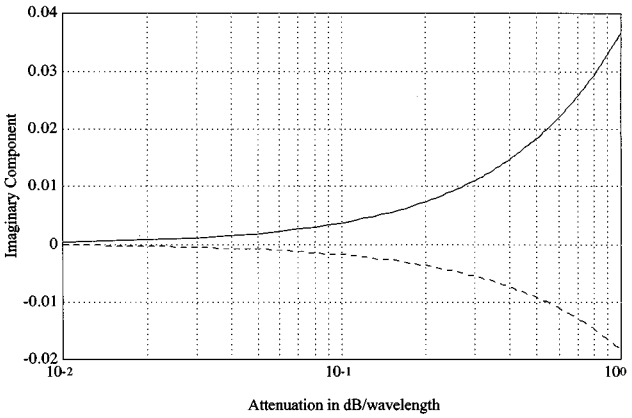


FIG. 1. The imaginary components of normalized sound speed and density as a function of medium loss in dB/wavelength. The solid line corresponds to density and the broken line corresponds to sound speed.

As can be seen from the figure, the proportional effect of loss on the density is greater than that on the sound speed. In the following sections we will show how this changes the plane-wave reflection coefficient, and the Green's function for propagation in a range-independent fluid channel with a pressure release surface and a lossy, fluid basement.

F. The plane-wave reflection coefficient

In what follows we will adopt the notation that V_c denotes the reflection coefficient for lossy fluids accounting for complex sound speed, but not complex density (this being the standard representation for losses) and V_{cp} accounts for both complex variables. The reflection coefficient for the interface in a two-layered fluid system with absorption in the lower half-space is found by requiring that the field due to a plane-wave incident on the interface satisfy certain boundary conditions. The requisite conditions are the continuity of pressure and component of particle velocity normal to the interface. We can express these conditions in terms of the velocity potential, defined by $-\nabla\psi=\mathbf{v}$. From Eq. (15) we have $p=i\omega\overline{\rho}\overline{\psi}$ and the boundary conditions can be written as

$$\left.\begin{aligned}\frac{\partial\psi_1}{\partial z}&=\frac{\partial\psi_2}{\partial z}\\ \rho_1\psi_1&=\overline{\rho_2}\psi_2\end{aligned}\right\}\ \text{at the interface,} \tag{25}$$

where the interface normal is parallel to the z axis, the subscripts 1, 2 respectively denote values in the upper and lower fluids, and we have assumed the upper fluid to be lossless so that $\overline{\rho_1}=\rho_1$. Complex sound speed enters the problem by requiring the velocity potential to satisfy the reduced wave equation. Expressing the incident, reflected and transmitted velocity potentials in the form of plane waves, applying the above boundary conditions at $z=0$ and some algebraic manipulation results in an expression for the ratio of the reflected and incident velocity potentials at the interface:

$$V_{cp}=\frac{\sin(\theta_1)/(c_1\rho_1)-\sin(\theta_2)/(\overline{c_2\rho_2})}{\sin(\theta_1)/(c_1\rho_1)+\sin(\theta_2)/(\overline{c_2\rho_2})}, \tag{26}$$

where θ is the grazing angle. The interface boundary conditions yield the additional requirement that

$$\cos(\theta_1)/c_1 = \cos(\theta_2)/\overline{c_2}. \tag{27}$$

Equation (27) is Snell's law. Equation (26) can be rewritten as

$$V_{c\rho} = \frac{m_\rho \beta \sin(\theta_1) - \sqrt{\beta \cos(\theta_c)^2 - \cos(\theta_1)^2}}{m_\rho \beta \sin(\theta_1) - \sqrt{\beta \cos(\theta_c)^2 - \cos(\theta_1)^2}}, \tag{28}$$

where we have made the substitutions $m_\rho = \rho_2/\rho_1$ and $\cos(\theta_c) = c_1/c_2$.

Since $\alpha \ll 1$ for realistic values of absorption, we can determine an approximate expression for the effect of absorption on the reflection coefficient by forming a Taylor series expansion of $V_{c\rho}$ and V_c to first order in α.

The resulting expressions are

$$V_c = V\left(1 - \frac{i \sin(\theta_1)\cos^2(\theta_c) m_\rho \alpha}{(m_\rho^2 \sin^2(\theta_1) + \cos^2(\theta_1) - \cos^2(\theta_c))(\cos^2(\theta_c) - \cos^2(\theta_1))^{1/2}}\right) \tag{29}$$

and

$$V_{c\rho} = V\left(1 + \frac{i \sin(\theta_1)(\cos^2(\theta_c) - 2\cos^2(\theta_1)) m_\rho \alpha}{(m_\rho^2 \sin^2(\theta_1) + \cos^2(\theta_1) - \cos^2(\theta_c))(\cos^2(\theta_c) - \cos^2(\theta_1))^{1/2}}\right), \tag{30}$$

where V is the lossless plane-wave reflection coefficient [given by Eq. (26) with $\beta = 1$]. The expansions are valid for all angles except the critical grazing angle, are least accurate near the critical angle and improve in accuracy for a fixed absorption as the critical angle increases. The behavior of the expansions is determined by the value of the radical in the denominator. For $\theta_1 < \theta_c$ the radical is imaginary and the factor multiplying the reflection coefficient is real and less than unity. For $\theta_1 > \theta_c$ the radical is real and the main effect is a change in phase.

A plot of the amplitude and phase of V (dash-dot line), V_c (dashed line), and $V_{c\rho}$ (solid line) for a typical shallow water seabed are given in Figs. 2 and 3. The fluid parameters for the figures are $c_1/c_2 = 15/17$, $\rho_2/\rho_1 = 1.5$, and $\Omega = 0.5$ dB/wavelength corresponding to $\alpha = 0.009$. These parameters are the same as the Acoustical Society of America benchmark wedge[19] with a lossy basement. The effect of absorption is to decrease the magnitude of the reflection coefficient for angles less than the critical, and remove the discontinuity in slope at the critical angle.

To illustrate the differences between V_c and $V_{c\rho}$, we have plotted the ratio $|V_{c\rho}/V_c|$ calculated using the above parameters as a function of grazing angle (see Fig. 4). The solid line is an exact result and the dashed line was calculated using the Taylor series expansions (20) and (21). The expansions are singular at the critical angle ($\theta_c \approx 0.49$ rad) and this point has been removed from the plot. The most significant effect of complex density occurs for angles less than the critical. In this region the effect is a reduction in the magnitude of the reflection coefficient computed using the complex sound speed only. The maximum reduction in magnitude is about 2% for the example considered. As we will demonstrate in the following section, a reduction of this size has a significant effect on the absorption of acoustic modes in waveguides.

An interesting result is that the magnitude of the reflection coefficient for large grazing angles is actually increased over its value computed using the complex sound speed only. This marginal increase results in a small decrease in the absorption of the continuous spectrum.

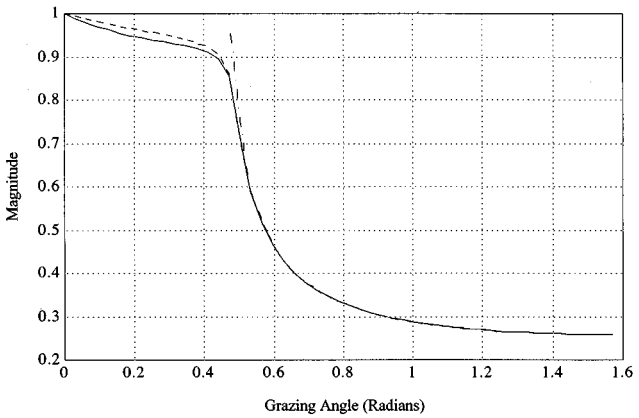


FIG. 2. Reflection coefficient magnitude as a function of grazing angle for a lossless fluid (dash-dot line), lossy fluid with complex sound speed (dashed line) and lossy fluid with complex sound speed and complex density (solid line). The fluid parameters are listed in the text.

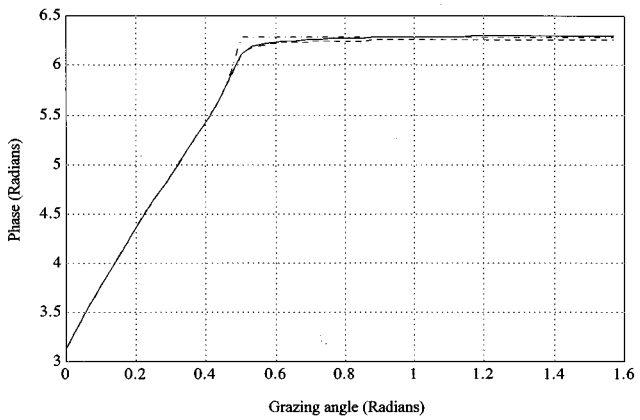


FIG. 3. Reflection coefficient phase as a function of grazing angle for a lossless fluid (dash-dot line), lossy fluid with complex sound speed (dashed line), and lossy fluid with complex sound speed and complex density (solid line). The fluid parameters are listed in the text.

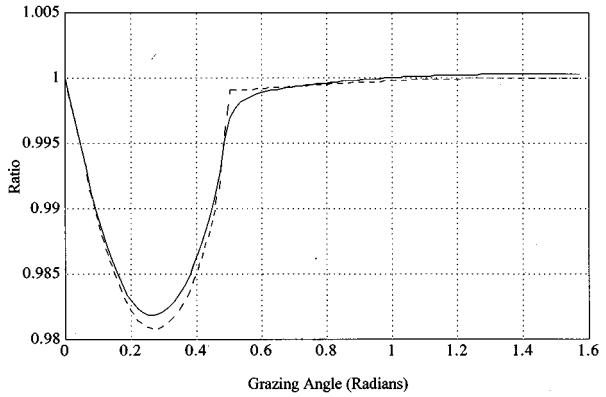


FIG. 4. Ratio of reflection coefficient magnitudes $|V_{cp}/V_c|$ as a function of grazing angle. The fluid parameters are listed in the text.

II. THE GREEN'S FUNCTION FOR PROPAGATION IN A WAVEGUIDE WITH A LOSSY BASEMENT

The dispersion relation for mode eigenvalues in a range-independent waveguide can be expressed as (see Brekhovskikh and Lysanov,[11] equation 5.5.2)

$$1+R_b\,\exp(2i\nu h)=0, \tag{31}$$

where h is the depth of the waveguide, ν is the mode vertical wave number (which we wish to solve for) and R_b is the plane-wave reflection coefficient at the water–seafloor boundary. The advantage of starting with the dispersion relation in this form is we can account for the effect of frictional losses via our modified expression for the plane-wave reflection coefficient. Making the substitution $R_b=V_{c\rho}$, the mode eigenvalue equation becomes

$$V_{c\rho}=-\exp(-2i\nu h). \tag{32}$$

Substitution of Eq. (26) for $V_{c\rho}$ into the above expression and some algebra results in the transcendental equation

$$\tan(\nu h)=\frac{-i\overline{\rho_2}\nu}{\rho_1\sqrt{k_2^2-k_1^2+\nu^2}}, \tag{33}$$

where k_1 and k_2, respectively, are the wave number in the waveguide and basement. With β set to unity, and a change of sign on the right-hand side, Eq. (33) reduces to the result given by Pekeris.[20] The change of sign is a result of our choosing the transform kernel $\exp(-i\omega t)$ rather than Pekeris' $\exp(i\omega t)$ when transforming the time-domain linear acoustic equations. Equation (33) differs from the dispersion relation used by Kornhauser and Raney[14] to analyze the effect of absorption in that the ratio of real densities has been modified to include the factor β. This modification is a direct consequence of including the internal friction loss term in the equation of motion when satisfying the boundary conditions at the waveguide fluid–fluid interface. Equation (33) yields values for the mode horizontal wave number via the relation

$$\xi=\sqrt{k_1^2-\nu^2}. \tag{34}$$

Kornhauser and Raney investigated the effect of absorption in the seafloor by performing a perturbation analysis of the lossless eigenvalue equation about its real-valued solutions in terms of the imaginary part of the vertical wave number. Compressional wave absorption was modeled in their analysis by making the sound speed in the seafloor complex. The analysis demonstrated that $\mathrm{Im}(\xi)$, which governs the absorption rate of modes in the horizontal, is proportional to α. This result can be expressed in the form of a dimensionless function F relating $\mathrm{Im}(\xi)$ to $\mathrm{Im}(k_2)$. Note that $k_2=\omega/\overline{c_2}=\omega/c_2\beta^{1/2}\approx\omega/c_2(1+i\alpha/2)$. The function is expressed in terms of the mode number n and the dimensionless variable $\sigma=\omega h/c_1$, so that

$$\mathrm{Im}(\xi)=\mathrm{Im}(k_2)F(n,\sigma). \tag{35}$$

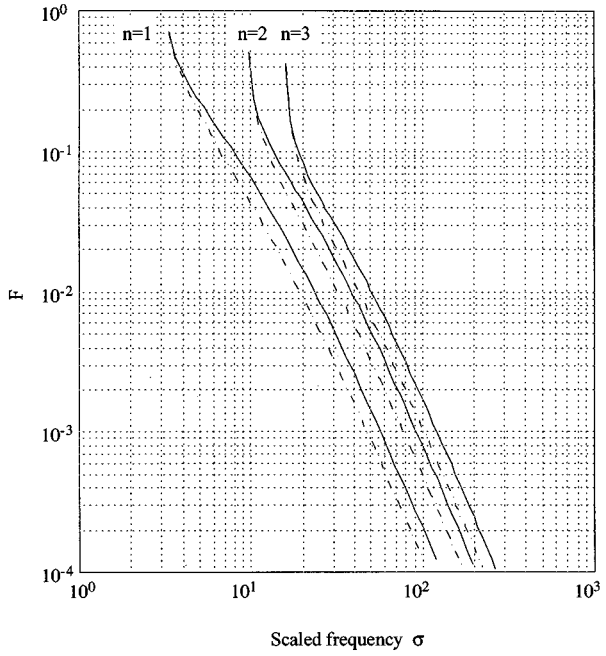


FIG. 5. $F(n,\sigma)$ vs σ for $n=1$–3 with $c_1/c_2=1.15$ and $\rho_2/\rho_1=2$. The broken line is Kornhauser and Raneys' results and the solid line includes the effect of complex density.

We have repeated Kornhauser and Raneys' analysis taking into account both complex sound speed and complex density. The analysis is much the same, but begins with the dispersion relation including complex density, Eq. (33). The result of including complex density is to multiply the function F by the factor

$$\Theta=1+\frac{2}{n_c^2\sigma^2}\,[\sigma^2(1-n_c^2)-x_n^2], \tag{36}$$

where $n_c=c_1/c_2$ and x_n is the real part of the vertical wave number [found by solving Eq. (33) with $\beta=1$] multiplied by h. The effect of including this factor can be found in Fig. 5, where we have plotted F as a function of σ for modes 1, 2, and 3. We used the same geoacoustic parameters as Kornhauser and Raney,[14] which were $c_2/c_1=1.15$ and $\rho_2/\rho_1=2$.

To further investigate the effect of complex density on the Green's function we have done a double Taylor series expansion of the dispersion relation in terms of the vertical wave number and absorption. Equation (33) can be written as[21]

$$\nu = \frac{m\pi}{h} - \arctan\left(\frac{\overline{\rho_2}\nu}{\rho_1 h\sqrt{k_1^2 - k_2^2}}\right) \tag{37}$$

and the inverse tangent function expanded in a Taylor series and solved for ν to yield the approximate solution

$$\nu \approx \frac{m\pi}{h}\left(1 + \frac{\overline{\rho_2}}{\rho_1 h\sqrt{k_1^2 - k_2^2}}\right)^{-1}, \tag{38}$$

where $m = 1,2,3,...$ is the mode number.

Equation (38) is most accurate for small vertical wave numbers, corresponding to low mode number. The effects of complex sound speed and complex density on ξ can be investigated for small values of absorption by substituting Eq. (38) into (34) and expanding the resulting expression in a Taylor series in α. Such an expansion leads to the result (correct to first order in α)

$$\xi = \xi_r + i(\xi_c + \xi_\rho), \tag{39}$$

where

$$\xi_r = \sqrt{k_1^2 - \frac{m^2\pi^2}{h^2\Lambda^2}}, \tag{40}$$

$$\xi_c = \frac{m^2 m_\rho \pi^2}{h^3 k_1^2 \Lambda^3 \xi_r} \frac{n_c^2}{2(1-n_c^2)^{3/2}}\,\alpha, \tag{41}$$

and

$$\xi_r = \frac{m^2 m_\rho \pi^2}{h^3 k_1^2 \Lambda^3 \xi_r} \frac{1}{(1-n_c^2)^{1/2}}\,\alpha. \tag{42}$$

To simplify the above equations we have made the substitution

$$\Lambda = 1 + \frac{m_\rho}{k_1^2 h\sqrt{1-n_c^2}}. \tag{43}$$

The effect of complex sound speed is accounted for by Eq. (41) and of complex density by Eq. (42). The overall effect is to introduce an imaginary component to the horizontal wave number, resulting in absorption of the mode for increasing horizontal distance form the source. As can be seen from the expansions, the effect of complex density is of the same order as the effect of complex sound speed. The terms are equal for $c_2 = \sqrt{3/2}c_1$, or $c_2 = 1837$ m/s if $c_1 = 1500$ m/s.

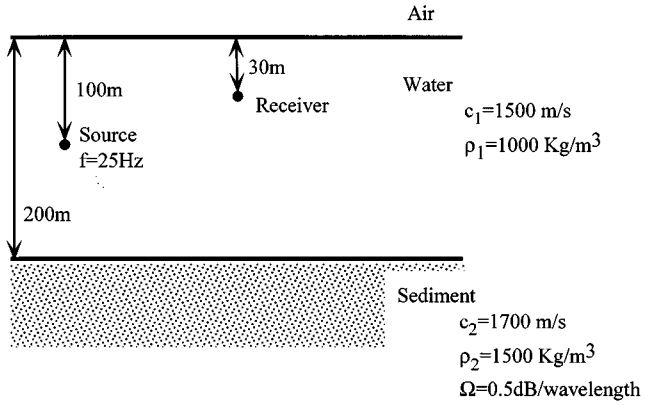


FIG. 6. Geometry of the lossy Pekeris wave guide used for the mode eigenvalue and Green's function calculations.

TABLE I. Real and imaginary parts of complex horizontal wave number and mode attenuation in dB/km as a function of mode number. Case 1: Complex sound speed only.

m	Real(k)	Imag(k)	loss dB/km
1	0.10382	0.0000072977	0.0634
2	0.10105	0.0000279867	0.2431
3	0.09622	0.0000735192	0.6386
4	0.08913	0.0009392614	8.1583
5	0.07736	0.0023749482	20.628

To illustrate the effect we have calculated mode eigenvalues for the geometry shown in Fig. 6. The eigenvalues were calculated by an iterative solution of Eq. (33) rather than using the approximate expansions. The geoacoustic parameters are as for the Acoustical Society of America benchmark wedge[19] with a lossy basement. There are three propagating modes present. The real and imaginary parts of the complex horizontal wave number for the first five modes are listed in Tables I (without complex density) and II (with complex density). The mode absorption expressed in terms of decibels per kilometer are also listed. Including complex density increases the absorption of the modes, by 50% for the first mode and 40% for the second. The effect is somewhat less for the third mode (17%), and the evanescent modes are almost unchanged.

The propagation loss functions as a function of range for the geometry in Fig. 6 and computed using V_c (solid line) and $V_{c\rho}$ (dashes) with $\Omega = 0.5$ dB/wavelength are plotted in Fig. 7. There is a clear difference between the two plots, amounting to a few decibels over the 50-km track. The difference, corresponding mainly to the mean amplitude of the pressure field, is consistent with the mode eigenvalues listed in Tables I and II.

III. THE INTERPRETATION OF ACOUSTIC ATTENUATION MEASUREMENTS

We have added a parameter to the model for sound absorption by including two loss mechanisms in the linear acoustic equations. The introduction of an additional parameter naturally leads to the question of how it can be determined. For propagation in a homogeneous medium without boundaries, the effect of both mechanisms can be accounted for by combining the parameters into a single complex sound speed [cf. Eq. (19) and Eq. (20) with $\overline{\rho}$ constant everywhere].

TABLE II. Real and imaginary parts of complex horizontal wave number and mode attenuation in dB/km as a function of mode number. Case 2: Complex sound speed and complex density.

m	Real(k)	Imag(k)	loss dB/km
1	0.10382	0.0000111391	0.0968
2	0.10105	0.0000390023	0.3388
3	0.09622	0.0000860930	0.7478
4	0.08914	0.0009364452	8.1339
5	0.07741	0.0023677485	20.566

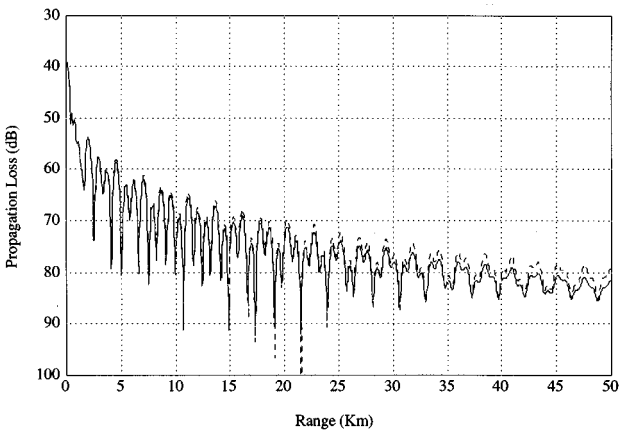


FIG. 7. Propagation loss versus range for the geometry of Fig. 6. The solid line includes the effect of both complex sound speed and complex density. The dashed line includes complex sound speed only.

In this case there is no way (and no need) to distinguish between the two mechanisms. For problems which involve boundaries between lossy media, such as the lossy Pekeris waveguide, internal friction results in different mode absorption values as a function of mode number than are obtained from internal relaxation. Thus the measurement of relative mode absorption may provide a means of distinguishing between the two mechanisms.

IV. CONCLUDING REMARKS

There are two mechanisms associated with compressional wave absorption, internal relaxation, and internal friction. Internal relaxations are modeled by modifying the equation of state for the lossy medium whereas internal frictions modify the equation of motion. For the special case of propagation in an homogeneous, unbounded medium, complex density is unimportant and both mechanisms can be modeled with complex sound speed. For the more realistic case of propagation in a stratified environment, complex sound speed does not give a correct description for losses caused by internal friction. In this case, frictional losses can be modeled by the introduction of a complex quiescent density in addition to complex sound speed.

An important application to ocean acoustics is the calculation of transmission loss in shallow water regions. In these regions the effect of complex density is commensurate with the effect of complex sound speed, and increases the absorption of propagating modes. Assuming that all absorption is due to internal friction losses, low order modes can have their absorption increased significantly (as much as 50% for the example given) and neglecting complex density may lead to amplitude errors of several decibels. Energy conservation corrections which eliminate amplitude errors in range-dependent propagation models have recently been published.[22,23] The complex density correction for the case of frictional losses is similar in nature in that the main effect is to shift the level of the transmission loss curves.

There is insufficient evidence at this time to decide if an equivalent fluid description of absorption in marine sediments is correctly formulated by a modification to the equation of state or the equation of motion. As propagation in realistic ocean environments depends somewhat on which modification is adopted, it is important to determine the origin of the losses.

ACKNOWLEDGMENTS

I am pleased to acknowledge my hosts at The Naval Research Laboratory, Washington, DC. I was a visiting scientist at NRL when this work was completed. I also wish to acknowledge helpful discussions with a number of colleagues, particularly Professor Michael Buckingham at Scripps Institution of Oceanography and Dr. Michael Collins at NRL. This work has been supported by the Office of Naval Research, Grant No. N00014-91-J-1118.

[1] E. L. Hamilton, ''Acoustic properties of sediments,'' in *Acoustics and the Ocean Bottom* (CSIS, Madrid, 1987).
[2] J. E. White, *Underground Sound: Application of Seismic Waves* (Elsevier, Amsterdam, 1983), p. 253.
[3] M. N. Tokösoz, D. H. Johnston, and A. Timur, ''Attenuation of seismic waves in dry and saturated rocks. I. Laboratory measurements,'' Geophysics **44**, 681–690 (1979).
[4] D. H. Johnston, M. N. Toksöz, and A. Timur, ''Attenuation of seismic waves in dry and saturated rocks, II. Mechanisms,'' Geophysics **44**, 691–711 (1979).
[5] A. C. Kibblewhite, ''Attenuation of sound in marine sediments: A review with emphasis on new low frequency data,'' J. Acoust. Soc. Am. **86**, 716–738 (1989).
[6] R. D. Stoll, *Sediment Acoustics* (Springer-Verlag, Berlin, 1989), Lecture Notes on Earth Sciences, p. 153.
[7] M. A. Biot, ''Theory of propagation of elastic waves in a fluid-saturated porous solid: I. Low-frequency range,'' J. Acoust. Soc. Am. **28**, 168–178 (1956).
[8] M. A. Biot, ''Theory of propagation of elastic waves in a fluid-saturated porous solid: II. Higher frequency range,'' J. Acoust. Soc. Am. **28**, 179–191 (1956).
[9] M. A. Biot, ''Generalized theory of acoustic propagation in porous dissipative media,'' J. Acoust. Soc. Am. **33**, 1254–1264 (1962).
[10] R. D. Stoll and G. M. Bryan, ''Wave attenuation in saturated sediments,'' J. Acoust. Soc. Am. **47**, 1440–1447 (1969).
[11] L. M. Brekhovskikh and Y. Lysanov, *Fundamentals of Ocean Acoustics* (Springer-Verlag, Berlin, 1982), Springer Series in Electrophysics.
[12] G. V. Frisk, *Ocean and Seabed Acoustics: A Theory of Wave Propagation* (Prentice-Hall, Englewood Cliffs, NJ, 1994).
[13] M. D. Verweij, ''Modeling space-time domain acoustic wave fields in media with attenuation: The symbolic manipulation approach,'' J. Acoust. Soc. Am. **97**, 831–843 (1995).
[14] E. T. Kornhauser and W. P. Raney, ''Attenuation in shallow-water propagation due to an absorbing bottom,'' J. Acoust. Soc. Am. **27**, 689–692 (1955).
[15] R. M. Morse, ''Acoustic propagation in granular media,'' J. Acoust. Soc. Am. **24**, 696–700 (1952).
[16] A. D. Pierce, *Acoustics: An Introduction to its Physical Principals and Applications* (McGraw-Hill, New York, 1981).
[17] P. G. Bergmann, ''The wave equation in a medium with a variable index of refraction,'' J. Acoust. Soc. Am. **17**, 329–333 (1946).
[18] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).
[19] L. Felsen, ''Benchmarks: An option for quality assessment, '' J. Acoust. Soc. Am. **87**, 1497–1498 (1980).
[20] C. L. Pekeris, ''Theory of propagation of explosive sound in shallow water,'' Geological Soc. Am. Memorandum **27**, 1–117 (1948).
[21] M. J. Buckingham, ''A theoretical model of ambient noise in a low-loss, shallow water channel,'' J. Acoust. Soc. Am. **67**, 1186–1192 (1980).
[22] M. D. Collins, ''An energy conserving parabolic equation for elastic media,'' J. Acoust. Soc. Am. **94**, 975–982 (1993).
[23] B. P. Porter, F. B. Jensen, and C. M. Ferla, ''The problem of energy conservation in one-way models,'' J. Acoust. Soc. Am. **89**, 1058–1067 (1991).

Robust maximum-likelihood source localization in an uncertain shallow-water waveguide

J. Tabrikian and J. L. Krolik
Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina 27708-0291

H. Messer
Department of Electrical Engineering–Systems, Tel-Aviv University, Tel-Aviv 69978, Israel

(Received 23 January 1996; accepted for publication 13 September 1996)

This paper addresses the problem of matched-field source localization in the presence of uncertainties in the ocean environment. Because signal wavefront mismatch can cause anomalous source location estimates, development of robust localization methods is critically important. In this paper, a robust maximum-likelihood estimator is proposed. It is based on a decomposition of the field into predictable and unpredictable subspaces of the acoustic normal mode representation. The estimator uses the predictable subspace for source localization. Identification of the predictable modes is made according to the second-order joint statistics of the horizontal wave numbers. The performance of the method is evaluated and compared to other matched-field methods using simulations and acoustic array data from the Mediterranean Sea. In the presence of mismatches, the algorithm has superior probability of correct localization than the maximum-likelihood, matched-mode-processing, and Bartlett methods. © *1997 Acoustical Society of America.* [S0001-4966(97)04701-2]

PACS numbers: 43.30.Wi, 43.60.Gk [MGP]

INTRODUCTION

Matched-field methods exploit complex multipath propagation models to localize underwater acoustic sources. As demonstrated in several articles,[1,2] a major difficulty facing this approach is that the localization process is extremely sensitive to errors in the assumed environmental conditions.

Several approaches to this problem have been investigated. In one approach, the problem is formulated by assuming a parametrized model in which the vector of the environmental and source location parameters is unknown. In the optimum uncertain field processor (OUFP),[3,4] prior statistical information about the source location, propagation, and array parameters is used in order to obtain optimal estimates of the source location parameters. This processor involves a computationally intensive integration over the model parameter space. In the maximum-likelihood (ML) estimator, a multidimensional search over the parameter space is required. In most situations the model parameter space is extremely large, thus solution via an exhaustive search is not a viable option. Furthermore, the objective function usually contains many local minima/maxima precluding the use of gradient descent methods. Therefore, this approach has been investigated mainly using simulated annealing and genetic algorithms.[5,6]

Two other approaches for matched-field processing methods which are robust to environmental uncertainty, are: (1) minimum variance beamformer with environmental perturbation constraints (MV-EPC),[4,8] and (2) matched mode processing (MMP) methods.[7,9–12] In the MV-EPC method, the second-order statistics of the signal wavefront, averaged over an ensemble of environmental perturbations, are used to design constraints for an MV adaptive beamformer. The robustness of the MMP methods is achieved by retaining only the part of the field which is less sensitive to environmental mismatch. As a first step, the MMP requires estimation of all the mode amplitudes. Then estimation of the source location is achieved by matching only the predictable modes to the corresponding modes in the replica field. Its main drawback is the difficulty of the preliminary step in making accurate estimates of all the modal amplitudes, particularly at lower signal-to-noise-ratios (SNR's). In addition, it requires that the number of modes be lower than the number of sensors.

In this paper, a robust version of the ML estimator for source range and depth is proposed. The method is based on the conventional ML estimator,[13,14] but with a different signal wavefront model. As in MMP, it is also based on a decomposition of the field into predictable and unpredictable subspaces of the acoustic normal mode representation of the field. In the proposed estimator, instead of neglecting the unpredictable modes, the measurements are projected into the unpredictable null subspace. The performance of the proposed method is evaluated and compared to other existing methods by simulations and experimental data. It is shown that the robust ML achieves higher probability of correct localization than the ML, MMP, and Bartlett methods.

The paper is organized as follows: The next section describes the scenario and the problem formulation. In Sec. II, the proposed estimator is presented. In Sec. III, the Cramer–Rao lower bound (CRLB) is developed to demonstrate the dependence of localization accuracy on the number of reliable modes. Section IV provides a discussion on identification of the predictable modes. Sections V and VI present the results of a performance evaluation of the proposed method

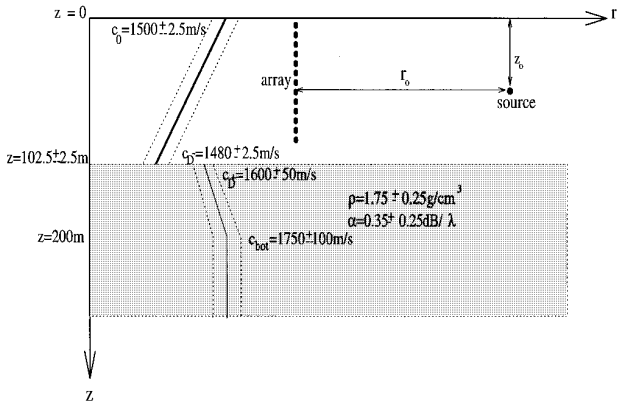


FIG. 1. The NRL workshop ''genlmis'' scenario configuration.

using simulations and experimental data, respectively. The conclusions are summarized in Sec. VII.

I. PROBLEM FORMULATION

Consider a point source at depth z_0 and range r_0 which radiates a monochromatic signal at angular frequency ω in a time invariant shallow-water waveguide. The acoustic field is sampled by a vertical array of N sensors. The depth of the ith sensor from the upper surface is denoted by z_i. The sensor locations are assumed to be known. Figure 1 depicts the environmental configuration and the source-array geometry. The environmental scenario is one of the more complex benchmark cases used in the May 1993 NRL Workshop on Acoustic Models in Signal Processing.[15]

Using an adiabatic normal mode model, the field measured by sensor i at time t, $(-(T/2)<t<(T/2))$ can be expressed by (see, e.g., Ref. 16):

$$\tilde{y}_i(t) \triangleq \tilde{y}(z_i,t) = b\sum_{m=1}^{M} \phi_m(z_i)\phi_m^{(0)}(z_0)\,\frac{e^{j\kappa_m r_0}}{\sqrt{\kappa_m r_0}}\,e^{j\omega t} + \epsilon_i e^{j\omega t}, \quad i=1,\ldots,N, \tag{1}$$

where $\phi_m^{(0)}(\cdot)$ and $\phi_m(\cdot)$, $m=1,\ldots,M$ are the modal depth eigenfunctions at the source and the array, respectively, and κ_m denotes the average horizontal wave number of the mth mode in the channel:

$$\kappa_m \triangleq \frac{1}{r_0}\int_0^{r_0}\kappa_m(r)\,dr, \quad m=1,\ldots,M.$$

Here, M is the number of the propagating modes in the channel and ϵ_i stands for the additive noise complex amplitude at the ith sensor. The complex amplitude of the received signal, b, is unknown. The received noise at the sensors is assumed to be zero-mean, Gaussian, with known covariance matrix, $\mathbf{Q}$, whose elements are given by

$$\mathbf{Q}_{ik} = E\{\epsilon_i \epsilon_k^*\}, \quad i,k=1,\ldots,N, \tag{2}$$

where E denotes the expectation operation. The complex Fourier coefficient of (1) at a single frequency ω is

$$y_i = b\sum_{m=1}^{M}\phi_m(z_i)\phi_m^{(0)}(z_0)\,\frac{e^{j\kappa_m r_0}}{\sqrt{\kappa_m r_0}} + \epsilon_i, \quad i=1,\ldots,N. \tag{3}$$

The modal eigenfunctions, $\phi_m(z_i)$, and horizontal wave numbers, κ_m, depend on environmental parameters which describe the bathymetry, geoacoustic properties of the bottom, and sound speed in the water column. In practice, these parameters are not precisely known. For example, the uncertain environmental parameters considered in this paper are shown in Table I. Each parameter is assumed to be uniformly distributed in the range of values given. The objective here is to estimate the source location parameters (r_0, z_0) from the measurements y_i, $i=1,\ldots,N$, when the complex signal amplitude b is unknown, and the environment is uncertain. The vector of the uncertain environmental parameters η includes: the sound velocity at depths $z=0$, D^-, D^+, D_{bot}, i.e., $c(0)$, $c(D^-)$, $c(D^+)$, $c(D_{\text{bot}})$, the waveguide depth D, the bottom attenuation α and the bottom density ρ.

Denoting the vector of the source parameters by $\mathbf{u}$: $\mathbf{u} \triangleq [r_0 \quad z_0]^T$, Eq. (3) can be written in matrix notation as

$$\mathbf{y} = b\mathbf{T}\mathbf{s}(\mathbf{u}) + \boldsymbol{\epsilon}, \tag{4}$$

where the elements of the matrix $\mathbf{T}$ and the vector $\mathbf{s}(\mathbf{u})$ are given by $T_{im} = \phi_m(z_i)$, $s_m(\mathbf{u}) = \phi_m^{(0)}(z_0)e^{j\kappa_m r_0}/\sqrt{\kappa_m r_0}$, $i=1,\ldots,N$, $m=1,\ldots,M$, and $b=[1 \quad j]\mathbf{b}$, where $\mathbf{b} \triangleq [\text{Re}(b) \quad \text{Im}(b)]^T$. The superscript T denotes the transpose operation.

II. THE ROBUST MAXIMUM LIKELIHOOD ESTIMATOR

Applying the conventional ML estimator to the problem stated above directly involves a multidimensional search over the source location and the environmental parameters $(\mathbf{u}, \eta)$. In order to avoid such a computationally intensive search procedure, the method proposed here identifies the part of the field that is less predictable due to environmental mismatch. The revised model ignores the information on source location carried by the unpredictable part and the localization is performed according to the information carried by the predictable part. Applying the ML estimator with this model yields a robust and computationally efficient procedure.

A. Model decomposition

In normal mode models, some modal amplitudes remain more correlated than others in the presence of environmental uncertainties. One of the larger uncertainties in acoustic propagation modeling is the effect of boundary interactions. Typically, higher-order modes have more significant coupling to the boundary and are thus less predictable. In typical shallow-water channels, the higher-order modes are associated with rays that have a higher number of reflections from

TABLE I. Uncertain shallow-water environmental parameters.

Environmental parameter	Symbol	Value range
Surface sound speed	$c(0)$	1500.0 ± 2.5 m/s
Bottom sound speed	$c(D^-)$	1480.0 ± 2.5 m/s
Bottom depth	D	102.5 ± 2.5 m
Sub-bottom sound speed	$c(D^+)$	1600 ± 50 m/s
Lower halfspace sound speed	$c(D_{\text{bot}})$	1750 ± 100 m/s
Bottom attenuation	α	0.35 ± 0.25 dB/λ
Bottom density	ρ	1.75 ± 0.25 g/cm^3

the boundaries and larger group delays in the medium.[17] Therefore, these modes are often affected by medium variations more than the lower-order modes. In other words, uncertainties in the channel parameters cause the higher-order modes to be less correlated than the lower-order modes. This means that the available information on the source location is weaker in the higher modes than the lower modes. The MMP techniques in[10,12] were developed relying on this assumption.

In this work, the signal wavefront is decomposed into two subspaces. The first subspace includes the modes which remain correlated in the presence of environmental variability. These modes will be referred to as ''predictable'' modes. The other modes, referred to as ''unpredictable,'' make up the second subspace. Identification of the predictable modes will be discussed in Sec. IV. In the proposed model, the complex amplitudes of the modes in the second subspace are assumed to be nuisance parameters. It is assumed that these modes do not carry any useful information concerning the source location parameters and the localization is performed according to the information carried by the modes in the first subspace. These assumptions lead to a suboptimal but computationally efficient solution.

Assume that M_0 modes, constructing the space Ω_p, are predictable while the $M-M_0$ modes in the space Ω_u are unpredictable. Then the field can be written as a sum of predictable and unpredictable parts:

$$y_i = b\sum_{m\in\Omega_p} s_m(\mathbf{u})\phi_m(z_i) + b\sum_{m\in\Omega_u} s_m\phi_m(z_i) + \epsilon_i . \qquad (5)$$

In the first term, the functional relationship between the coefficients $\{s_m(\mathbf{u})\}_{m\in\Omega_p}$ and the source location is known *a priori*. In the second part, the coefficients $\{s_m\}_{m\in\Omega_u}$ have a weak or essentially unknown relationship to the source location and environmental parameters. In order to get a robust estimate in an imperfectly known ocean waveguide, the coefficients $\{s_m\}_{m\in\Omega_u}$ are modeled as nuisance parameters. Equation (5) can be written in vector notation in the form of Eq. (4) by

$$\mathbf{y} = b\mathbf{T}_1\mathbf{s}_1(\mathbf{u}) + \mathbf{T}_2\mathbf{q} + \boldsymbol{\epsilon} = \underbrace{[\mathbf{T}_1\mathbf{s}_1(\mathbf{u}) \quad \mathbf{T}_2]}_{\mathbf{A}(u)}\begin{bmatrix} b \\ \mathbf{q}\end{bmatrix} + \boldsymbol{\epsilon}, \qquad (6)$$

where $\mathbf{T}=[\mathbf{T}_1 \quad \mathbf{T}_2]$, in which $\mathbf{T}_1$ and $\mathbf{T}_2$ are $N\times M_0$ and $N\times(M-M_0)$ matrices, respectively, and $\mathbf{s}^T(\mathbf{u}) = [\mathbf{s}_1^T(\mathbf{u}) \quad \mathbf{q}^T/b]$, where $\mathbf{s}_1(\mathbf{u})$ and $\mathbf{q}$ are $M_0\times 1$ and $(M-M_0)\times 1$ vectors, respectively. The vector $\mathbf{q}$ contains the nuisance parameters $\{s_m\}_{m\in\Omega_u}$. We assume that all the eigenfunctions $\phi_m(z)$ at the receiving array are known, i.e., the matrices $\mathbf{T}_1$ and $\mathbf{T}_2$ are known. The justification for this assumption is that environmental measurements are typically available at the receiver. In contrast, the errors in $\mathbf{q}$ are due to uncertainties in the horizontal wave numbers, κ_m, which are determined by variable channel properties between the source and the receiver. Since horizontal wave numbers are multiplied by the source range r_0 to determine the model phases in (3), errors in $\mathbf{q}$ are usually the most significant reason for signal wavefront mismatch.

B. The ML estimator

Based on the model described in (6), the ML estimator of the vector $\mathbf{u}$ is given by

$$\hat{\mathbf{u}}_{\mathrm{ML}} = \arg\min_{\mathbf{u}}\{\min_{\mathbf{b},\mathbf{q}}[\mathbf{y}-b\mathbf{T}_1\mathbf{s}_1(\mathbf{u})-\mathbf{T}_2\mathbf{q}]^H\mathbf{Q}^{-1} \times[\mathbf{y}-b\mathbf{T}_1\mathbf{s}_1(\mathbf{u})-\mathbf{T}_2\mathbf{q}]\}, \qquad (7)$$

where the elements of $\mathbf{Q}$ are defined in (2). The linear dependence upon the nuisance parameter vector considerably simplifies the ML estimator. By minimization over $[b \quad \mathbf{q}^T]$ one obtains

$$\begin{bmatrix}\hat{b}\\ \hat{\mathbf{q}}\end{bmatrix}_{\mathrm{ML}} = (\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-1}\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{y}, \qquad (8)$$

where $\mathbf{A}\triangleq[\mathbf{T}_1\mathbf{s}_1 \quad \mathbf{T}_2]$. The ML estimator of $\mathbf{u}$ becomes

$$\begin{aligned}\hat{\mathbf{u}}_{\mathrm{ML}} &= \arg\min_{\mathbf{u}}\{(\mathbf{y}-\mathbf{A}(\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-1}\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{y})^H\mathbf{Q}^{-1} \times(\mathbf{y}-\mathbf{A}(\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-1}\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{y})\} \\ &= \arg\min_{\mathbf{u}}\|(\mathbf{P}_A^{\perp}(\mathbf{u})\mathbf{Q}^{-1/2}\mathbf{y})\|^2, \end{aligned} \qquad (9)$$

where $\mathbf{P}_{\mathbf{A}}^{\perp}(\mathbf{u})$ is a projection matrix into the null space of $\mathbf{Q}^{-1/2}\mathbf{A}(\mathbf{u})$:

$$\mathbf{P}_{\mathbf{A}}^{\perp}(\mathbf{u}) \triangleq \mathbf{I}-\mathbf{P}_{\mathrm{A}}(\mathbf{u}) = \mathbf{I}-\mathbf{Q}^{-1/2}\mathbf{A}(\mathbf{u}) \times(\mathbf{A}^H(\mathbf{u})\mathbf{Q}^{-1}\mathbf{A}(\mathbf{u}))^{-1}\mathbf{A}^H(\mathbf{u})\mathbf{Q}^{-1/2}, \qquad (10)$$

where the matrix $\mathbf{I}$ is an identity matrix of size N.

This estimator requires the number of sensors N, be greater or equal to $M-M_0+1$, in contrast with the MMP which requires $N>M$. As mentioned above, the proposed estimator is more robust than the regular ML estimator because it properly nulls modes which are the most sensitive to environmental mismatch. By projection of the data and the model into the null space of these modes, the information on the source location carried by these modes is lost. Therefore, the performance of the estimator can be expected to degrade as the number of the predictable modes decreases. This means that there is a trade-off between robustness to environmental mismatch and the local accuracy of the estimator. This trade-off will be studied via the CRLB in Sec. III. The robust ML is similar to MMP[12] which is based on ''matching'' the predictable modes to the measured modes to obtain a robust estimate of the source location. However unlike the MMP, it does not require making accurate estimates of all the modal amplitudes which becomes difficult at lower SNR's and instead of ignoring the unpredictable space it properly nulls it.

III. THE CRAMER–RAO BOUND ON SOURCE LOCATION PARAMETERS

To illustrate the trade-off between robustness and accuracy, the non-Bayesian Cramer–Rao lower bound (CRLB) is derived here for the model of (6). If the environmental parameters are known, this bound describes the performance of the ML estimator at sufficiently large SNR and/or observation time.

The CRLB is an indication of the minimum variance of any unbiased estimator of the vector of unknown parameters given by

$$\theta \triangleq [\alpha^T, \mathbf{q}^T]^T, \tag{11}$$

where $\alpha^T \triangleq [\mathbf{u}^T, \mathbf{b}^T]$. For a deterministic signal and Gaussian noise, in a model of the general form, $\mathbf{y}=\mathbf{G}(\theta)+\boldsymbol{\epsilon}$, the entries of the Fisher information matrix (FIM) of the vector θ are given by

$$[J_{\theta\theta}]_{ij} = 2\,\mathrm{Re}\left[\frac{\partial \mathbf{G}^H(\theta)}{\partial \theta_i}\,\mathbf{Q}^{-1}\,\frac{\partial \mathbf{G}(\theta)}{\partial \theta_j}\right]. \tag{12}$$

Partitioning $J_{\theta\theta}$ according to (11), and using the model in (6), gives

$$J_{\theta\theta} = 2\begin{bmatrix} \mathrm{Re}\left(\dfrac{\partial b^* \mathbf{s}_1^H(\alpha)}{\partial \alpha}\,\mathbf{T}_1^H\mathbf{Q}^{-1}\mathbf{T}_1\,\dfrac{\partial b\mathbf{s}_1(\alpha)}{\partial \alpha}\right) & \mathrm{Re}\left(\dfrac{\partial b^* \mathbf{s}_1^H(\alpha)}{\partial \alpha}\,\mathbf{T}_1^H\mathbf{Q}^{-1}\mathbf{T}_2\right) \\ \mathrm{Re}\left(\mathbf{T}_2^H\mathbf{Q}^{-1}\mathbf{T}_1\,\dfrac{\partial b\mathbf{s}_1(\alpha)}{\partial \alpha}\right) & \mathrm{Re}(\mathbf{T}_2^H\mathbf{Q}^{-1}\mathbf{T}_2) \end{bmatrix} \triangleq \begin{bmatrix} J_{\alpha\alpha} & J_{\alpha\mathbf{q}} \\ J_{\mathbf{q}\alpha} & J_{\mathbf{qq}} \end{bmatrix}. \tag{13}$$

The CRLB on α using the predictable modes, is given by the upper submatrix of $J_{\theta\theta}^{-1}$ which, from (13), can be expressed as

$$\mathrm{CRP}(\alpha) = J_{\alpha\alpha}^{-1} + \Delta\mathrm{CRLB}_1. \tag{14}$$

Using a well-known formula for the inverse of a partitioned matrix[18] where $\Delta\mathrm{CRLB}_1$ is given by

$$\Delta\mathrm{CRLB}_1 \triangleq J_{\alpha\alpha}^{-1}J_{\alpha\mathbf{q}}(J_{\mathbf{qq}} - J_{\mathbf{q}\alpha}J_{\alpha\alpha}^{-1}J_{\alpha\mathbf{q}})^{-1}J_{\mathbf{q}\alpha}J_{\alpha\alpha}^{-1} \geqslant 0. \tag{15}$$

The matrix $J_{\alpha\alpha}$ is the FIM for source localization assuming the field only consists of the predictable modes. Its relation to the FIM for source localization where all the modes are used is established in the sequel. The matrix $\Delta\mathrm{CRLB}_1$ expresses the performance loss resulting from the estimation of the additional nuisance parameters in the model (the vector $\mathbf{q}$). If the noise at the sensors is independent, identically distributed (i.i.d.), then $\mathbf{Q}=\sigma^2\mathbf{I}$. Further, in the case of a dense array which spans the entire water column, the acoustic modes are orthogonal, so $\mathbf{T}_2^H\mathbf{T}_1=\mathbf{0}$, and the resulting $J_{\theta\theta}$ is a block diagonal matrix with $J_{\mathbf{q}\alpha} = J_{\alpha\mathbf{q}}^T = \mathbf{0}$. In this case, $\Delta\mathrm{CRLB}_1=\mathbf{0}$, so asymptotically there is no performance loss due to the need to estimate $\mathbf{q}$.

Although in the case of a fully spanning dense array, $\Delta\mathrm{CRLB}_1$ of (15) is zero, treating the unpredictable modes as nuisance parameters still results in performance degradation relative to the case of a known environment. This loss can be studied by considering the CRLB in the perfectly matched case where

$$\mathbf{y} = b\mathbf{T}_1\mathbf{s}_1(\alpha) + b\mathbf{T}_2 s_2(\alpha) + \mathbf{n}. \tag{16}$$

This model assumes no environmental uncertainties. The second term in (16) represents the modes which are not used by the suggested robust method. Let $\mathrm{CRE}(\alpha)$ be the CRLB on α using the exact model of (16). Then, $\mathrm{CRE}(\alpha)$ is given by

$$\begin{aligned}\mathrm{CRE}(\alpha) &= \left\{2\,\mathrm{Re}\left[\left(\mathbf{T}_1\,\frac{\partial b\mathbf{s}_1(\alpha)}{\partial \alpha} + \mathbf{T}_2\,\frac{\partial b\mathbf{s}_2(\alpha)}{\partial \alpha}\right)^H \right.\right. \\ &\quad \left.\left. \times \mathbf{Q}^{-1}\left(\mathbf{T}_1\,\frac{\partial b\mathbf{s}_1(\alpha)}{\partial \alpha} + \mathbf{T}_2\,\frac{\partial b\mathbf{s}_2(\alpha)}{\partial \alpha}\right)\right]\right\}^{-1}, \\ &= (J_{\alpha\alpha} + \mathbf{E})^{-1},\end{aligned} \tag{17}$$

where

$$\begin{aligned}\mathbf{E} \triangleq 2\,\mathrm{Re}\Bigg[&\frac{\partial b^*\mathbf{s}_1^H(\alpha)}{\partial \alpha}\,\mathbf{T}_1^H\mathbf{Q}^{-1}\mathbf{T}_2\,\frac{\partial b\mathbf{s}_2(\alpha)}{\partial \alpha} \\ &+ \frac{\partial b^*\mathbf{s}_2^H(\alpha)}{\partial \alpha}\,\mathbf{T}_2^H\mathbf{Q}^{-1}\mathbf{T}_1\,\frac{\partial b\mathbf{s}_1(\alpha)}{\partial \alpha}\Bigg] \\ &+ \frac{\partial b^*\mathbf{s}_2^H(\alpha)}{\partial \alpha}\,\mathbf{T}_2^H\mathbf{Q}^{-1}\mathbf{T}_2\,\frac{\partial b\mathbf{s}_2(\alpha)}{\partial \alpha}.\end{aligned} \tag{18}$$

In general, $\mathbf{E}$ is not necessarily non-negative-definite, however, under the assumption of a dense array and i.i.d. noise, $\mathbf{E}$ can be approximated by

$$\mathbf{E} \approx 2\,\mathrm{Re}\left[\frac{1}{\sigma^2}\,\frac{\partial b^*\mathbf{s}_2^H(\alpha)}{\partial \alpha}\,\mathbf{T}_2^H\mathbf{T}_2\,\frac{\partial b\mathbf{s}_2(\alpha)}{\partial \alpha}\right] > 0. \tag{19}$$

Indeed, in this case, use of less modes for source localization results in performance reduction. Using another matrix identity:

$$\mathrm{CRE}(\alpha) = (J_{\alpha\alpha} + \mathbf{E})^{-1} = J_{\alpha\alpha}^{-1} - (J_{\alpha\alpha} + J_{\alpha\alpha}\mathbf{E}^{-1}J_{\alpha\alpha})^{-1}, \tag{20}$$

thus

$$J_{\alpha\alpha}^{-1} = \mathrm{CRE}(\alpha) + \Delta\mathrm{CRLB}_2, \tag{21}$$

where

$$\Delta\mathrm{CRLB}_2 = (J_{\alpha\alpha} + J_{\alpha\alpha}\mathbf{E}^{-1}J_{\alpha\alpha})^{-1} > \mathbf{0}. \tag{22}$$

Substituting Eq. (21) into (14) shows that modeling the unpredictable modes as a noisy component in the received data leads to a CRLB on the source localization which can be expressed as

$$\mathrm{CRP}(\alpha) = \mathrm{CRE}(\alpha) + \Delta\mathrm{CRLB}_1 + \Delta\mathrm{CRLB}_2, \tag{23}$$

where for fully spanning dense array case, $\Delta\mathrm{CRLB}_1=0$, so $\Delta\mathrm{CRLB}_2$ describes the performance loss in the robust ML relative to the exact ML for a fully calibrated channel.

In the general case of an array which does not span the water column, the CRP can be expected to increase relative to CRE, because of both the effects of:

(1) adding the unknown nuisance parameters $\mathbf{q}$ to the model,

(2) ignoring the source location information carried by the rejected modes,

which are represented by $\Delta\mathrm{CRLB}_1$ and $\Delta\mathrm{CRLB}_2$, respectively. Because in the general case, $\mathbf{E}$ in (18) is not necessarily positive semidefinite, the simple sum in (23) cannot be used directly. Nevertheless, intuitively these factors lead to a trade-off between robustness and accuracy. In other words, although the robustness is achieved by excluding the unpredictable modes from the space Ω_p, this results in a degradation in local accuracy performance.

IV. IDENTIFICATION OF THE PREDICTABLE MODES

Determination of the predictable and unpredictable components of a hypothesized signal wavefront requires the ability to identify the normal modes which are least affected by perturbations in the uncertain environmental parameters. This task involves both determining the number of predictable modes, M_0, as well as which M_0 of the M possible modes are most robust to environmental mismatch. In this section, the latter problem is addressed. Given the value of M_0, the method proposed here identifies the most robust M_0-sized subset of the M modes. The problem of selecting M_0 is briefly discussed at the end of this section.

Environmental mismatch causes errors in both the modal horizontal wave numbers and eigenfunctions. To identify the predictable modes, treat the uncertain environmental parameters as random variables. Under this assumption, the horizontal wave numbers and modal eigenfunctions in the propagation model are also random. The predictable modes can then be determined by finding those that exhibit the lowest variability. Once identified, these predictable modes are then approximated as being deterministically known in the model of Sec. II.

In a shallow-water waveguide, where the source range is much larger than the channel depth, the effect of errors in the modal horizontal wave numbers is the most significant error, since the modal phases are products of the horizontal wave numbers and the source range. Hence, a simple approximation which can be used to identify the predictable modes is to assume that the environmental mismatch causes errors primarily in the modal phases (i.e., the stable modes can be determined according to the stability of the horizontal wave numbers). In Sec. V the effect of mismatch on the modal amplitude errors will be evaluated numerically.

The correlation of the modes in the presence of environmental uncertainty can be evaluated by computing the second order statistics of the horizontal wave numbers. For this purpose, KRAKEN,[19] a normal mode propagation model program, can be used to generate a database containing independent random realizations of the uncertain environmental parameters. Assume a candidate subset of modes with horizontal wave-number vector, $\mathbf{k}$, of dimension $M_0\times 1$. Let $\mathbf{k}_l$, $l=1,\ldots,L$ be the lth realization of the vector $\mathbf{k}$. The covariance matrix of $\mathbf{k}$ can be estimated by

$$\mathbf{\Gamma}\triangleq\mathrm{cov}(\mathbf{k})\approx\frac{1}{L}\sum_{l=1}^{L}(\mathbf{k}_l-\hat{\mathbf{k}})(\mathbf{k}_l-\hat{\mathbf{k}})^H, \tag{24}$$

where $\hat{\mathbf{k}}$ is the unbiased estimator of $\overline{\mathbf{k}}\triangleq E(\mathbf{k})$:$\hat{\mathbf{k}}=1/L\Sigma_{l=1}^{L}\mathbf{k}_l$.

In general, some errors in $\mathbf{k}$ affect source range bias while others affect distortion of the ambiguity/likelihood function shape. To see this, decompose the horizontal wave numbers by letting

$$\mathbf{k}=\overline{\mathbf{k}}+\Delta\mathbf{k}=\overline{\mathbf{k}}+\Delta\mathbf{k}^{\perp}+\epsilon_1\overline{\mathbf{k}}+\epsilon_2\mathbf{1}, \tag{25}$$

where $\mathbf{1}$ is a vector of ones. The vector $\Delta\mathbf{k}^{\perp}$ is orthogonal to the vectors $\mathbf{k}$ and $\mathbf{1}$ (i.e., $\Delta\mathbf{k}^{\perp}\perp[\mathbf{k}\quad\mathbf{1}]$). Letting $\mathbf{F}$ be defined by $\mathbf{F}\triangleq[\mathbf{k}\quad\mathbf{1}]$, then, $\Delta\mathbf{k}^{\perp}$ is given by

$$\Delta\mathbf{k}^{\perp}=\Delta\mathbf{k}-\mathbf{F}(\mathbf{F}^H\mathbf{F})^{-1}\mathbf{F}^H\Delta\mathbf{k}=\mathbf{P}_{\mathbf{F}}^{\perp}\Delta\mathbf{k}, \tag{26}$$

where $\mathbf{P}_{\mathbf{F}}^{\perp}$ is

$$\mathbf{P}_{\mathbf{F}}^{\perp}\triangleq\mathbf{I}-\mathbf{P}_{\mathbf{F}}=\mathbf{I}-\mathbf{F}(\mathbf{F}^H\mathbf{F})^{-1}\mathbf{F}^H. \tag{27}$$

Substituting (25) into $be^{j\mathbf{k}r}$ gives

$$\begin{aligned}be^{j\mathbf{k}r}&=be^{(j\overline{\mathbf{k}}r+j\epsilon_1\overline{\mathbf{k}}r+j\epsilon_2\mathbf{1}r+j\Delta\mathbf{k}^{\perp}r)}\\&=be^{j\epsilon_2 r}e^{(j\overline{\mathbf{k}}(1+\epsilon_1)r+j\Delta\mathbf{k}^{\perp}r)}=\tilde{b}e^{(j\overline{\mathbf{k}}\tilde{r}+j\Delta\mathbf{k}^{\perp}r)},\end{aligned} \tag{28}$$

where $\tilde{b}\triangleq be^{j\epsilon_2 r}$ and $\tilde{r}\triangleq(1+\epsilon_1)r$. Equation (28) demonstrates that the horizontal wave-number errors collinear with the vector $\mathbf{1}$ result in an error in the estimate of the complex amplitude b, whereas horizontal wave-number errors collinear with the vector $\mathbf{k}$ result in a bias in the source range estimate. Horizontal wave-number errors in other directions (i.e., in the space defined by $\Delta\mathbf{k}^{\perp}$) result in distortions of the ambiguity/likelihood function.

In order to exclude only unpredictable modes that result in distortions of the ambiguity/likelihood function, take the projection of the estimated covariance matrix of the horizontal wave-number Γ into the null space of the matrix $\mathbf{F}$, i.e., let

$$\begin{aligned}\mathbf{\Gamma}_{\mathrm{proj}}\triangleq\mathrm{cov}(\Delta\mathbf{k}_l^{\perp})&\approx\frac{1}{L}\sum_{l=1}^{L}\Delta\mathbf{k}_l^{\perp}\Delta\mathbf{k}_l^{\perp H}\\&=\mathbf{P}_{\mathbf{F}}^{\perp}\left(\frac{1}{L}\sum_{l=1}^{L}(\mathbf{k}_l-\overline{\mathbf{k}})(\mathbf{k}_l-\overline{\mathbf{k}})^H\right)\mathbf{P}_{\mathbf{F}}^{\perp}.\end{aligned} \tag{29}$$

The resulting covariance matrix, $\mathbf{\Gamma}_{\mathrm{proj}}$ is associated with horizontal wave-number errors in the column space of $\Delta\mathbf{k}^{\perp}$.

Let S_{M_0} be the space of all M_0-sized subsets of the M modes, and let $\Gamma_{\mathrm{proj}}(\Omega_c)$ denote the covariance matrix Γ_{proj} calculated for a particular subset selection, Ω_c, of M_0 modes. Then the most predictable selection of M_0 modes, Ω_p, is determined as that subset which minimizes the trace of $\Gamma_{\mathrm{proj}}(\Omega_c)$:

$$\Omega_p=\arg\min_{\Omega_c\in S_{M_0}}\mathrm{trace}(\mathbf{\Gamma}_{\mathrm{proj}}(\Omega_c)). \tag{30}$$

The eigenvalues of $\mathbf{\Gamma}_{\text{proj}}$ express the variability of $\mathbf{\Delta k}^{\perp}$ in an orthogonal space. Therefore, predictability of a subset of modes depends on the sum of eigenvalues of $\mathbf{\Gamma}_{\text{proj}}$. The subset of modes, Ω_c, which minimizes the trace (sum of eigenvalues) of $\mathbf{\Gamma}_{\text{proj}}(\Omega_c)$ is selected as the most predictable subset of modes. Note that the minimization in (30) is performed over the selection of M_0 sized subsets of M modes. Thus the total number of subsets to evaluate in order to find the most predictable modes is $\binom{M}{M_0} = M!/M_0!(M-M_0)!$.

Although the above procedure for identifying the most predictable modes assumes that M_0 is known, in practice, a method for determining M_0 is required. Choice of M_0 is analogous to the order determination problem in parametric spectral estimation where information theoretic criteria are used for finding the best robustness versus resolution trade-off. As discussed in the previous section, reduction of M_0 improves robustness to mismatch but causes a degradation in the local accuracy of the robust ML method. Although useful for illustrating this trade-off, the CRLB is not a tight bound on the large errors which are common in matched-field problems. A more appropriate bound for predicting large error performance would be the Barankin bound,[20] whose application to matched-field processing remains a subject for future work. In the following section, results are reported using values of M_0 which were found to give the best robust ML localization performance.

V. SIMULATION RESULTS

In this section, the performance of the proposed estimator is evaluated and compared to the conventional ML, Bartlett and MMP. In these simulations, the environmental configuration shown in Fig. 1 was used.[15] The point source is narrow band at frequency 250 Hz. The field is sampled by a vertical line array containing 20 hydrophones spaced 5 m apart ranging from 5 to 100 m in depth. The ambiguity surfaces are computed for source ranges between 5 and 10 km away from the array and at source depths 0 to 100 m. The additive noise at the sensors is zero mean, Gaussian and i.i.d.

The predictable modes were identified according to the variance of the horizontal wave numbers, κ_m [i.e., Eq. (30)]. One hundred independent realizations of the environment were used in order to estimate the variance of κ_m. The number of predictable modes used was $M_0=7$. The estimated horizontal wave-number standard deviation of the modes is shown in the upper graph of Fig. 2. The lower graph depicts the wave-number standard deviation of the selected modes (numbers 3 through 9) after rejection of the components collinear with the vectors $\mathbf{k}$ and $\mathbf{1}$.

As previously noted, the proposed estimator rejects the unpredictable modes by projecting the data into their orthogonal complement space. This operation is performed by multiplication of the received data by the projection matrix $\mathbf{P}_{\mathbf{A}}^{\perp}$. This is actually a mode filtering operation. Nulling the unpredictable modes is not perfect, since in construction of $\mathbf{P}_{\mathbf{A}}^{\perp}$ the supposed eigenfunctions are different from the true eigenfunctions because of local environmental mismatch. Therefore, it is interesting to evaluate the effect of errors in the local eigenfunctions in the modal filtering response. Figure 3 shows the modal filter response for an environmental realization that was randomly selected assuming uniform prior distributions with limits given in Table I. In the selected realization, 13 modes are predicted in total. In Fig. 3 suppression of the rejected modes (i.e., mode numbers: 1, 2, 10, 11, 12, 13) can be observed. This ''sidelobe leakage'' is negligible at low SNR's but it becomes dominant at SNR's above 20 dB.

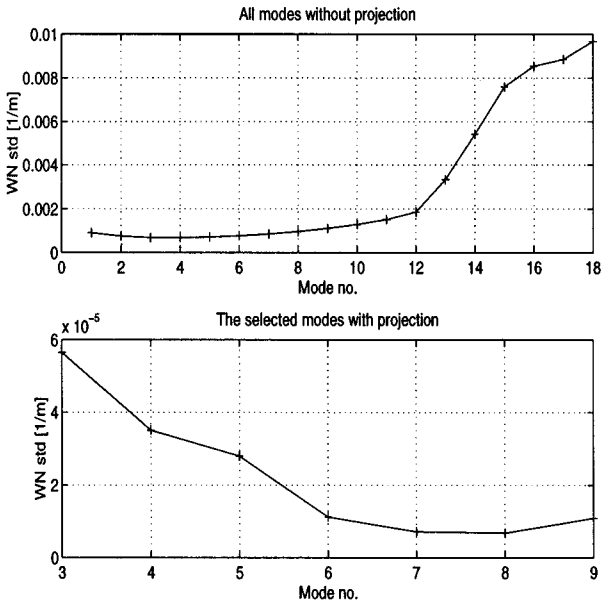


FIG. 2. The horizontal wave-number standard deviations of all modes and the selected modes with projection.

Define the probability of correct localization (P_{CL}) as the probability that the absolute source location estimation error is less than 400 m in range and 2 m in depth. To estimate the P_{CL}, 100 independent realizations of the random environmental parameters described by Fig. 1 and Table

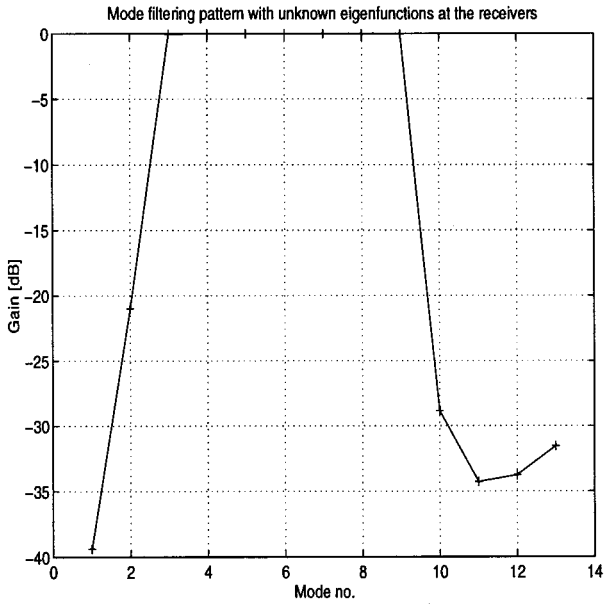


FIG. 3. A typical modal filter response for a realization of a local modal eigenfunctions.

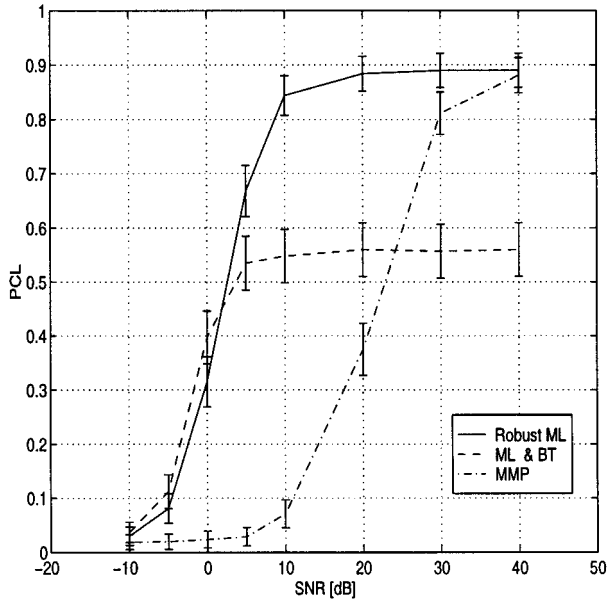


FIG. 4. The probability of correct localization for the robust ML estimator (solid line), the regular ML (dashed line), and MMP (dashed-dotted line).

I were generated for each SNR. Figure 4 compares the P_{CL} of the proposed estimator to the MMP, Bartlett (BT) and conventional ML in which the environmental parameters are not estimated. The bars in the figure denote the standard deviation of the P_{CL} estimation errors. The method used to calculate the standard deviation of the P_{CL} appears in the Appendix. The conventional ML and Bartlett estimators provide P_{CL} that is less than 0.56 for any SNR. The MMP achieves higher values of P_{CL} at high SNR's, but it fails at low SNR's. The poor performance of the MMP can be explained by the fact that it requires estimation of all the modes even though this is not necessary for source location estimation. The robust ML proposed here achieves high values of P_{CL} at high SNR's, and it performs as well as the ML and Bartlett at low SNR's. One of the reasons that the P_{CL} of the robust ML does not increase beyond a threshold SNR, is the effect of errors in the eigenfunctions at the receivers. As discussed above, the errors in the eigenfunctions become dominant at SNR's greater than 20 dB. Therefore, by increasing the SNR beyond approximately 20 dB, the performance is not improved.

In Fig. 5 the likelihood/ambiguity surfaces of the ML, MMP, and the robust ML estimator are shown for a case where the MMP and ML estimators fail at an SNR of 10 dB. Note that at this SNR, the ML and the MMP achieve P_{CL}'s of 0.56 and 0.07, respectively, whereas the proposed estimator provides a P_{CL} of 0.84.

VI. EXPERIMENTAL RESULTS WITH MEDITERRANEAN DATA

The algorithm proposed in this paper was also evaluated using experimental data collected in the Mediterranean Sea by the NATO SACLANT Center. A detailed description of the data set may be found in Ref. 5. The data are also available on the World-Wide-Web at http://spib.rice.edu. The data were collected on 26 October 1993 in a shallow-water

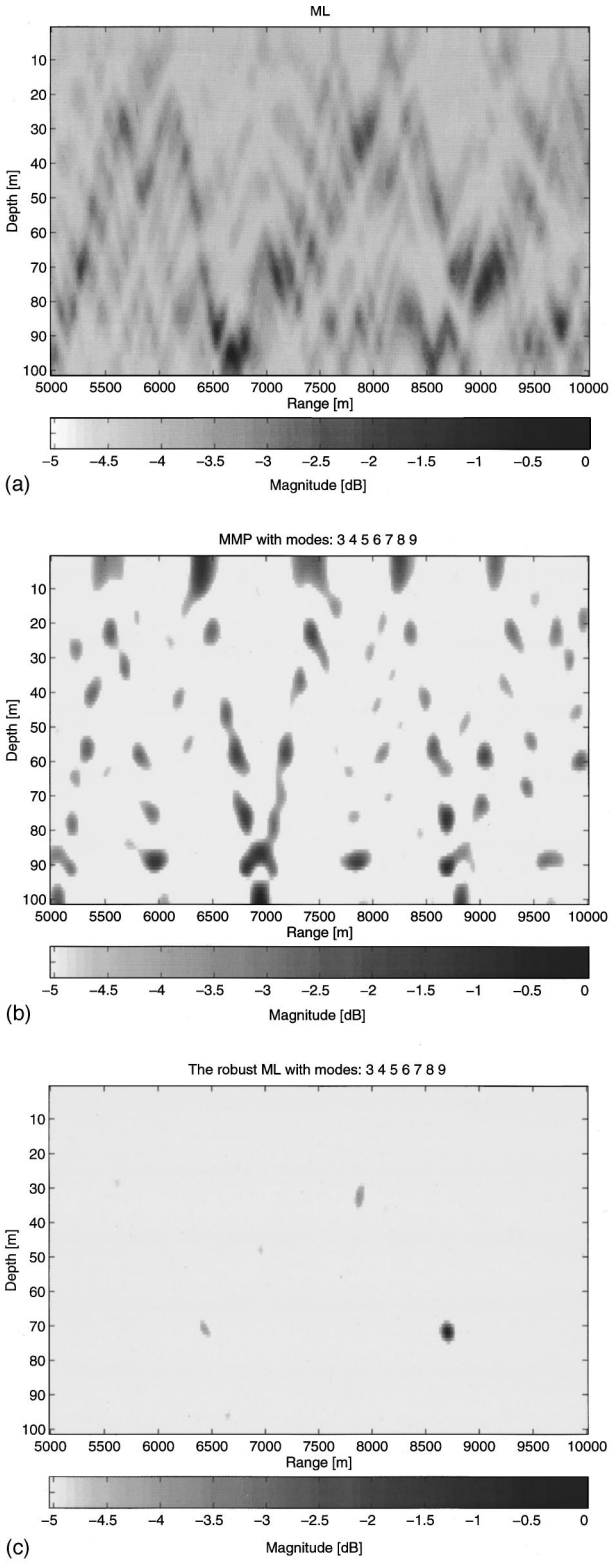


FIG. 5. The localization surfaces of the ML estimator, the MMP and the robust ML estimator for the simulations of the NRL workshop scenario. True source location: $(r,z)_0=(8780, 73)$, ML estimate: $(r,z)_{\mathrm{ML}}=(6660, 95)$, MMP estimate: $(r,z)_{\mathrm{MMP}}=(6940, 100)$, the proposed estimator: $(r,z)_{\mathrm{MLrobust}}=(8700, 72)$.

area north of the island of Elba off the Italian west coast. The environmental model consisted of an ocean layer overlying a sediment layer and a bottom layer. All layers were assumed to be range independent. The water sound-speed profile was

TABLE II. Environmental parameters for the North Elba experiment site.

Environmental parameter	Lower bound	Upper bound	Units
Bathymetry	125	130	m
Sediment thickness	0.0	6.0	m
Sediment upper sound speed	1450	1550	m/s
Sediment lower sound speed	1500	1600	m/s
Sediment attenuation	0.0	0.4	dB/λ
Sediment density	1.2	2.2	g/cm^3
Bottom sound speed	1550	1650	m/s
Bottom attenuation	0.0	0.4	dB/λ
Bottom density	1.2	2.2	g/cm^3

measured near the vertical array and assumed to be known. The environmental configuration is similar to the one used in the simulations (see Fig. 1) except that only the geoacoustic parameters were unknown. The range of uncertainties in the environmental parameters is given in Table II.

The vertical array had a total aperture of 94 m, containing 48 hydrophones with 2-m spacing, where the top hydrophone was at a depth of 18.7 m. The wideband pseudo-random source was located at a range of between 5.4 and 5.8 km from the array and at a depth of between 78 and 82 m. The data was acquired at a sample rate of 1 kHz over 65.536 s. The data at each hydrophone was transformed into the frequency domain using a Fourier transform over the entire data length. Since the narrow-band problem is of interest here, only the frequency bin near 169.9 Hz was used to obtain a narrow-band snapshot of the sensor outputs.

Identification of the predictable modes was performed using a Monte-Carlo estimate of $\Gamma_{\text{proj}}(\Omega_c)$ and the procedure described in (30). One hundred independent realizations of the field were used to estimate the horizontal wave-number covariance matrix. The number of predictable modes used was $M_0=3$. The modes 2, 4, and 5 were found to be predictable in this sense whereas the other mode amplitudes were modeled as nuisance parameters. In order to statistically evaluate the performance of the proposed estimator with the collected data, 100 independent realizations of environmental parameters were selected within a range given in Table II to represent the environmental parameters assumed by the processors. The estimated P_{CL} of the robust ML estimator was 0.79 whereas the Bartlett and the ML achieved a P_{CL} of 0.6. Figure 6 depicts the localization surfaces of the proposed estimator, the MMP and the conventional ML estimators for one of the cases in which the MMP and ML estimators fail. The proposed estimator achieves results similar to those of the genetic algorithm[5] in which the source location and the environmental parameters are estimated simultaneously: $(r,z)_{\text{MLrobust}}=(5400, 77)$ where $(r,z)_{\text{GA}}=(5437, 74.6)$ for the source location at $(r,z)_0=(5600\pm200, 80\pm2)$. Note that the robust ML technique does not require the computationally intensive global optimization performed by genetic algorithms.

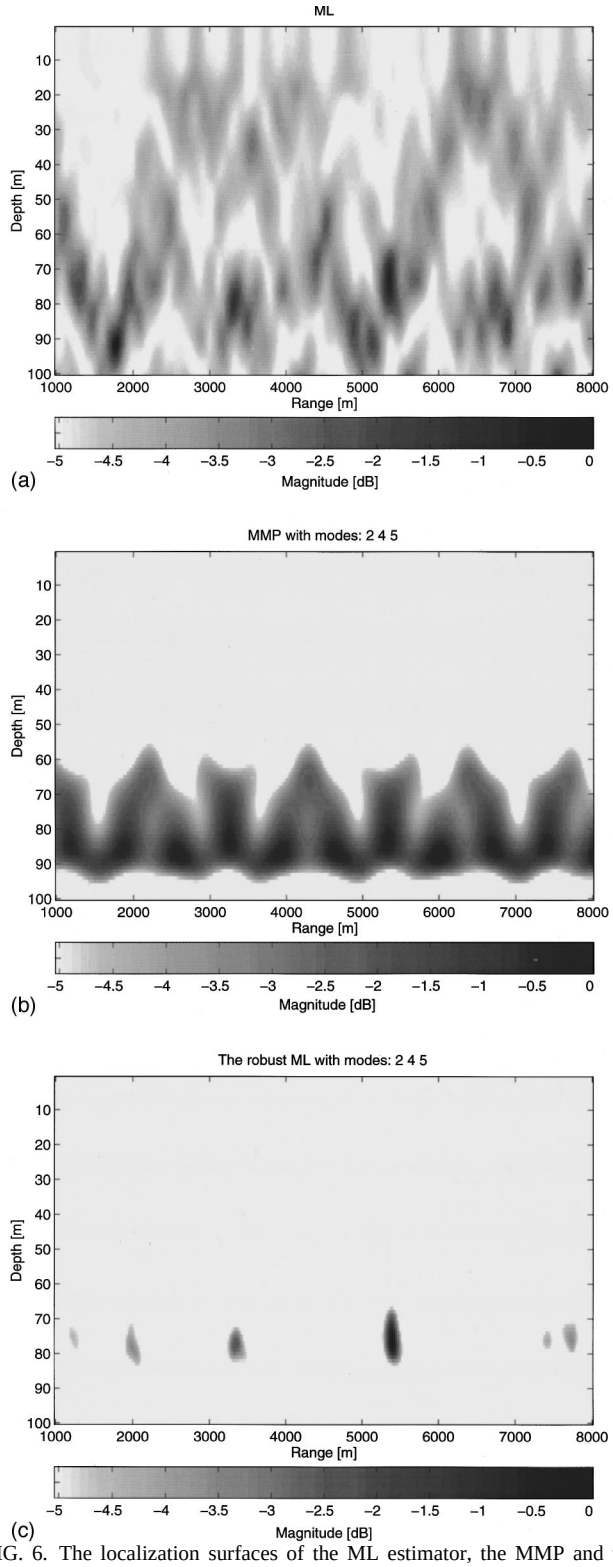


FIG. 6. The localization surfaces of the ML estimator, the MMP and the robust ML estimator for the SACLANT experimental data. True source location: $(r,z)_0=(5600\pm200, 80\pm2)$, ML estimate: $(r,z)_{\text{ML}}=(1775, 92)$, MMP estimate: $(r,z)_{\text{MMP}}=(6000, 88)$, the robust ML estimate: $(r,z)_{\text{MLrobust}}=(5400, 77)$.

VII. CONCLUSIONS

In this paper, a robust ML estimator for source localization in a shallow-water environment has been developed. The method exploits predictable modes for source localization while rejecting unpredictable modes. It has been compared to other methods using simulations and experimental data and has been demonstrated to achieve higher values of P_{CL}, particularly at mid-range SNR values.

In this work, the second-order statistics of the horizontal wave numbers were used to evaluate the predictability/

coherency level of the modes. Using the CRLB it was shown that in the case of a dense array which spans the entire water column, reducing the number of modes used in the replica degrades the local accuracy in the source location estimate, even though it improves robustness to larger errors. This suggests that there exists an optimal number of modes which should be used in the replica field. Further work is required in order to more precisely quantify this important trade-off.

ACKNOWLEDGMENTS

This work is partially supported under ONR Grant No. N66001-95-C-6032. The authors would also like to thank Dr. D. F. Gingras of ONR for making the SACLANT data available. The authors would also like to express their thanks to the anonymous reviewers of this paper for their very constructive suggestions.

APPENDIX

The P_{CL} is estimated according to N independent results of the algorithm. We define N random variables, $\{a_i\}_{i=1}^{N}$ as

$$a_i \triangleq \begin{cases} 1, & \text{for correct localization in } i\text{th experiment,} \\ 0, & \text{for false localization in } i\text{th experiment.} \end{cases} \tag{A1}$$

The P_{CL} is estimated by

$$\hat{P}_{\mathrm{CL}} = \frac{1}{N} \sum_{i=1}^{N} a_i . \tag{A2}$$

Then the variance of this estimator is

$$\mathrm{var}(\hat{P}_{\mathrm{CL}}) = E(\hat{P}_{\mathrm{CL}} - P_{\mathrm{CL}})^2 = E\left(\frac{1}{N} \sum_{i=1}^{N} a_i - P_{\mathrm{CL}} \right)^2 . \tag{A3}$$

Since the $\{a_i\}_{i=1}^{N}$ are independent, identically distributed (i.i.d.), then

$$\begin{aligned} \mathrm{var}(\hat{P}_{\mathrm{CL}}) &= \frac{1}{N^2} \sum_{i=1}^{N} E(a_i - P_{\mathrm{CL}})^2 \\ &= \frac{1}{N^2} \left[\sum_{i=1}^{N} E(a_i^2) - 2 \sum_{i=1}^{N} E(a_i) P_{\mathrm{CL}} + N P_{\mathrm{CL}}^2 \right]. \end{aligned} \tag{A4}$$

Note that $a_i \in \{0,1\}$, such that $a_i^2 = a_i$, and $E(a_i^2) = E(a_i) = P_{\mathrm{CL}}$. Therefore (34) can be written as

$$\begin{aligned} \mathrm{var}(\hat{P}_{\mathrm{CL}}) &= \frac{1}{N^2} \left[N P_{\mathrm{CL}} - 2 N P_{\mathrm{CL}}^2 + N P_{\mathrm{CL}}^2 \right] \\ &= \frac{1}{N} P_{\mathrm{CL}} (1 - P_{\mathrm{CL}}). \end{aligned} \tag{A5}$$

[1] D. F. Gingras, ''Methods of predicting sensitivity of matched-field processors to mismatches,'' J. Acoust. Soc. Am. **86**, 1940–1949 (1989).
[2] A. Tolstoy, ''Sensitivity of matched-field processing to sound speed profile mismatch for vertical arrays in a deep water pacific environment,'' J. Acoust. Soc. Am. **85**, 2394–2404 (1989).
[3] A. M. Richardson and L. W. Notle, ''*A posteriori* probability source localization in an uncertain sound speed, deep ocean environment,'' J. Acoust. Soc. Am. **89**, 2280–2284 (1991).
[4] J. A. Shorey, L. W. Notle, and J. L. Krolik, ''Computationally efficient monte carlo estimation algorithms for matched field processing in uncertain ocean environments,'' J. Comput. Acoust. **2**(3), 285–314 (1994).
[5] D. F. Gingras and P. Gerstoft, ''Inversion for geometric and geoacoustic parameters in shallow water: Experimental results,'' J. Acoust. Soc. Am. **97**, 3589–3598 (1995).
[6] C. E. Lindsay and N. R. Chapman, ''Matched-field inversion for geoacoustic model parameters using simulated annealing,'' IEEE J. Ocean. Eng. **18** 224–231 (1993).
[7] C. W. Bogart and T. C. Yang, ''Comparative performances of a matched mode and matched field localization in a range-dependent environment,'' J. Acoust. Soc. Am. **92**, 2051–2068 (1992).
[8] J. L. Krolik, ''Matched-field minimum variance beamforming in a random ocean channel,'' J. Acoust. Soc. Am. **92**, 1408–1419 (1992).
[9] T. C. Yang, ''A method for range and depth estimation by model decomposition,'' J. Acoust. Soc. Am. **82**, 1736–1745 (1987).
[10] M. J. Hinich and E. J. Sullivan, ''Maximum-likelihood passive localization using mode filtering,'' J. Acoust. Soc. Am. **85**, 214–219 (1979).
[11] E. C. Shang, C. S. Clay, and Y. Y. Wang, ''Passive harmonic source ranging in waveguides by using mode filter,'' J. Acoust. Soc. Am. **78**, 172–175 (1985).
[12] G. R. Wilson, R. A. Koch, and P. J. Vidmar, ''Matched mode localization,'' J. Acoust. Soc. Am. **84**, 310–320 (1988).
[13] M. J. Hinich, ''Maximum-likelihood signal processing for a vertical array,'' J. Acoust. Soc. Am. **54**, 499–503 (1973).
[14] M. J. Hinich, ''Maximum-likelihood estimation of the position of a radiating source in a waveguide,'' J. Acoust. Soc. Am. **66**, 480–483 (1979).
[15] M. B. Porter and A. Tolstoy, ''The matched field processing benchmark problems,'' J. Comput. Acoust. **2**(3), 161–185 (1994).
[16] I. Tolstoy and C. S. Clay, *Ocean Acoustics* (McGraw-Hill, New York, 1966).
[17] L. B. Felsen and A. H. Kamel, ''Hybrid ray-mode formulation of parallel plane waveguide Green's functions,'' IEEE Trans. Antennas Propag. **AP-29(3)**, 637–649 (1981).
[18] K. Ogata, *Modern Control Engineering* (Prentice-Hall, Englewood Cliffs, NJ, 1990), 2nd ed., pp. 931–932.
[19] M. B. Porter and E. L. Reiss, ''A numerical method for ocean acoustic normal modes,'' J. Acoust. Soc. Am. **76**, 244–252 (1984).
[20] R. J. McAulay and E. M. Hofstetter, ''Barankin bounds on parameter estimation,'' IEEE Trans. Inf. Theory **17**, 669–676 (1971).

The schlieren image of two-dimensional ultrasonic fields and cavity resonances

Paul A. Chinnery, Victor F. Humphrey, and Carolyn Beckett
School of Physics, University of Bath, Bath BA2 7AY, United Kingdom

(Received 9 May 1996; accepted for publication 15 August 1996)

The schlieren technique can be used to visualize the two-dimensional ultrasonic standing wave fields associated with circular cylindrical shells under resonance conditions. However, the interpretation of the schlieren image is not simple due to the complex relationship between the acoustic and optical fields. A model for predicting the optical image in an ideal schlieren system is presented and used to investigate the influence of the acoustic pressure and optical spatial filtering on the resultant image, demonstrating the conditions under which the image is a meaningful representation of the acoustic wave field. A low-frequency (>100 kHz), wide-aperture, laboratory schlieren system is used to image the fluid column resonances of a circular cylindrical shell. Experimental results agree well with the predictions, validating the theory. Although the schlieren image is two-dimensional, limiting investigations to targets having translational symmetry, the technique is noninvasive and can potentially provide greater insight into the acoustic resonance behavior of more complex scattering geometries. [S0001-4966(97)00601-2]

PACS numbers: 43.35.Sx, 43.20.Ks, 43.20.Ye [HEB]

INTRODUCTION

Acoustic resonances are known to play an important part in the scattering of waves from discrete objects immersed in water and a number of techniques have been developed to identify the resonance modes of scattering objects. For liquid-filled cylindrical shells Maze *et al.*[1] have used the method of isolation and identification of resonances (MIIR) to experimentally obtain resonance spectra from the scattered field. Several different types of resonance were identified, the majority of resonances being attributed to the fluid column within the shell.

In contrast to the scattered field, the interior field in the fluid cavity has received little attention in the literature. One technique that has been used to visualize and study acoustic fields in fluids is the schlieren technique.[2] This approach can also be used to study the resonance and scattering behavior of cylindrical objects in detail, and importantly, in a noninvasive way. The acoustic fields both inside and outside insonified cylindrical shells at resonance exhibit distinctive standing wave patterns. Several classes of resonance (fluid column modes, shell resonances, and Stoneley wave resonances) and individual modes within each class can be identified from the symmetry of the resulting wave fields. The acoustic fields in the interior and exterior fluids can easily be predicted theoretically, but are not easily measured experimentally, especially within the fluid column. For targets having translational symmetry, such as straight cylinders, the schlieren technique provides a noninvasive means of experimentally imaging resonances, and, therefore, of identifying individual modes, and thus obtaining the resonance spectrum of the scattering object.

The schlieren technique is sensitive to any phenomena that produce refractive index changes in the medium through which the light beam passes. It has mainly been used in the visualization of high speed air flow and in underwater ultrasonics. Toepler[3] was among the first to use the schlieren technique to visualize the sound waves emitted by an electric spark (in air). The earliest known application to the visualization of ultrasound is described in a publication by Tawil[4] who obtained images of ultrasonic waves in gases. More recent applications of schlieren in underwater acoustics have involved the visualization of acoustic scattering phenomena; see, for example, Neubauer[5] and Newman.[6] In particular, the pioneering work of Neubauer and colleagues has provided valuable insight into the mechanisms involved in scattering by planar and curved surfaces, including cylindrical structures. A related technique involving stress birefringence in optically transparent solids[7] has been used to study the elastic resonances of solid cylinders.[8–10]

Much of the work involving schlieren has been performed at frequencies well above one megahertz. Although visualization at these frequencies can be achieved with moderate ease, using a variety of experimental techniques based on schlieren, individual wave fronts have not normally been resolved. In addition, for the low ka ranges of interest in the marine environment, extremely small scale models of scattering targets would be required if studies were to be undertaken in the megahertz region. With the aim of studying scattering by discrete objects of reasonable size at low ka, a low-frequency schlieren system has been developed at the University of Bath.[11–13] This system permits the visualization of pulsed and continuous acoustic fields with carrier frequencies down to about 100 kHz, and has been used to study the scattering and resonance behavior of circular cylindrical objects.[12,13] The location and identification of resonance modes is particularly simple using this technique. The schlieren system developed at Bath has also been used to investigate the resonance behavior of elliptical[14,15] and stadium shaped[16] cavities. A low-frequency experimental ar-

rangement using light diffraction techniques has also been reported by Wolf *et al.*[17]

The present study concerns itself with the interpretation of cavity resonance images obtained with the schlieren system. Of particular concern are the effects that the spatial filtering arrangement and acoustic pressure levels have on the optical image. Under certain conditions the optical image produced by an ideal schlieren system is known to approximate the square of the acoustic pressure field, and comparisons with theoretical predictions of pressure squared show reasonable agreement.[18] Unfortunately, the acoustic pressure in the fluid column is unknown and of variable intensity, and thus the validity of this comparison over the entire field is uncertain. The present study uses a Fourier optics approach to investigate this relationship between the acoustic field and optical image. The Fourier optics technique has previously been used by Bucaro *et al.* to study schlieren imaging of sinusoidal acoustic signals[19] and Gaussian-like pulses.[20] For this one-dimensional case the effects of spatial filtering arrangements[19] and pressure amplitude[20] were investigated in an attempt to aid interpretation and extract quantitative information about the acoustic field from the optical image. Fourier techniques have also been used by Bucaro and Dardy[21] to investigate the sensitivity of an ideal schlieren system for visualizing low-frequency ($\ll$1-MHz) waves.

This paper presents an exact theoretical model for predicting the optical image produced by an ideal schlieren system imaging a general two-dimensional acoustic field under conditions of Raman–Nath diffraction. This model is not restricted to central order filtering or low acoustic pressures. Example results are presented that illustrate the application of schlieren to the study of resonances of water-filled cylindrical cavities of circular cross section. Computed results for the fluid column resonance modes of a circular cylindrical shell are presented and compared with images obtained from the experimental system. The theoretical results are also compared with the pressure-squared approximation to test its validity.

I. SCHLIEREN

A. Principle

The principle of the schlieren technique is that light is refracted as it passes through an acoustic wave field in a fluid, since the refractive index is a function of fluid density, and hence pressure. This interaction of an optical wave with an acoustic wave has been the subject of much study since the effect was first explained by Raman and Nath in 1935.[22] More recent discussions of the interaction include those by Klein and Cook[23] and Van Den Abeele and Leroy.[24] It is known that if the intensity of the acoustic field is sufficiently low, the diffraction effect is small and the light suffers only phase variations. This is known as the Raman–Nath diffraction. In this case the acoustic field can be considered to be a ''phase grating'' which produces an interference pattern containing information about the acoustic field. This information can be processed to produce an ''image'' corresponding to the acoustic field. For a given wavelength of light, piezo-optic coefficient, and interaction length, the relationship that this image has with the acoustic field depends on the optical processing applied in the diffraction plane and the amplitude of the acoustic pressure field through which the light passes.

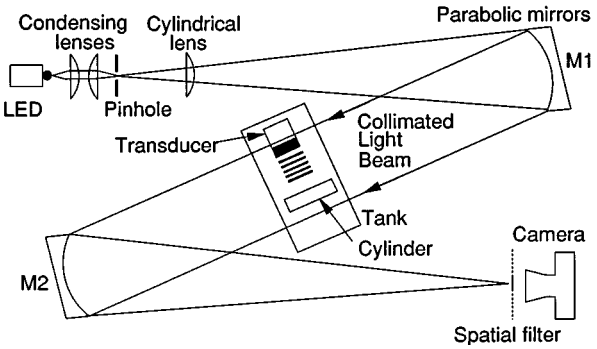


FIG. 1. The schlieren visualization system.

B. Experimental system

The experimental arrangement (Fig. 1) is based upon a standard Z configuration.[12,17] A pair of condensing lenses focus the light from a high-power light emitting diode onto a pinhole, which acts as an effective optical source. The light transmitted by the pinhole is collimated by the first parabolic mirror and passed through the water-filled tank containing the acoustic field and test object. A cylindrical lens is placed between the pinhole and the mirror to help reduce the aberration caused by the use of parabolic mirrors off-axis.

The source of acoustic waves in the water is a transducer positioned with its axis perpendicular to the light beam. The light emerging from the tank is focused by the second parabolic mirror and forms a diffraction pattern in the focal plane of that mirror. In the absence of an acoustic field an image of the pinhole is observed in the focal plane. If an acoustic field is present, the diffraction pattern has side orders which contain the light that has been deflected out of the main beam.

The system can be used in two different modes: continuous and pulsed. In the continuous mode of operation both the acoustic transducer and the optical source are driven continuously. In the pulsed mode the light source is pulsed in synchronization with the acoustic source, allowing pulses to be visualized; by varying the delay between the acoustic and optical pulses an acoustic pulse can be observed as it propagates in apparent slow motion. The pulsed mode of operation is particularly informative when studying scattering by hollow cylindrical structures.[12]

If Raman–Nath diffraction is assumed, then a uniformly bright image would be obtained were all the light in the focal plane used to form an image. This is because the light emerging from the tank has only phase variations, and is not altered in amplitude. If, however, a part of the diffraction pattern is removed by a spatial filter, and the remaining light allowed to recombine, a meaningful ''image'' can be obtained on film or video. The type of image produced, and its relationship to the acoustic field, is critically dependent on the filtering applied in the focal plane of the second parabolic mirror. Two commonly used filtering methods are central order filtering and knife-edge filtering, both of which have advantages and disadvantages. Central order filtering in-

volves the use of a circular stop to remove only the central order in the diffraction pattern; knife-edge filtering involves the removal of half of every diffraction order by a straight edge stop. Although commonly used for imaging one-dimensional fields, knife-edge filtering is less useful for direct visualization of two-dimensional acoustic fields, as demonstrated later.

At low frequencies the resolution of the diffraction orders can become problematic. The separation of the orders (s) in the focal (transform) plane is given by

$$s = \frac{kh}{\kappa}, \tag{1}$$

where h is the focal length of the second mirror, κ is the optical wave number, and k is the acoustic wave number. Thus, for light of a fixed frequency and mirrors of fixed focal length, the order separation decreases with the acoustic frequency, and orders begin to overlap due to their finite size.

While the use of smaller pinholes results in narrower diffraction orders, it also reduces the amount of light passing through the system. For this reason an array of pinholes is sometimes used in the current system to increase the light level and produce narrower orders. The use of an incoherent light source prevents unwanted interference effects. An array of eight 100-μm pinholes and an array of eighteen 50-μm pinholes have both been used. When multiple pinhole arrays are being used, central order spatial filtering is achieved by using a matching array of spatial filters. Aberrations in the system result in distortion of the diffraction orders, therefore close-up photographs of the pinhole image without an acoustic field in the tank are used as spatial filters. The arrays are given a random basis, so as to minimize the inevitable overlapping between the side orders of one pinhole pattern and those of another.

Although the use of a single pinhole source and a knife-edge filter is the easiest way of obtaining an image of the acoustic field, it does—as will be demonstrated—produce an image bearing a complicated relation to the (two-dimensional) acoustic field. The use of central order filtering produces images that are easier to interpret (provided the pressure levels are not too high), and this method is generally used with the present experimental system when resonances are being imaged. The use of multiple pinhole arrays to increase the light level in the system also precludes the use of knife-edge filtering.

C. Measurement technique

The schlieren system can be used as a tool to locate, identify, and record the resonances of cylindrical shells in the following manner.

The test cylinder is placed in the water tank, correctly aligned and insonified by the transducer; the frequency of the continuous wave drive signal is swept, either manually, or automatically by the function generator. A video camera, placed immediately behind the spatial filter, is used to form an image of the object plane and monitor the acoustic field. At frequencies corresponding to resonances of the test object, acoustic standing waves are clearly seen in the fluid column, and sometimes in the outer fluid. At other frequencies, where little energy penetrates the shell, only the scattered field is seen.

To record the resonance patterns, the video camera is replaced by a still camera, the frequency locked, the drive level adjusted until an acceptable image obtained, and several timed exposures made.

Care must be taken when interpreting the observed patterns and identifying the resonance modes as other resonances can be excited by overtones of the drive frequency, or possibly by nonlinearly generated harmonics in the acoustic field. Care must also be taken to correctly identify resonance modes of different type, but similar appearance, such as certain fluid column modes and shell resonances. Also, resonance modes which are close in frequency may overlap, making identification of either mode difficult. This occurs particularly at high frequencies where the density of resonances is greatest.[25]

The video camera output can be fed into a personal computer incorporating a line-grabbing digitization board and software. This facility allows the optical intensity at a point within the image to be monitored, and intensity profiles of the resonances to be obtained. The use of this facility, combined with the theoretical model described in Sec. II, offers the potential of using the system to obtain more quantitative information from the schlieren system than has hitherto been possible.

D. Interpretation of image

Interpretation of the image of a one-dimensional progressive wave is quite simple and the optical image can easily be predicted. In this case, studies of the acousto-optic interaction and spatial filtering arrangements show that as the pressure amplitude changes, the amount of light falling into the different orders in the diffraction plane varies, causing the image to change.[11,19] For certain values of the pressure, the central optical order completely disappears as all the light is diffracted into the higher orders. As this is tantamount to changing the spatial filtering arrangement, the optical distribution changes accordingly. For central order filtering the optical distribution varies as the square of the acoustic pressure if the pressure is low, but changes significantly as the pressure is increased. Knife-edge filtering on the other hand produces an optical image whose intensity varies linearly with pressure over a broader range of pressures.

For two-dimensional standing waves, such as in the fluid cavities of cylindrical shells, the situation is not so simple. At certain places in the acoustic standing wave, the pressure amplitude is always zero, and we would expect that light will pass through the field undeflected in these regions, so that the central order would never completely disappear. Additionally, in the standing wave fields of interest in the current study, the acoustic fields are two-dimensional and the pressure amplitude varies throughout the cross section of the cavity, potentially creating a very complex dependence of the optical image upon the acoustic field. Knife-edge filtering produces an image that is harder to interpret as only the wave fronts parallel to the knife are imaged. For these reasons a more general theoretical investigation into the image forma-

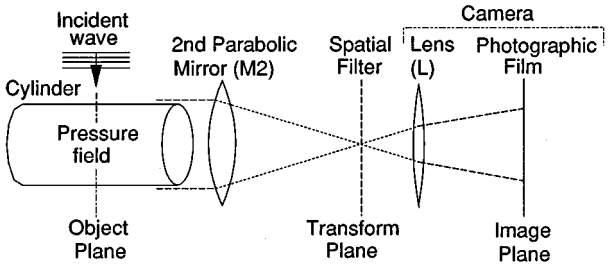


FIG. 2. Location of the field planes in the schlieren system.

tion is required. This theory, applicable to two-dimensional fields and standing waves, is presented in Sec. II.

II. THEORY

The approach used to calculate the optical distribution in the schlieren image is based on a Fourier optics technique, similar to that used by Bucaro *et al.*[19–21] for one-dimensional fields. The field planes in the schlieren system are shown in Fig. 2.

For a continuous wave excitation the acoustic pressure field within the fluid can be written

$$P(\mathbf{r},t)=e^{-i\omega t}p(\mathbf{r})e^{i\delta(\mathbf{r})}, \tag{2}$$

where $p(\mathbf{r})$ is the spatial dependence of the pressure field, and $\delta(\mathbf{r})$ is its phase. Under Raman–Nath conditions the acoustic pressure field produces phase variations (σ) in the light wave that passes through it. These phase variations are proportional to the real part of the pressure distribution:

$$\sigma(\mathbf{r},t)\propto \mathrm{Re}\{P\}=\tau p(\mathbf{r})(\cos\ \delta(\mathbf{r})\cos\ \omega t + \sin\ \delta(\mathbf{r})\sin\ \omega t), \tag{3}$$

where

$$\tau=\kappa\left(\frac{\partial\mu}{\partial p}\right)_s l, \tag{4}$$

and where κ is the optical wave number, $(\partial\mu/\partial p)_s$ is the adiabatic piezo-optic coefficient (which is independent of the pressure and constant for a given medium), and l is the acousto-optic interaction length (which is approximately equal to the width of the transducer face). The parameter τ is dependent on the strength of the acousto-optic interaction and is a constant for a given experimental arrangement. For a propagating field, the maximum value of $\tau p(r)$ corresponds to the commonly used Raman–Nath parameter.

The optical field emerging from the tank (in the object plane; see Fig. 2) can then be written

$$E(\mathbf{r},t)=E_0e^{i\sigma}=E_0e^{i\tau p(\mathbf{r})\cos\ \delta(\mathbf{r})\cos\ \omega t}e^{i\tau p(\mathbf{r})\sin\ \delta(\mathbf{r})\sin\ \omega t}, \tag{5}$$

or, by expanding the exponentials in series of Bessel functions,

$$E(\mathbf{r},t)=E_0\sum_m\ i^m J_m(\tau p(\mathbf{r})\cos\ \delta(\mathbf{r}))e^{im\omega t} \times\sum_n\ J_n(\tau p(\mathbf{r})\sin\ \delta(\mathbf{r}))e^{in\omega t}. \tag{6}$$

This then is the optical field in the object plane under conditions of Raman–Nath diffraction. A related approach to the diffraction of light by ultrasound has been reported by Hargrove[26] for the case of a Gaussian light beam and an arbitrary one-dimensional periodic acoustic wave, and by Mayer and Neighbors[27] for ultrasonic pulses.

After reflection in the second mirror (M2) the optical field in the transform plane (E_T) is given by the spatial Fourier transform (FT) of E,

$$E_T(\boldsymbol{\eta},t)=E_0\sum_{m,n}\ i^m e^{i(n+m)\omega t} \times\mathrm{FT}\{J_m(\tau p\ \cos\ \delta)J_n(\tau p\ \sin\ \delta)\}, \tag{7}$$

where η is the coordinate vector in the transform plane. The diffraction pattern observed in the transform plane is given by $E_TE_T^*$, where E_T^* is the complex conjugate of E_T. Different directions in the diffraction pattern correspond to different spatial frequencies in the object plane.

A spatial filter, $\mathfrak{T}(\boldsymbol{\eta})$, placed in the transform plane removes part of the optical field, leaving the transmitted field ($E_{T'}$):

$$E_{T'}(\boldsymbol{\eta},t)=E_0\sum_{m,n}\ i^m e^{i(m+n)\omega t}\aleph_{mn}(\boldsymbol{\eta}), \tag{8}$$

where

$$\aleph_{mn}(\boldsymbol{\eta})=\mathfrak{T}(\boldsymbol{\eta})\mathrm{FT}\{J_m(\tau p\ \cos\ \delta)J_n(\tau p\ \sin\ \delta)\}. \tag{9}$$

The transmitted waves then pass through the imaging lens (camera) to produce the final optical field in the image plane. This field (E_I) is the inverse Fourier transform of $E_{T'}$

$$E_I(\mathbf{r}',t)=\mathrm{FT}^{-1}(E_{T'}) =E_0\sum_{m,n}\ i^m e^{i(m+n)\omega t}\ \mathrm{FT}^{-1}\{\aleph_{mn}(\boldsymbol{\eta})\}. \tag{10}$$

The intensity of the final image is given by

$$I(\mathbf{r}',t)=E_IE_I^*=E_0^2\sum_{m,n,p,q}\ i^m e^{i(m+n-p-q)\omega t}\ \mathrm{FT}^{-1}(\aleph_{mn}) \times[i^p\ \mathrm{FT}^{-1}(\aleph_{pq})]^*, \tag{11}$$

but the actual image observed by the eye, or recorded on film, is the time average of I, $1/T\int^T I\ dt$, in which only terms having $m+n-p-q=0$ will contribute.

This then is the final optical intensity distribution seen in the image plane for a spatial filtering arrangement described by the function $\mathfrak{T}(\boldsymbol{\eta})$, a pressure distribution described by $p(\mathbf{r})$ and $\delta(\mathbf{r})$, and the constant τ.

III. RESULTS

The theory described in Sec. II can be used to predict the optical image resulting from different acoustic fields, pressure levels, and spatial filtering arrangements used in the schlieren system. Results are illustrated by examining the fluid column resonances of a brass cylindrical shell having outer and inner radii (a) 15.85 mm and (b) 14.25 mm. The theoretical values of $p(\mathbf{r})$ and $\delta(\mathbf{r})$ were obtained using the normal mode series solution[28] for a plane wave normally

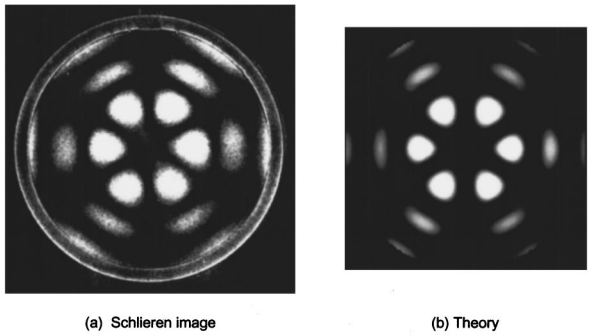


FIG. 3. The (3,3) fluid column resonance; central order filtering.

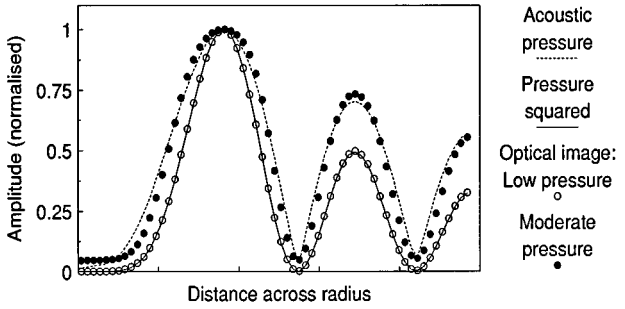


FIG. 4. Computed cross sections of the acoustic field (lines) and optical image (points) for the (3,3) fluid column mode and central order filtering.

incident on an infinite cylindrical shell. The models for the acoustic fields and optical image were both numerically implemented using Fortran and NAG FFT routines. Measurements were made with the transducer driven in continuous mode and the shell insonified normal to its axis. The samples used in the experimental measurements were 100 mm long.

A. Central order filtering: Low/moderate pressures

Figure 3(a) shows a fluid column resonance found at an experimental kb value of 11.46. This image was obtained using an array of 18 pinholes of diameter 50 μm and a matching spatial filter providing central order filtering. This is a typical schlieren observation: the drive level was adjusted so that a reasonable image was obtained; too low a value and only the regions of highest amplitude are visible (or no image at all), too high a value and other effects complicate the image (as illustrated later). The continuous wave excitation is incident from the left in the figures.

Figure 3(b) shows the theoretical prediction of the optical image calculated using Eq. (11). The computed kb value for this resonance is 11.448. The amplitude of the incident wave was not known exactly, but a typical value was inferred from analysis of the diffraction orders in a study of one-dimensional progressive waves.[11] Taking the wavelength of the light to be 660 nm, the interaction length to be approximately equal to the diameter of the transducer face (5 cm), and the piezo-optic coefficient for water to be 1.5×10^{-10} Pa^{-1},[29] the value of τ [Eq. (4)] was estimated to be 7×10^{-5} Pa^{-1}. With these estimates of the parameter τ and the pressure amplitude, the maximum value of τp in the fluid column was 4.3. This is the value at the highest pressure in the acoustic standing wave field and is proportional to the incident pressure. At resonance the pressure in the fluid column is much higher than that in the incident field. For this reason the image of the incident wave is extremely faint, or not visible at all.

In order to test the pressure-squared approximation the optical image was calculated for a very low value of $\tau p (\ll 1)$. Cross sections of the calculated optical and acoustic fields are shown in Fig. 4; the square of the pressure amplitude is shown as the solid line, the calculated optical field by the circles. The profiles are taken along the horizontal radii of the fields (orientation as in Fig. 3) and the vertical scales have been normalized for comparison. As expected, the agreement is exact; the model predicts optical images proportional to pressure squared when the acoustic pressure is very low ($\tau p \ll 1$). However, in the experimental schlieren system, typical pressures in the fluid column at resonance are thought to be higher than this.

Also shown in Fig. 4 is a comparison of the magnitude of the pressure distribution with the optical distribution calculated with a higher value of $\tau p (=3.4)$. The agreement, although not exact, is quite good, suggesting that at a certain value of τp the predicted optical image is a good approximation to the magnitude of the pressure distribution itself. The minor differences seen in the figure would not be observable in practice anyway. The values of τp present in the experimental measurements are thought to be of this order[11] and so it appears that the experimental result (where $\tau p > 1$) is a closer approximation to the magnitude of the acoustic pressure amplitude than to its square. Care must be taken in interpreting these results, however. For the standing wave field inside the fluid column, different regions have different pressure amplitudes, and therefore different values of τp (unlike the progressive plane wave). The value of τp referred to here corresponds to the maximum value of pressure in the fluid column; other regions of the field have lower values of τp, and hence the optical image is a complicated function of the acoustic wave field and cannot strictly be compared with the acoustic pressure amplitude or its square.

B. Higher pressure amplitudes

If the transducer is driven harder, the acoustic pressure in the fluid column increases and the optical image changes. An example of this effect is shown in Fig. 5.

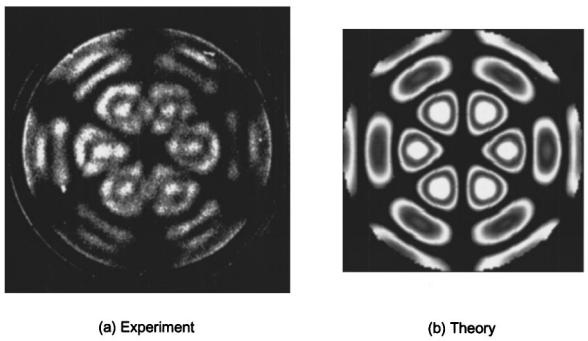


FIG. 5. The (3,3) fluid column resonance at high acoustic drive level.

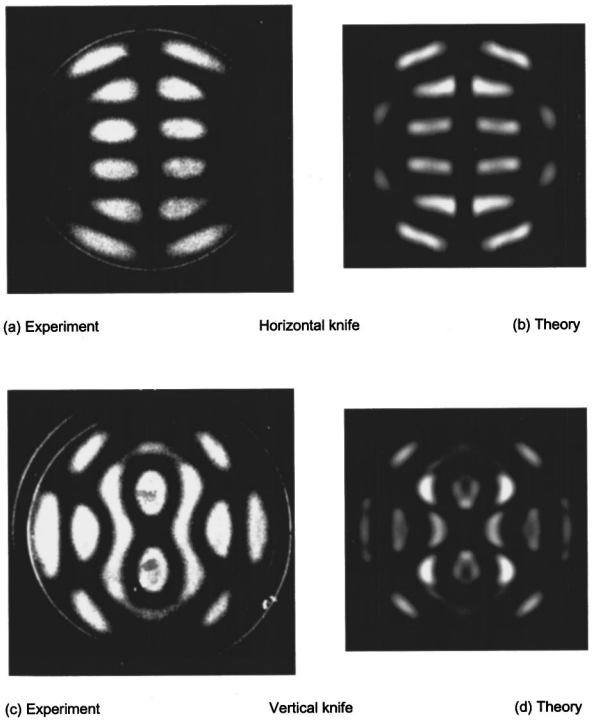


FIG. 6. The (3,3) fluid column resonance visualized using horizontal [(a) and (b)] and vertical [(c) and (d)] knife-edge filtering. The horizontal and vertical knife edges are parallel and perpendicular (respectively) to the axis of the incident beam.

Figure 5(a) shows the appearance of the experimental image for a higher acoustic drive level. Experimentally, the higher pressure levels resulted in the diffraction orders of the individual pinholes in the array overlapping with adjacent pinhole side orders, so an array of 8×100 μm pinholes, more widely spaced, had to be used. This array, however, results in reduced clarity for low-frequency fields due to the overlapping of orders in the individual pinhole images, which results in the removal of part of the first side order by the filter designed only to remove the central order.

The predicted optical distribution (calculated for a peak value of $\tau p=11.2$) displays mostly the same features as the experimental image, although the agreement is not as good as previously. The regions of highest amplitude in the standing wave are seen to broaden and split into several parts in the optical image. This complicates the identification of the mode. For this reason, when a resonance is located, the drive level is turned down as far as possible to ensure a meaningful image of the acoustic field is obtained.

C. Spatial filtering

All the images discussed in the previous sections involved central order filtering. More dramatic changes to the schlieren image of the acoustic field are evident when different spatial filtering is used in the transform plane. Figure 6 shows the (3,3) resonance again, this time imaged using knife-edge filtering. A single pinhole of 400-μm diameter was used to illuminate the acoustic field in the schlieren system. Figure 6(a) and (b) show the appearance of the resonance when the knife edge passes horizontally through the diffraction pattern (i.e., parallel to the axis of the incident wave). Figure 6(c) and (d) were produced with the knife in the vertical position (at right angles to the axis of the incident beam). Insonification is from the left. All the features seen in the experimental results can be seen in the theoretical predictions if they are examined closely. The lack of symmetry between the two cases is due to the fact that the knife edge passes along different symmetry planes in the diffraction pattern; these filtering arrangements are therefore fundamentally different.

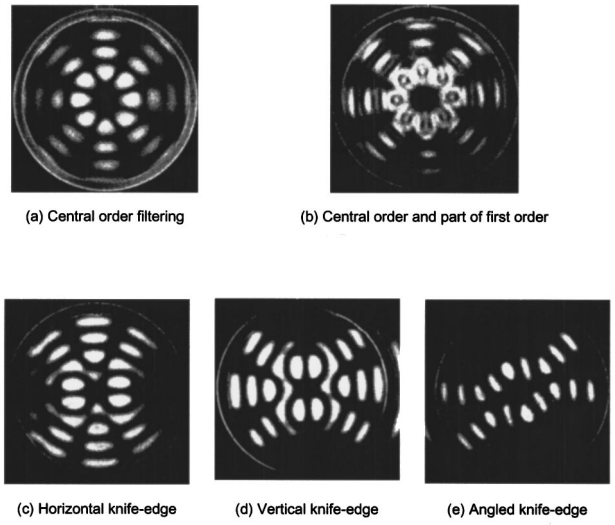


FIG. 7. The (4,4) fluid column resonance visualized using different spatial filtering arrangements; experimental observations.

Figure 7 shows the (4,4) fluid column mode visualized experimentally, again using different spatial filtering in the stop plane. The result of central order filtering is shown in (a), while (b) shows the result of using a filter that is larger than the central order, and also blocks a part of the first side orders. Knife-edge filtering was used to produce the remaining images; (c) and (d) are the result of using the knife edge horizontally and vertically. Unlike the previous case considered (Fig. 6), both of these images are the same, apart from a 90° rotation. This reflects the symmetry of the (4,4) resonance and its diffraction pattern. To obtain a different image with the knife edge it must be placed at some other angle, as is shown in (e).

IV. CONCLUSION

A theoretical model has been presented which predicts the optical image in an ideal schlieren system visualizing a general two-dimensional acoustic field. The model accurately predicts the images obtained for fluid column resonances within circular cylindrical shells for a variety of spatial filtering arrangements. It should be noted that the theory contains no allowance for the finite size pinhole used in the experimental system. Differences from the experimental results are expected for this, and other reasons, including the aberrations that exist in the optical system.

Comparison has been made between the pressure-squared distributions and the exact predictions of the optical image. Good agreement has been found when very low pressures are present in the fluid column, but departures become

obvious for more moderate pressures, and drastic for high pressures. It would appear that the experimental images are often obtained at pressures where the square law prediction no longer holds and where the exact relationship between the acoustic field and the optical image is a complex one. However, in these cases the schlieren image is still a meaningful and easily interpreted representation of the acoustic field.

Unlike central order filtering, where the acoustic pressure level affects the appearance of the image considerably, knife-edge filtering produces images that do not change appreciably as the pressure is increased to moderate levels. This is of significance when we consider that the standing wave field does not have a uniform pressure amplitude throughout the fluid column. Unfortunately, however, as shown here, the resulting images are a complex representation of the acoustic field and do not easily permit identification of the resonance modes.

The use of the experimental system for locating and identifying the resonance modes of two-dimensional cavities has also been demonstrated. The advantage of this technique is that it is noninvasive, allows immediate identification of resonance modes, and provides insight into phenomena that could not easily be measured or observed with a conventional acoustic approach using a hydrophone.

The experimental system can also be used to investigate many aspects of scattering and resonance behavior involving cylindrical objects of various geometries, and is of particular value when considering those geometries that cannot easily be modeled. Of particular interest here are the hybridization of resonance modes at high frequencies,[25] and the modes of cavities of noncircular geometry, such as the ellipse and stadium.[14,16]

The theory developed here helps in the understanding and interpretation of the images obtained using the schlieren system.

ACKNOWLEDGMENTS

The theoretical model for predicting the optical image in the schlieren system was developed by V. F. Humphrey and C. Beckett. The experimental apparatus was originally developed by S. M. Knapp. This work was supported by the Procurement Executive of the Ministry of Defence.

[1] G. Maze, J. L. Izbicki, and J. Ripoche, ''Acoustic scattering from cylindrical shells: guided waves and resonances of the liquid column,'' Ultrasonics **24**, 354–362 (1986).
[2] R. B. Barnes and C. J. Burton, ''Visual methods for studying ultrasonic phenomena,'' J. Appl. Phys. **20**, 286–294 (1949).
[3] A. Töpler, ''Optische studien nach der methode der schlieren beobachtung,'' Poggendorff' d Annal. **131**, 33–55 (1867).
[4] E. P. Tawil, ''Ultrasonic stationary waves in gases rendered visible by the method of striae,'' C. R. Acad. Sci. **191**, 92 (1930).
[5] W. G. Neubauer, ''Observation of acoustic radiation from plane and curved surfaces,'' in *Physical Acoustics*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1973), Vol. X, pp. 61–126.
[6] D. R. Newman, ''Observations of cylindrical waves reflected from a plane interface,'' J. Acoust. Soc. Am. **53**, 1174–1176 (1973).
[7] L. Bergmann, *Der Ultraschall* (Hirzel, Stuttgart, 1954).
[8] L. Bergmann, ''Eigenschwingungen von glaszylindern,'' Z. Phys. **125**, 405–417 (1949).
[9] E. K. Sittig and G. A. Coquin, ''Visualization of plane-strain vibration modes of a long cylinder capable of producing sound radiation,'' J. Acoust. Soc. Am. **48**(5), 1150–1159 (1970).
[10] H. D. Dardy, L. Flax, C. F. Gaumond, J. V. Subrahmanyam, S. Ashrafi, P. K. Raju, and H. Überall, ''Acoustically induced stresses in elastic cylinders and their visualization,'' J. Acoust. Soc. Am. **82**(4), 1378–1385 (1987).
[11] S. M. Knapp, University of Bath Internal Reports, Numbers 3 & 4 (unpublished, 1988).
[12] S. M. Knapp and V. F. Humphrey, ''Schlieren visualisation of low frequency ultrasonic fields,'' in *Ultrasonics International 89 Conference Proceedings* (Butterworth, Guildford, 1989), pp. 1089–1094.
[13] V. F. Humphrey, S. M. Knapp, and C. Beckett, ''Schlieren visualisation of the resonances of a fluid-filled cylindrical shell using a low frequency schlieren system,'' in *Physical Acoustics*, edited by O. Leroy and M. A. Breazeale (Plenum, New York, 1991), pp. 371–376.
[14] P. A. Chinnery and V. F. Humphrey, ''Resonances of deformed cylindrical shells—experimental visualisation and identification,'' in *Proceedings of the 2nd European Conference on Underwater Acoustics*, Copenhagen, edited by L. Björnö (European Commission, 1994), pp. 159–164.
[15] P. A. Chinnery and V. F. Humphrey, ''Resonances in acoustic scattering by cylindrical objects,'' Proc. I. O. A. **16**(6), 51–58 (1994).
[16] P. A. Chinnery and V. F. Humphrey, ''Experimental visualization of acoustic resonances within a stadium-shaped cavity,'' Phys. Rev. E **53**(1), 272–276 (1996).
[17] J. Wolf, T. H. Neighbors, and W. G. Mayer, ''Optical probing of ultrasonic pulses,'' Ultrasonics **27**, 150–154 (1989).
[18] S. M. Knapp, C. Beckett, and V. F. Humphrey, J. Acoust. Soc. Am. (in preparation).
[19] J. A. Bucaro, L. Flax, H. D. Dardy, and W. E. Moore, ''Image profiles in schlieren observations of acoustic wave fronts,'' J. Acoust. Soc. Am. **60**, 1079–1084 (1976).
[20] J. A. Bucaro, L. Flax, and E. F. Carome, ''Quantitative schlieren visualization of unipolar acoustic transients,'' J. Acoust. Soc. Am. **63**, 49–54 (1978).
[21] J. A. Bucaro and H. D. Dardy, ''Sensitivity of the schlieren method for visualization of low-frequency ultrasonic waves,'' J. Acoust. Soc. Am. **63**, 768–773 (1978).
[22] C. V. Raman and N. S. N. Nath, Proc. Indian Acad. Science, Vols. 2 and 3, Parts I–V (1935).
[23] W. R. Klein and B. D. Cook, ''Unified approach to ultrasonic light diffraction,''IEEE Trans. Sonics Ultrason. **SU-14**, 123–134 (1967).
[24] K. Van Den Abeele and O. Leroy, ''Light diffraction by ultrasonic pulses: analytical and numerical solutions of the extended Raman-Nath equations,'' J. Acoust. Soc. Am. **88**, 2298–2315 (1990).
[25] P. A. Chinnery and V. F. Humphrey, ''On the overlapping acoustic resonances of a fluid-filled cavity: schlieren visualization of an insonified circular-cylindrical shell,'' submitted to J. Acoust. Soc. Am. (1996).
[26] L. E. Hargrove, ''Effects of ultrasonic waves on Gaussian light beams with diameter comparable to ultrasonic wavelength,'' J. Acoust. Soc. Am. **43**, 847–851 (1968).
[27] W. G. Mayer and T. H. Neighbors, ''Acousto-optic method for ultrasonic pulse characterization,'' Ultrasonics **25**, 83–86 (1987).
[28] N. D. Veksler, *Resonance Acoustic Spectroscopy* (Springer-Verlag, Berlin, 1993).
[29] V. Raman and K. S. Venkataraman, ''Determination of the adiabatic piezo-optic coefficient of liquids,'' Proc. R. Soc. London, Ser. A **171**, 137–147 (1939).

Response of an embedded fiber optic ultrasound sensor

John Dorighi, Sridhar Krishnaswamy, and Jan D. Achenbach
Center for Quality Engineering and Failure Prevention, Northwestern University, Evanston, Illinois 60208

(Received 9 January 1996; revised 3 May 1996; accepted 28 August 1996)

In this work the response of an embedded fiber optic Fabry–Perot ultrasound sensor has been investigated. Calculations were performed that model the fiber as a cylindrical elastic inclusion subjected to obliquely incident harmonic plane waves. The phase shift of the light traveling in the fiber was calculated along the principal strain axes as a function of ultrasonic frequency and incident angle. Experiments were performed to validate the calculations. Induced phase shifts were measured for the cases of normally incident ultrasound impinging upon a fiber sensor submerged in water as well as embedded in epoxy. The case of obliquely incident ultrasound at angles up to 4 deg was also investigated for a fiber embedded in epoxy. The experimental measurements were performed over a bandwidth from 2 to 8 MHz. © *1997 Acoustical Society of America.* [S0001-4966(97)01801-8]

PACS numbers: 43.35.Sx [HEB]

INTRODUCTION

Fiber optic sensors are capable of detecting a range of phenomena including strain, temperature, and magnetic fields.[1,2] The small size, light weight, immunity to electromagnetic interference, corrosion resistance, and high temperature properties of fiber optic sensors offer advantages over traditional sensing elements, and has motivated their development. A particular area of current interest which utilizes the advantages of fiber optic sensors are fiber optic ''smart structures.''[3–7] A fiber optic ''smart structure'' incorporates an embedded network of fiber optic sensors for health monitoring.

A tool widely used in the field of nondestructive evaluation (NDE) for the health monitoring of structures is ultrasound. Ultrasound has been used to detect a variety of material defects,[8] and has also been demonstrated as a practical tool for process monitoring.[9] Consequently, the development of a fiber optic ultrasound sensor offers the potential to use the full power of ultrasound NDE in fiber optic ''smart structure'' applications.

A number of researchers have developed fiber optic interferometers which are capable of detecting ultrasonic fields inside a structure. Liu and Measures[10] have used a fiber optic Michelson interferometer for the detection of elastic waves. A limitation of the fiber optic Michelson sensor is the nonlocal nature of the sensor which results when the two lead-in sections of fiber experience different states of strain. A truly localized sensor which requires the use of only a single lead-in fiber is the extrinsic fiber optic Fabry–Perot interferometer. This configuration incorporates an air gap between two lengths of optical fiber in a glass capillary.[11] Murphy *et al.*[12] have used an extrinsic fiber optic Fabry–Perot sensor for the detection of ultrasound. A disadvantage of the extrinsic Fabry–Perot sensor is the signal fading which results from diffraction losses at the fiber/air interface when static strains increase the length of the sensing region. Consequently, the extrinsic fiber optic Fabry–Perot sensor has different sensitivities when detecting ultrasound under different states of strain. A localized fiber interferometer which does not experience signal fading is the intrinsic fiber optic Fabry–Perot (FOFP). The sensing region of this type of interferometer is formed by a continuous length of fiber separated from the lead-in fiber by an internal partial mirror.[13] Alcoz *et al.*[14] have used an unstabilized FOFP for the detection of ultrasound. However, maintaining the most sensitive operating point, quadrature, in the presence of low-frequency strains is a critical issue which must be addressed. Dorighi *et al.*[15] have maintained the quadrature point using an active homodyne technique which tunes the laser frequency which permits the small displacements associated with ultrasound to be detected.

Once a localized fiber optic interferometer is configured for the detection of ultrasound inside a structure, it is important to determine the ultrasonic response of the embedded fiber sensor. A number of authors have investigated the response of a fiber optic hydrophone submerged in water and an extensive review is given by Bucaro *et al.*[16] The low-frequency acoustic response (Hz-kHz) of a submerged fiber optic sensor has been investigated analytically[17] and experimentally.[18,19] Similarly, the high-frequency ultrasonic response (MHz) of a submerged fiber optic sensor has been investigated both analytically[20] and experimentally.[21]

Similar to the situation of a fiber optic hydrophone, the response of an embedded fiber optic ultrasound sensor is determined by the mechanical interaction between the fiber sensor and the host. However, unlike the situation of a submerged fiber optic sensor, shear strains in the host material are present and must be considered for an embedded fiber sensor. An appropriate model is essential when considering an embedded sensor because of the wide range of host materials in which it could be placed. This is in contrast to a fiber hydrophone where water is the typical host material. Once a model is experimentally validated for an embedded fiber optic ultrasound sensor it can be easily applied to a range of host materials with differing mechanical properties.

In this paper the ultrasonic response of a fiber optic interferometer embedded in epoxy has been characterized. A theoretical model was developed which predicts the response of the embedded fiber sensor as a function of the frequency and the incident angle of impinging ultrasound. Experiments

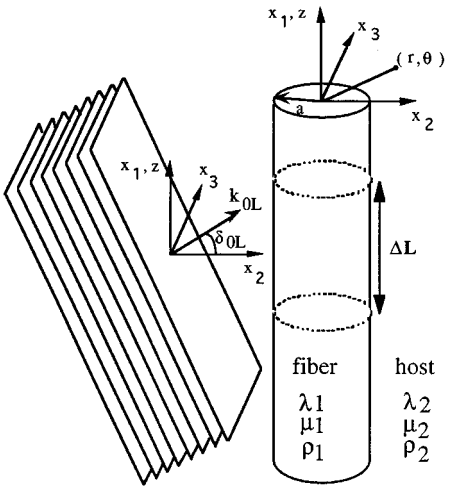


FIG. 1. Problem geometry.

carried out over a bandwidth from 2 to 8 MHz at incident angles of up to 4 deg, confirmed the theoretical calculations.

I. THEORY

The theoretical formulation is similar to that of Flax *et al.*,[20] who investigated a normally incident harmonic plane wave impinging upon an optical fiber submerged in water at ultrasonic frequencies. The situation we have considered is an obliquely incident harmonic plane wave impinging upon an embedded elastic cylinder, representing the optical fiber. The host and inclusion are assumed to be linearly elastic, isotropic, and homogeneous with different material properties. Additionally, we have assumed the differences in the elastic constants between the fiber core and fiber cladding to be negligible. A schematic of the geometry is shown in Fig. 1, where the axis of the cylinder is assumed to extend to infinity in the x_1 direction. For a normally incident plane wave, axial strains are not present in the theoretical formulation. However, axial strains must be considered for a harmonic plane wave obliquely incident upon an embedded fiber sensor.

The frequency response of the embedded fiber sensor was calculated in two steps. First, strains in the fiber core produced by the incident elastic wave were determined from elastodynamics.[22] Then the resulting phase shift of light was calculated using a phase-strain relation.[23] The response of the finite sensor was finally determined by integrating the induced phase shift over the insonified length of fiber.

A. Incident plane wave

The mechanical interaction between the elastic inclusion and the host must be modeled to determine the strains in the fiber core. Displacements are expressed in terms of a scalar and a vector wave potential as follows

$$\mathbf{u}=\nabla\Phi+\nabla\times\boldsymbol{\Psi}, \tag{1}$$

where $\mathbf{u}$ is the displacement, Φ is the scalar potential, and $\boldsymbol{\Psi}$ is the vector potential. Longitudinal waves are expressed in terms of the scalar potential while transverse waves are represented by the components of the vector potential. The obliquely incident longitudinal plane wave represented in cylindrical coordinates is given as[22]

$$\Phi^{\text{in}}=\frac{P_o}{\rho_2\omega^2}\sum_{n=0}^{\infty} e_n i^n J_n(\alpha_{0L}r)\cos(n\theta)e^{i(\beta z-\omega t)}, \tag{2}$$

where e_n is the neumann factor $e_O=1$, $e_n=2$ for $n=1,2,3$, etc.), i is $\sqrt{-1}$, α_{0L} is the radial propagation constant for the longitudinal wave in the host, β is the axial propagation constant, ω is the circular frequency, P_O is the pressure, and ρ_2 is the density of the host medium.

B. Scattered and refracted waves

At the fiber/host interface the incident wave excites longitudinal and transverse waves in the cylinder and scattered longitudinal and transverse waves which travel away from the cylinder. This requires the use of both scalar and vector wave potentials to represent the scattered and refracted waves. Solution forms for the scattered and refracted wave potentials in the host medium and scatterer are given as[24]

$$\Phi^{\text{refract}}=\sum_{n=0}^{\infty} A_n J_n(\alpha_{1L}r)\cos(n\theta)e^{i(\beta z-\omega t)}, \tag{3}$$

$$\Psi_z^{\text{refract}}=\sum_{n=0}^{\infty} B_n J_n(\alpha_{1T}r)\sin(n\theta)e^{i(\beta z-\omega t)}, \tag{4}$$

$$\Psi_r^{\text{refract}}=\sum_{n=0}^{\infty} C_n J_{n+1}(\alpha_{1T}r)\sin(n\theta)e^{i(\beta z-\omega t)}, \tag{5}$$

$$\Psi_\theta^{\text{refract}}=-\sum_{n=0}^{\infty} C_n J_{n+1}(\alpha_{1T}r)\cos(n\theta)e^{i(\beta z-\omega t)}, \tag{6}$$

$$\Phi^{\text{scatt}}=\sum_{n=0}^{\infty} D_n H_n^{(1)}(\alpha_{2L}r)\cos(n\theta)e^{i(\beta z-\omega t)}, \tag{7}$$

$$\Psi_z^{\text{scatt}}=\sum_{n=0}^{\infty} E_n H_n^{(1)}(\alpha_{2T}r)\sin(n\theta)e^{i(\beta z-\omega t)}, \tag{8}$$

$$\Psi_r^{\text{scatt}}=\sum_{n=0}^{\infty} F_n H_{n+1}^{(1)}(\alpha_{2T}r)\sin(n\theta)e^{i(\beta z-\omega t)}, \tag{9}$$

$$\Psi_\theta^{\text{scatt}}=-\sum_{n=0}^{\infty} F_n H_{n+1}^{(1)}(\alpha_{2T}r)\cos(n\theta)e^{i(\beta z-\omega t)}, \tag{10}$$

where A_n-F_n are the unknowns to be solved for.

The solution forms expressed in Eqs. (3)–(10) arise from solving the displacement equation of motion in cylindrical coordinates in terms of the scalar and vector wave potential using separation of variables. It is required that strains in the fiber as $r\to 0$ must remain finite, and that the scattered wave in the host must travel away from the embedded fiber. This determines whether the Bessel or Hankel function is chosen. Furthermore, the relation between the axial and radial propagation constants in Eqs. (3)–(10) is deduced from the analysis and shown Fig. 2. They are related to the incident wave number as follows:

$$\beta=k_{0L}\sin(\delta_{0L})=k_{iL}\sin(\delta_{iL})=k_{iT}\sin(\delta_{iT}), \tag{11}$$

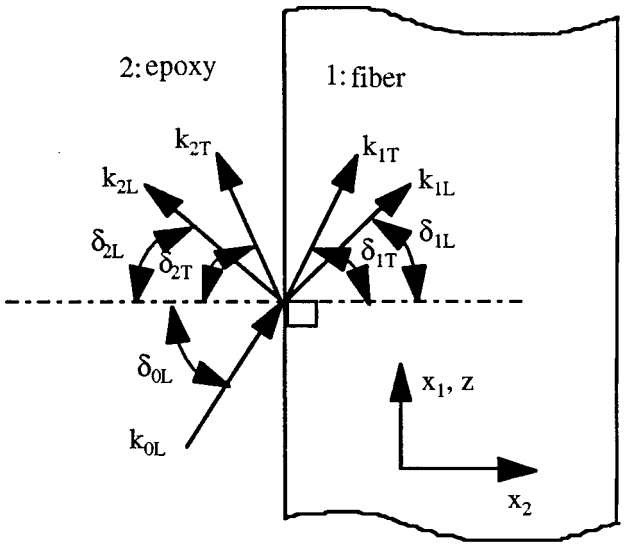


FIG. 2. Schematic of scattered and refracted waves at the host/fiber interface.

$$\alpha_{iL}=k_{iL}\cos(\delta_{iL}), \quad i=(1,2), \tag{12}$$

$$\alpha_{iT}=k_{iT}\cos(\delta_{iT}), \tag{13}$$

where k_{0L} is the incident wave number, α_{2L} and α_{1L} are the longitudinal wave radial propagation constants, respectively, in the host and cylinder, α_{2T} and α_{1T} are the transverse wave radial propagation constants, respectively, in the host and cylinder.

The refracted field inside the fiber is obtained by solving for A_n, B_n, and C_n using the traction and displacement continuity conditions at the cylinder and host interface. The continuity conditions are expressed as

$$\tau_{rr}^{\text{in}}+\tau_{rr}^{\text{scatt}}=\tau_{rr}^{\text{refract}} \quad \text{at} \quad r=a, \tag{14}$$

$$\tau_{r\theta}^{\text{in}}+\tau_{r\theta}^{\text{scatt}}=\tau_{r\theta}^{\text{refract}} \quad \text{at} \quad r=a, \tag{15}$$

$$\tau_{rz}^{\text{in}}+u_{rz}^{\text{scatt}}=\tau_{rz}^{\text{refract}} \quad \text{at} \quad r=a, \tag{16}$$

$$u_{r}^{\text{in}}+u_{r}^{\text{scatt}}=u_{r}^{\text{refract}} \quad \text{at} \quad r=a, \tag{17}$$

$$u_{\theta}^{\text{in}}+u_{\theta}^{\text{scatt}}=\tau_{\theta}^{\text{refract}} \quad \text{at} \quad r=a, \tag{18}$$

$$u_{z}^{\text{in}}+u_{z}^{\text{scatt}}=u_{z}^{\text{refract}} \quad \text{at} \quad r=a, \tag{19}$$

where a is the radius of the embedded fiber. Using Hooke's law and Eq. (1), the stresses and displacements in Eqs. (14)–(19) can be expressed in terms of the scalar and vector wave potentials. Upon substitution of Eqs. (3)–(10) into Eqs. (14)–(19) a set of six equations for the six sets of unknowns A_n, B_n, C_n, D_n, E_n, and F_n results. The required coefficients can be determined by applying the orthogonality condition, to eliminate the summation, and then inverting the resulting system of equations for the required values of n.

C. Strain in the fiber core

In order to evaluate the phase shift of light in the fiber, we need to obtain the strains in the fiber once the scalar and vector potentials are determined. Simplified expressions for the strains in the fiber core result when we recognize that the core diameter of a single mode fiber is on the order of 4–5 μm, while the cladding diameter is 125 μm. This suggests that the limit of the strains should be investigated as $r\to 0$, where the Bessel functions can be replaced by their small value approximations as $r\to 0$. Upon simplification, the strains in the fiber core can be expressed as

$$\varepsilon_j=\overline{\varepsilon_j}e^{i(\beta z-\omega t)}, \quad j=(1,2,3), \tag{20}$$

where

$$\overline{\varepsilon_1}=[-A_0\beta^2-iC_0\beta\alpha_{1T}],$$

$$\overline{\varepsilon_2}=\left[-A_0\frac{\alpha_{1L}^2}{2}+iC_0\frac{\alpha_{1T}\beta}{2}+\left(A_2\frac{\alpha_{1L}^2}{4}+B_2\frac{\alpha_{1T}^2}{4}\right)\right],$$

$$\overline{\varepsilon_3}=\left[-A_0\frac{\alpha_{1L}^2}{2}+iC_0\frac{\alpha_{1T}\beta}{2}-\left(A_2\frac{\alpha_{1L}^2}{4}+B_2\frac{\alpha_{1T}^2}{4}\right)\right],$$

where ε_2 and ε_3 are the secondary principal strains, and ε_1 is the strain along the axis of the fiber.

D. Phase strain relation

The following expressions relate the strain in the fiber core to the phase shift of light traveling in the fiber:[23]

$$\Delta\phi^{x2}=2K_o n\int_{L1}^{L2}\left\{\varepsilon_1-\frac{n^2}{2}[p_{11}\varepsilon_2+p_{12}(\varepsilon_1+\varepsilon_3)]\right\}dx_1, \tag{21}$$

$$\Delta\phi^{x3}=2K_o n\int_{L1}^{L2}\left\{\varepsilon_1-\frac{n^2}{2}[p_{11}\varepsilon_3+p_{12}(\varepsilon_1+\varepsilon_2)]\right\}dx_1, \tag{22}$$

where $\Delta\phi^{x2}$ and $\Delta\phi^{x3}$ are the phase shift of light polarized along the two secondary principal strain axes, n is the refractive index of the fiber core, K_o is the free-space wave number of light, p_{11} and p_{12} are the Pockel's strain-optic coefficients of the fiber material, and L_2-L_1 is the gauge length of the sensor over which the ultrasound interacts with the fiber. Upon inspection of Eqs. (21) and (22) it is noticed that the only effects which contribute to the phase shift of light are changes in the sensor gauge length produced by the axial strain, ε_1, and changes in the refractive index of the fiber produced by both the axial strain, ε_1, and the secondary principal strains, ε_2 and ε_3.[23] Both these effects must be considered because an obliquely incident wave produces both transverse and axial strains in the fiber. Substituting the strains given in Eq. (20) into the phase-strain relation Eqs. (21) and (22) and performing the required integration, yields the magnitudes of the phase shifts in the fiber core which are expressed as follows:

$$|\Delta\phi^{x2}|=|\Delta\overline{\phi^{x2}}|\Delta L\frac{|\sin(\beta\Delta L/2)|}{(\beta\Delta L/2)}, \tag{23}$$

$$|\Delta\phi^{x3}|=|\Delta\overline{\phi^{x3}}|\Delta L\frac{|\sin(\beta\Delta L/2)|}{(\beta\Delta L/2)}, \tag{24}$$

where

TABLE I. Physical constants used to calculate the frequency response of fiber sensor embedded in epoxy.

radius of embedded fiber—a	62.5 μm
density of epoxy—ρ_2	1180 kg/m^3
density of fiber—ρ_1	2200 kg/m^3
Lame elastic constant of epoxy—λ_2	4.0 GPa
Lame elastic constant of fiber—λ_1	13.5 GPa
Lame elastic constant of epoxy—μ_2	2.4 GPa
Lame elastic constant of fiber—μ_1	30.3 GPa
Pressure of incident plane wave—P_O	10 kPa
refractive index of fiber core—n	1.458
Pockel's strain-optic constant of fiber—p_{11}	0.113
Pockel's strain-optic constant of fiber–p_{12}	0.252
gauge length of sensor–ΔL	0.01 m

$$|\Delta\overline{\phi^{x2}}| = |2K_o n\{\overline{\varepsilon_1} - (n^2/2)[p_{11}\overline{\varepsilon_2} + p_{12}(\overline{\varepsilon_1} + \overline{\varepsilon_3})]\}|,$$

$$|\Delta\overline{\phi^{x3}}| = |2K_o n\{\overline{\varepsilon_1} - (n^2/2)[p_{11}\overline{\varepsilon_3} + p_{12}(\overline{\varepsilon_1} + \overline{\varepsilon_2})]\}|,$$

and ΔL is the gauge length of the sensor. It is seen from Eqs. (23), (24) that the induced phase shift is zero when

$$\frac{\beta\Delta L}{2} = \frac{k_{0L}\sin(\delta_{0L})\Delta L}{2} = m\pi, \tag{25}$$

where m is an integer. Physically this implies there are an integral number of ultrasonic wavelengths along the axis of the fiber, resulting in a net phase shift of zero.

Using the above formulation, phase shifts were computed for ultrasonic frequencies up to 10 MHz and incident angles up to 4 deg for the physical parameters given in Table I.

II. EXPERIMENT

The fiber optic interferometer was fabricated by fusion splicing a length of polarization maintaining (PM) single mode fiber to a short length ($\sim$5 cm) of ordinary single mode fiber with a partial dielectric mirror deposited on the output end. A schematic of the fiber sensor in an epoxy block is given in Fig. 3. The low finesse Fabry–Perot interferometer is formed by reflections between the air/glass interface at the fiber input and the partial dielectric mirror on the other end of the fiber. While the fiber interferometer is sensitive along its entire length ($\sim$60 cm), the experimental setup dictates that ultrasound only interacts with the ordinary single mode fiber embedded in the epoxy block. Consequently, the effective gauge length of the sensor is the length over which ultrasound interacts with the embedded portion of the fiber and is determined by the ultrasonic beamwidth at the sensor. The sensor gauge length used in the theoretical calculations is similar to the ultrasonic interaction length in the experiment in order to facilitate a comparison between the two. The $-$6-dB beamwidth of the ultrasonic transducer at the epoxy block was measured to be 1.0 cm, which is consistent with the gauge length of 1 cm used for the theoretical calculations.

There are two additional assumptions made in the theoretical analysis which must be met in the experiment. Specifically, the polarization of light interacting with the ultrasound must be linear and must coincide with either of the ultrasound-induced secondary principal strain axes. Also, the segment of the fiber that is insonified by the ultrasound must be a homogeneous isotropic cylinder. It is difficult to obtain both these conditions simultaneously in the experiment. This is because the polarization of light traveling down an ordinary single mode will in general become elliptical.[25] To avoid this, polarization maintaining (PM) fibers can be used. However, such fibers have birefringence-inducing stress inclusions on either side of the fiber core. This results in a fiber that can no longer be simply treated as a homogeneous cylindrical inclusion as assumed in the elastodynamic calculations. To overcome these difficulties, the fiber interferometer used in these experiments consisted of a PM fiber in the lead-in segment of the sensor to which was fused a short length of ordinary single mode fiber. The PM lead-in fiber segment ensured that the polarization of light reaching the interaction length is linear. Furthermore, only the segment of the sensor that was made from ordinary single mode fiber was insonified by the ultrasound. This arrangement therefore satisfactorily met the theoretical assumptions as discussed in Dorighi.[26]

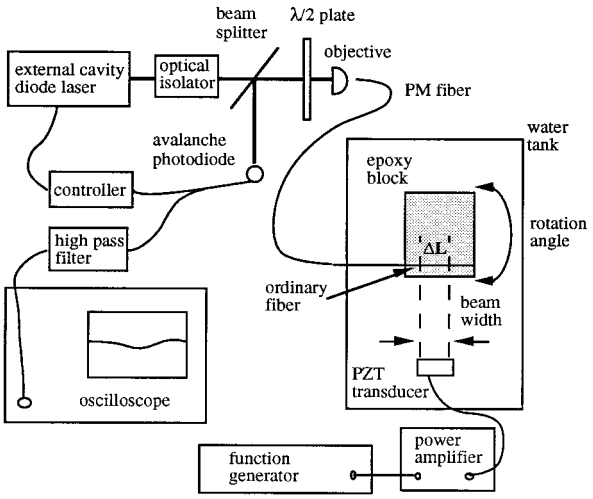


FIG. 3. Schematic of experimental setup.

A schematic of the experimental setup is shown in Fig. 3. Light from an external cavity diode laser (New Focus at 780 nm) passes through an optical isolator and is focused into the fiber optic sensor using a microscope objective. A half-wave plate was used which allowed the polarization of light coupled into the sensor to be adjusted. Light reflected back from the fiber sensor was directed toward an avalanche photodiode (APD) using a beamsplitter.

The APD was used to detect both the low-frequency (Hz-kHz) intensity variations produced by strain and temperature effects as well as the high-frequency (MHz) variations produced by ultrasound. High-frequency signals were stored on a digital oscilloscope, while the low-frequency signals were input to a feedback loop which actively controlled the laser frequency to maintain the quadrature point.[15] Quadrature is the linear portion of the interferometer response curve and produces the largest change in reflected intensity for a given phase shift. It should be noted that once the interferometer is stabilized at quadrature the detected intensity variations are proportional to the phase shift of light in the fiber sensor for the small phase shifts associated with

ultrasound. This allows for a comparison between the experiment and theoretical model to be made.

Ultrasound was generated using an unfocused immersion piezoelectric transducer (Panametrics) with a center frequency of 5 MHz and 0.5 in. diam. An arbitrary function generator was used to provide a constant amplitude frequency sweep from 2 to 8 MHz over 40 μs. The voltage from the function generator was passed through a 50-dB power amplifier to drive the transducer.

A. Fiber submerged in water

In a first set of experiments, measurements of the response of a fiber submerged in water were made. This allowed comparison with similar measurements performed by DePaula *et al.*[27] Great care needed to be taken in aligning the fiber polarization axes with the propagation direction of the incident ultrasound. The fiber sensor under investigation was held in a rotatable fiber chuck and was aligned in the holding fixture before measurements were performed. Once the fiber was properly aligned in the fiber chuck the phase shifts along the two principal strain axes could be investigated simply by rotating the half-wave plate.

Measurements of the response of a fiber submerged in water were only performed for the case of normally incident ultrasound.

B. Fiber embedded in epoxy

The response of a sensor embedded in epoxy was measured in a second set of experiments. The embedded fiber polarization axes were aligned parallel and perpendicular to the front face of the epoxy block by aligning the fiber sensor in the aluminum mold prior to pouring the epoxy. The dimensions of the epoxy block were 4 cm$\times$5 cm$\times$10 cm with the fiber embedded 6 to 10 mm from the front face of the block, depending on the specimen. The epoxy block was long enough to insure that the entire ultrasonic signal was detected before the first reflection from the backface arrived. It should be noted that all ultrasonic measurements were performed with the fiber sensor placed outside the transducer near-field distance as specified by the transducer manufacturer.

The epoxy block was submerged in a water tank and supported on a rotation fixture which enabled precise adjustment of the incident angle. The rotation fixture was adjusted for normally incident ultrasound by maximizing the amplitude of the signal detected at 8 MHz. The angles for obliquely incident ultrasound were determined by minimizing the ultrasonic response at the frequency where zero response was predicted by the theory. This procedure was necessary because the refraction of ultrasound at the water/epoxy interface made direct measurements of the incident angle in epoxy difficult. Ultrasonic signals were acquired with the light polarized parallel and perpendicular to the front face of the epoxy block for incident ultrasound angles from 0 to 4 deg. It should be noted that all of the ultrasonic signals detected using the fiber sensor were averaged 1000 times to improve the signal-to-noise ratio.

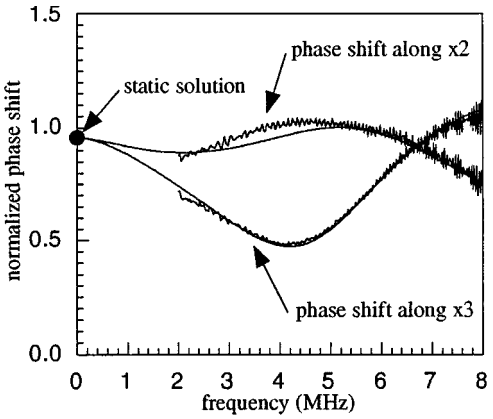


FIG. 4. Response of fiber sensor submerged in water to normally incident ultrasound.

III. DECONVOLUTION

The frequency response of the fiber sensor needed to be deconvolved from the response of the system before a comparison with the theoretical model could be made. The spectrum magnitude of the waveform detected by the embedded fiber sensor can be described by

$$V_{\mathrm{fib}}(\omega)=V_{\mathrm{in}}(\omega)G(\omega)S_{\mathrm{wat}}(\omega)T_{w/e}(\omega)S_{\mathrm{epox}}(\omega)F(\omega), \tag{26}$$

where $V_{\mathrm{fib}}(\omega)$ is the spectrum magnitude of the signal detected by the fiber, $V_{\mathrm{in}}(\omega)$ is the magnitude of the voltage spectrum input to the transducer, $G(\omega)$ is the transducer response, $S_{\mathrm{wat}}(\omega)$ is the frequency-dependent attenuation of water, $T_{w/e}(\omega)$ is the transmission coefficient into the epoxy, $S_{\mathrm{epox}}(\omega)$ is the frequency-dependent attenuation of epoxy, and $F(\omega)$ is the frequency response of the embedded fiber sensor. In order to determine the frequency response of the embedded fiber sensor, the attenuation of the water is neglected and transmission from the water into epoxy is assumed constant with frequency. Both of these are reasonable assumptions in the frequency range under consideration.

The remaining terms were measured in a separate set of experiments. The constant amplitude frequency sweep from the function generator was acquired and stored using a digital oscilloscope. The transducer response was measured using a pair of matched 5-MHz unfocused immersion transducers in a pitch–catch arrangement. In this configuration the generating transducer was driven with the same frequency sweep used for the fiber measurements, while the second transducer detected the signal.

The ultrasonic attenuation in epoxy was measured over the frequency range of interest using a technique described by Papadakis.[28] It should be noted that the attenuation measurement was performed using an epoxy sample sliced from the same epoxy block in which the fiber sensor was embedded. A linear fit of the measured epoxy attenuation versus frequency was used to simplify the deconvolution. Finally, the response of the embedded fiber sensor can be expressed in terms of the measured quantities as follows

$$F(\omega)=A\frac{V_{\mathrm{fib}}(\omega)}{S_{\mathrm{epox}}(\omega)\sqrt{V_{\mathrm{in}}(\omega)V_G(\omega)}}, \tag{27}$$

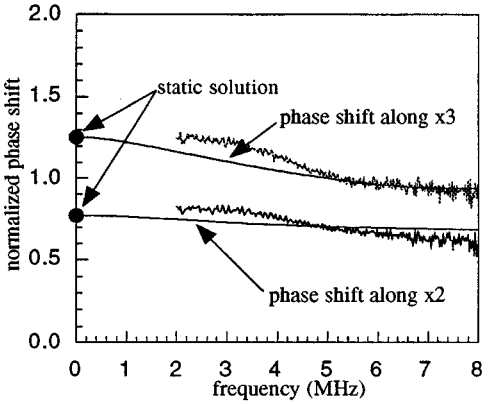


FIG. 5. Response of fiber sensor embedded in epoxy to normally incident ultrasound.

where $V_G(\omega)$ is the spectrum magnitude of the signal from the transducer calibration, and A is a constant related to the incident acoustic pressure which is unknown but can be removed by appropriate normalization.

IV. RESULTS AND DISCUSSION

A. Response of fiber submerged in water

The measured and calculated response of a fiber sensor submerged in water to normally incident ultrasound is shown in Fig. 4. It should be noted that the numerical results displayed in Fig. 4 are identical to those reported by DePaula *et al.*[27] A direct comparison between the theoretical and experimental data shown in Fig. 4 was facilitated by normalizing the calculated and measured results with respect to the response along the x_2 axis at 5 MHz. Ultrasound generated at the transducer center frequency has the largest amplitude and provided a good reference for normalization. From Fig. 4 we see excellent agreement between the measured and calculated results. At lower frequencies as the ultrasonic wavelength becomes larger than the fiber diameter, the response along both the secondary principal strain axes tend toward the same limit. This is expected since at lower frequencies the fiber experiences a uniform radial pressure which produces the same phase shift for all polarizations of light. However, at higher frequencies when the ultrasonic wavelength is comparable to the fiber diameter it is apparent that strain variations across the fiber diameter become important.

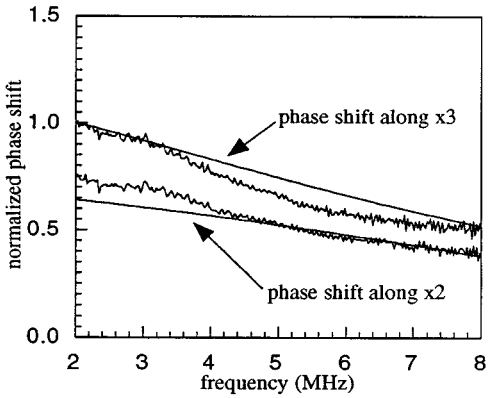


FIG. 6. Embedded sensor response at 1 deg.

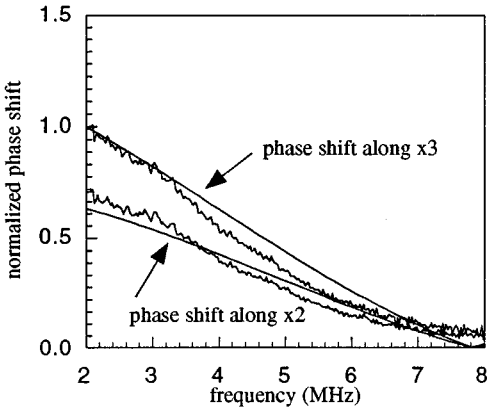


FIG. 7. Embedded sensor response at 2 deg.

Also shown in Fig. 4 is the plane strain static solution for a fiber experiencing a uniform radial pressure. It is seen that the elastodynamic solution tends toward the static case at lower frequencies.

B. Response of fiber embedded in epoxy

The measured and calculated response of an embedded fiber with ultrasound normally incident upon the front face of the epoxy block is shown in Fig. 5. In this figure both the experimental and theoretical results are normalized with respect to the response along the x_3 secondary principal strain axis at 5 MHz. Again we see good agreement between the calculated and measured results. It is clear from Fig. 5 that at low frequencies the phase shifts induced along the two secondary principal strain axes do not tend toward the same value. This is expected because the shear stresses present in the solid host prevent the fiber from experiencing a uniform radial pressure even as the wavelength becomes larger than the fiber diameter. Also shown in Fig. 5 is the plane strain static solution for an embedded fiber experiencing a uniaxial displacement applied along the x_2 axis. It is observed that the elastodynamic solution tends toward the static solution at lower frequencies.

Shown in Figs. 6–9 are the measurements and calculations for ultrasound incident upon the embedded fiber at angles from 1 to 4 deg. All of these plots were normalized relative to the response along the x_3 axis at 2 MHz. The fiber response at 2 MHz was the largest for obliquely incident

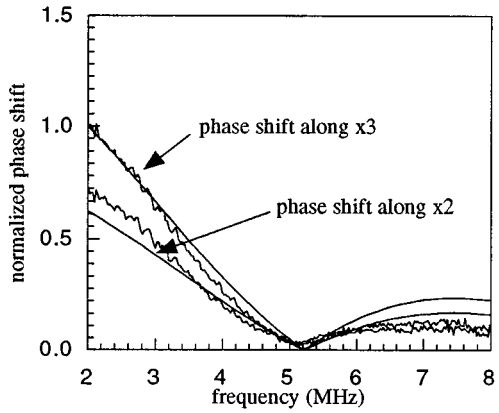


FIG. 8. Embedded sensor response at 3 deg.

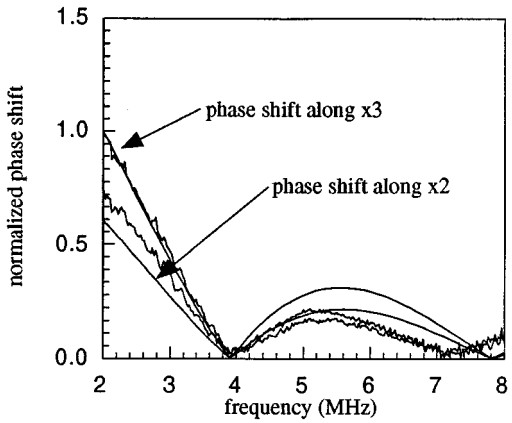


FIG. 9. Embedded sensor response at 4 deg.

ultrasound and provided a good reference for normalization. The measurements at the four different incident angles follow the trend predicted by the calculations. The fiber sensor response decreases with frequency as the incident angle of the ultrasound increases. Physically, this decrease can be attributed to the fact that the response goes to zero when the length of the interaction region is an integral number of ultrasonic wavelengths along the axis of the fiber. This is consistent with the measurements made and is in contrast to the case of normal incidence where the entire interaction length of the fiber experiences the same state of strain at a given point in time, causing phase shifts to accumulate.

V. CONCLUSIONS

The response of an embedded fiber optic ultrasound has sensor has been investigated. A theoretical model was developed which predicts the ultrasonic response of a fiber sensor embedded in a host material. The model was verified experimentally for the situation of an optical fiber embedded in epoxy. The agreement between the experimental results and the theoretical model suggest that the model can be applied to investigate the ultrasonic response of a fiber optic sensor embedded in a range of host materials.

The case of normally incident ultrasound impinging upon a fiber submerged in water and a fiber embedded in epoxy were considered, along with ultrasound obliquely incident upon an embedded fiber sensor. The results show that an embedded sensor has broadband response at normal incidence of ultrasound. For oblique incidence, however, the response of the embedded sensor rolls off rapidly at high frequencies. In addition, at oblique incidence the sensor is ''blind'' to ultrasound of certain frequencies. This suggests that embedded fiber optic ultrasound sensors are most useful at near normal incidence. Their broadband response at normal incidence can be taken advantage of in situations where the location of the ultrasonic source is controllable. Fortunately, this is the case in most NDE ultrasound applications.

ACKNOWLEDGMENT

This work was supported by the Air Force Office of Sponsored Research under Award F4960-92-0342AFOSR.

[1] D. A. Jackson, ''Monomode optical fiber interferometers for precision measurement,'' J. Phys. E **18**, 981–1001 (1985).
[2] D. A. Jackson, ''Recent progress in monomode fiber-optic sensors,'' Meas. Sci. Technol. **5**, 621–638 (1994).
[3] E. Udd (Ed.), ''Fiber Optic Smart Structures and Skins I,'' SPIE Proc. **986** (1988).
[4] E. Udd (Ed.), ''Fiber Optic Smart Structures and Skins II,'' SPIE Proc. **1170** (1989).
[5] E. Udd (Ed.), ''Fiber Optic Smart Structures and Skins III,'' SPIE Proc. **1370** (1990).
[6] E. Udd (Ed.), ''Fiber Optic Smart Structures and Skins IV,'' SPIE Proc. **1588** (1991).
[7] E. Udd (Ed.), ''Fiber Optic Smart Structures and Skins V,'' SPIE Proc. **1798** (1993).
[8] J. Krautkramer and H. Krautkramer, *Ultrasonic Testing of Materials* (Springer-Verlag, Berlin, 1969).
[9] L. C. Lynnworth, *Ultrasonic Measurements for Process Control* (Academic, Boston, 1989).
[10] K. Liu and R. M. Measures, ''Detection of high-frequency elastic waves with embedded ordinary single-mode fibers,'' *Fiber Optic and Laser Sensors IX* (SPIE Proceedings, Bellingham, WA, 1991), Vol. 1584, pp. 226–234.
[11] K. A. Murphy, M. F. Gunther, A. M. Vengsarkar, and R. O. Claus, ''Quadrature phase-shifted extrinsic Fabry–Perot optical fiber sensors,'' Opt. Lett. **16**(4), 273–275 (1991).
[12] K. A. Murphy, R. O. Claus, J. A. Greene, and T. A. Tran, ''Acoustic wave response of the extrinsic Fabry–Perot interferometric (EFPI) optical fiber sensor,'' *Adaptive Structures and Material Systems* (ASME, 1993), AD-Vol. 35, pp. 395–399.
[13] C. E. Lee, R. A. Atkins, andH. F. Taylor, ''Performance of a fiber-optic temperature sensor from −200 to 1050 °C,'' Opt. Lett. **13**(11), 1038–1040 (1988).
[14] J. J. Alcoz, C. E. Lee, and H. F. Taylor, ''Embedded Fiber–Optic Fabry–Perot Ultrasound Sensor,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **37**(4), 302–306 (1990).
[15] J. Dorighi, S. Krishnaswamy, and J. D. Achenbach, ''Stabilization of an embedded fiber optic Fabry–Perot sensor for ultrasound detection,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42**(5), 820–824 (1995).
[16] J. A. Bucaro, N. Lagakos, J. H. Cole, and T. G. Giallorenzi, *Fiber Optic Acoustic Transduction, Physical Acoustics*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1982), Vol. 16, Chap. 7, p. 385.
[17] B. Budiansky, D. C. Drucker, G. S. Kino, and J. R. Rice, ''Pressure sensitivity of a clad optical fiber,'' Appl. Opt. **18**(24), 4085–4088 (1979).
[18] J. A. Bucaro, H. D. Dardy, and E. F. Carome, ''Optical fiber acoustic sensor,'' Appl. Opt. **16**(7), 1761–1762 (1977).
[19] J. H. Cole, R. L. Johnson, and P. G. Bhuta, ''Fiber-optic detection of sound,'' J. Acoust. Soc. Am. **62**, 1136–1138 (1977).
[20] L. Flax, J. H. Cole, R. P. De Paula, and J. A. Bucaro, ''Acoustically induced birefringence in optical fibers,'' J. Opt. Soc. Am. **72**, 1159–1162 (1982).
[21] R. P. De Paula, J. H. Cole, and J. A. Bucaro, ''Broad-band ultrasonic sensor based on induced optical phase shifts in single-mode fibers,'' J. Lightwave Technol. **LT-1**(2), 390–393 (1983).
[22] R. M. White, ''Elastic Wave Scattering at a Cylindrical Discontinuity in a Solid,'' J. Acoust. Soc. Am. **30**, 771–785 (1958).
[23] J. S. Sirkis ''Unified approach to phase-strain-temperature models for smart structure interferometric optical fiber sensors: Part 1, development,'' Opt. Eng. (Bellingham) **32**(4), 752–761 (1993).
[24] L. Flax, V. K. Varadan, and V. V. Varadan, ''Scattering of an obiliquely incident acoustic wave by an infinite cylinder,'' J. Acoust. Soc. Am. **68**, 1832–1835 (1980).
[25] J. Wilson and J. F. B. Hawkes, *Optoelectronics, An Introduction* (Prentice-Hall, New York, 1989), p. 423.
[26] J. Dorighi, ''Development, Characterization, and Application of an Embedded Fiber Optic Ultrasound Sensor,'' Ph.D. thesis, Northwestern University, pp. 50–56, 1996.
[27] R. P. De Paula, J. H. Cole, and J. A. Bucaro, ''Broad-band ultrasonic sensor based on induced optical phase shifts in single-model fibers,'' J. Lightwave Technol. **LT-1**(2), 390–393 (1983).
[28] E. P. Papadakis, ''Buffer-rod system for ultrasonic attenuation measurements,'' J. Acoust. Soc. Am. **44**, 1437–1441 (1968).

Acoustic backscatter from materials with rough surfaces and finite size microstructure: Theory

Mehmet Bilgen[a)]
Department of Radiology, University of Kansas Medical Center, Kansas City, Kansas 66160

James H. Rose[b)]
Center for NDE, Iowa State University, Ames, Iowa 50011

(Received 19 February 1996; accepted for publication 16 July 1996)

Surface roughness changes the acoustic backscatter signal due to the microstructure of a sample. These changes have been previously reported for the backscatter power under the restrictive assumption that the microstructural length scales are much smaller than any other length scale in the problem. In this study this restriction is removed and an approximate analytic series solution is presented that describes the effects of surface roughness on the power and mean square of the acoustic backscatter from samples whose microstructure is characterized by an autocorrelation length L_m, which may be smaller than, comparable to, or larger than either the wavelength or the autocorrelation length L_s that describes the surface roughness. Both focused and unfocused phase-sensitive transducers are considered. A surprising result emerged. For focused probes at normal incidence and for moderate roughness, the backscattered power at the focal depth is almost entirely determined by the frequency and the rms surface height; it is independent of L_m, L_s, and the radius of the transducer. © *1997 Acoustical Society of America.* [S0001-4966(97)00201-4]

PACS numbers: 43.35.Zc [HEB]

INTRODUCTION

Ultrasound is widely used to inspect samples. Typical inspections use pulse-echo, phase-sensitive piezoelectric transducers. The part is commonly immersed in a water bath, which acts in part as a transmitting fluid and in part as an impedance matching layer. Most samples are inhomogeneous on a relatively fine length scale; e.g., the cell size in tissue and the microcrystallite size in polycrystalline metals. The acoustic signal backscattered from these fine inhomogeneities is referred to as the acoustic backscatter signal. The acoustic backscatter signal has been used to characterize various material properties such as the contractility of myocardium or the crystallite size of single-phase structural metals. The acoustic backscatter also interferes with the detection of larger discrete scatterers such as cracks, voids, and inclusions; in this context, it is often referred to as material noise.

An extensive literature exist both for the reflection of acoustic waves from rough surfaces and for acoustic backscatter from microstructure. Reviews of the reflection problem are provided by the books of Ogilvy[1] and of Clay and Medwin.[2] A review of acoustic backscatter from biological tissues can be found in the compendium by Shung and Thieme.[3] Several authors discuss the backscatter of ultrasound from the microstructure of metals.[4–7] Fewer references are available for the problem of acoustic transmission through rough surfaces (see a brief discussion in Ref. 2 and the work of Nagy and co-workers[8,9]).

The effects of rough surfaces on acoustic backscatter was broached in an experimental/theoretical paper by Nagy and Rose.[9] The present authors later published two theoretical papers concerning the effects of rough surfaces on the acoustic backscatter.[10,11] The first of these theoretical papers was a numerical study of one-dimensional randomly rough surfaces.[10] The second paper presented an approximate analytic series solution for the acoustic backscatter power for two-dimensional randomly rough surfaces.[10] In both papers, the problem was simplified by assuming that the characteristic dimensions of the microstructure were much smaller than any other length scale in the problem.

In this paper, we generalize our previous results in a straightforward manner by removing the simplifying restriction on the microstructural length scale. Rather, we assume that the microstructure is described by a Gaussian random process and by a Gaussian autocorrelation function with autocorrelation length L_m. An approximate analytic series solution for the power of the acoustic backscatter is presented for this case and evaluated for a variety of parameters. In particular, we assume that L_m may be smaller than, comparable to, or larger than either the wavelength or the autocorrelation length, L_s, that describes the surface roughness. However, we will continue to assume that the scattering from the microstructure is very weak and that the fluctuations (scintillations) in the beam arise entirely from the rough surface. Experimental tests of our predictions are not available.

There are many examples of materials with both rough surfaces and microstructure. Biological materials such as cartilage and myocardium come immediately to mind, as do structural materials such as rock and cement. Even structural metals are generally polycrystalline aggregates with the size

[a)]Permanent address: The Scientific and Technical Research Council of Turkey, Ankara, Turkey.
[b)]Electronic mail: jrose@cnde.iastate.edu

of the average microcrystal, $\sim L_m$, ranging from 10^{-3} mm to 10 mm in an industrial metal such as aluminum. Production processes for such metals produce rough surfaces at intermediate stages of production. For example, the sand casting of aluminum produces rough surfaces with average surface heights in the range of 0.01 mm to 0.1 mm and with surface autocorrelation lengths L_s ranging from 0.1 to 1.0 mm, while surface finishing may produce L_s values of a few millimeters. At 10 MHz, the wavelength of sound in water is 0.15 mm.

In the main text, we concentrate on the predictions of the normalized mean backscattered signal power since this quantity connects with current experimental techniques. However, our method of analysis also yields an approximate analytic series solution for the effects of roughness on the mean square of the backscatter signal. These results are included in the Appendix for completeness and are presented graphically as a function of L_m, L_s, and the radius of the transducer for both focused and unfocused beams.

Qualitatively, one expects rather different results for the acoustic backscatter, depending on the relative size of the microstructure and surface roughness and depending on whether the transducer is focused or not. The most important effects of the microstructure and the rough surface are (1) to randomize the phase of the ultrasonic wave and (2) to scatter the wave out of the beam pattern. In the absence of roughness, phase randomization is determined by the microstructure and its autocorrelation length L_m. Roughness randomizes the phase by a small amount compared to the more rapidly varying microstructure if $L_s \gg L_m$. In this case, since microstructure mainly causes the phase randomization, the acoustic backscatter is weakly affected by surface roughness. On the other hand, if $L_s \ll L_m$, roughness is the main cause of phase randomization. In this case, roughness can change the acoustic backscatter by large amounts. In addition to these general considerations based on the relative size of the roughness and microstructure, one must consider the tendency of the surface roughness to scatter energy outside the original beam pattern. This tendency dominates for focused probes due to the small size of the beam pattern in the focal region and leads to simple formulas for acoustic backscatter from the focal depth.

The structure of the paper is as follows. In Sec. I, we describe the problem and introduce various simplifying approximations. In Sec. II, we derive and present an approximate analytic series solutions for the power of the acoustic backscatter signal. Numerical calculations of the effects of roughness on acoustic backscatter are presented and illustrated graphically in Sec. III. Section IV contains a summary and discussion of the paper. An approximate series solution for the mean square of the backscatter signal is contained in the Appendix.

I. PROBLEM SETUP AND APPROXIMATIONS

The geometry analyzed in this paper and typical of ultrasonic immersion inspections is shown in Fig. 1. A phase-sensitive transducer is immersed in water and oriented normal to the rough surface of a sample plate at a set-off of several centimeters. A pressure toneburst consisting of several wavelengths is excited in water and propagates to the plate's surface. After interacting with the rough surface, the wave field transmits into the sample, and scatters from distributed volumetric scatterers. The resulting reflected waves then transmit through the rough surface again and are detected by the same transducer. As suggested by the figure, for a well-collimated transducer, the sound primarily forms a sharp pulse in the solid with a width of several wavelengths. Roughly, this pulse is at a distance beneath the surface, $z \approx c_0 t/2$, where c_0 is the average velocity of sound in the sample and t denotes the time since the pulse crossed the interface. We will use this time–distance relation, and the assumption of single scattering from the microstructure to approximate the time evolution of the backscattered signal. We denote the coordinate parallel to the surface by $\mathbf{r}$, and the general three-dimensional coordinate by $\mathbf{x}=(\mathbf{r},z)$.

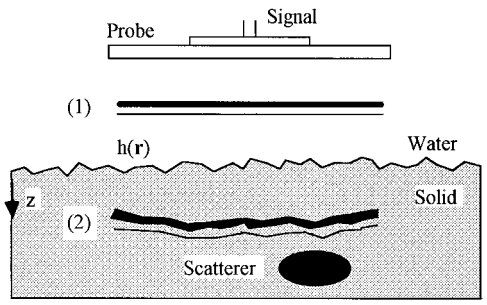


FIG. 1. Schematic geometry of hypothetical ultrasonic inspection. Shows backscatter of ultrasound from inhomogeneous background. (1) represents the plane wave and (2) represents the transmitted wave.

Backscattering is assumed to result from spatial variations in the sound velocity, $c(\mathbf{x})$, of the sample's microstructure. The effects of sound velocity variation on scattering enters via the velocity deviation

$$\delta\nu(\mathbf{x}) \equiv 1 - \frac{c_0^2}{c^2(\mathbf{x})}, \tag{1}$$

where $c_0^{-2} = \langle c^{-2}(\mathbf{x}) \rangle_m$. The angular bracket $\langle\ \rangle_m$ denotes an average over an ensemble of microstructures. The ensemble average will typically be estimated by spatial averages, which can be obtained by translating the transducer parallel to the surface in actual experiments. The velocity deviation, $\delta\nu$, is assumed to arise from a spatially uniform, Gaussian, random process of zero mean that is characterized by the autocorrelation function

$$K(\mathbf{x}-\mathbf{x}') \equiv \frac{\langle \delta\nu(\mathbf{x})\,\delta\nu(\mathbf{x}') \rangle_m}{\langle \delta\nu^2 \rangle_m}, \tag{2}$$

where $\langle \delta\nu^2 \rangle_m$ denotes the variance of the velocity deviation. We assume that the microstructural autocorrelation function can be approximated by a Gaussian function

$$K(\mathbf{x}-\mathbf{x}') = \exp(-|\mathbf{x}-\mathbf{x}'|^2/L_m^2). \tag{3}$$

Thus, the microstructure will be parameterized by its autocorrelation length L_m and variance $\langle \delta\nu^2 \rangle_m$.

The surface is chosen to be randomly rough, planar on the average at $z=0$. The surface's height variation, $h(\mathbf{r})$, is governed by a spatially uniform, zero-mean Gaussian ran-

dom process, and the autocorrelation function, Γ. The assumption of spatial uniformity (translational invariance) implies that the autocorrelation function can be written in the form

$$\Gamma(\mathbf{r}-\mathbf{r}')\equiv\frac{\langle h(\mathbf{r})h(\mathbf{r}')\rangle_s}{h^2}, \tag{4}$$

where $h^2=\langle h^2(\mathbf{r})\rangle_s$. The angular brackets denote an average over an ensemble of rough surfaces. The autocorrelation function will be represented by a Gaussian function, $\Gamma=\exp(-|\mathbf{r}-\mathbf{r}'|^2/L_s^2)$. Here, L_s defines the surface autocorrelation length.

The treatment of wave propagation and the interaction of the waves with microstructure are all handled in much the same way as in Ref. 11. Here we provide only a cursory summary. We use the following wave equation (and its time-domain analog)

$$\nabla^2 u(\omega,\mathbf{r},z)+\frac{\omega^2}{c^2(\mathbf{x})}\,u(\omega,\mathbf{r},z)=0 \tag{5}$$

to model the propagation of sound in water and in the sample. Here u denotes the pressure wave field, and ω denotes the angular frequency. The boundary conditions are the continuity of the wave field and its normal derivative at the water–solid interface.

One effect of rough surfaces is to change the phase of the incident wave as it crosses the surface. We will model this effect using the phase-screen approximation as described in Refs. 8–12. The phase-screen approximation assumes that both the rms height and the wavelength are small compared to the surface autocorrelation length. The phase-screen approximation for the roughness-induced phase change is $\delta\phi=\delta k h(\mathbf{r})$, where δk denotes the change in wave number upon crossing the surface while $h(\mathbf{r})$ denotes the local surface height. The mean-square value of $\delta\phi$,

$$\sigma^2=\langle\delta\phi^2\rangle_s=\delta k^2 h^2, \tag{6}$$

is an important parameter that appears often in our formulas.

Transverse modes in the solid are ignored since the incident wave is at nearly normal incidence and consequently these modes are negligible. We make the following approximations (see Ref. 11) for the propagation of the beam: (1) the phase-screen approximation for transmission through the rough surface; (2) the Fresnel approximation; (3) the approximation of the incident wave as a Gaussian beam and as a toneburst; and (4) the scalar wave approximation. We make the following approximations for the interaction of the beam with the microstructure: (1) acoustic contrast is due to fluctuations in the compressibility; (2) ultrasonic attenuation and absorption are neglected; and (3) the interaction between the beam and the microstructure is weak (Born approximation).

II. ANALYTICAL SERIES SOLUTIONS

An approximate analytic series solution for the acoustic backscatter power will be derived in this section. The introduction of a Gaussian autocorrelation function to describe the microstructure is the major difference from Ref. 11. Consequently, we start with Eq. (21) of Ref. 11, modified to include the finite size of the microstructure. Upon making the various approximations described in the last section, we obtain the following equation for the acoustic backscatter power (normalized by the electrical power input to the transducer) from microstructure at depth z:

$$P(z)=|C(\omega)|^2|T_{\mathrm{WS}}|^2|T_{\mathrm{SW}}|^2\int d^2\mathbf{r}\times\left[\int d^3\mathbf{x}'\langle\delta\nu(\mathbf{x})\,\delta\nu(\mathbf{x}')\rangle_m\langle u_0^2(\omega,\mathbf{x})u_0^{2\,*}(\omega,\mathbf{x}')\rangle_s\right]. \tag{7}$$

Here, u_0 denotes the pressure field propagated through the rough surface but in the absence of microstructure; i.e., the sample is treated as a roughened but otherwise uniform half-space. T_{WS} and T_{SW} denote the transmission coefficients from the water to the solid and *vice versa*. $C(\omega)=i\omega CF/[4\rho c_0^2E(\omega)]$, where ρ denotes the density of the sample, $E(\omega)$ denotes the electrical power input to the transducer, and CF denotes conversion factor from electric power into acoustic power.

The analytic approximation of Eq. (7) proceeds as follows. We start by substituting Eq. (3) and u_0, defined by Eqs. (16) and (20) of Ref. 11, into Eq. (7). Next, we note that the beam can be described by a rapidly varying part (e^{ikz}) and a much more slowly varying envelope function. We will assume that this envelope function varies slowly with depth compared to the autocorrelation length of the microstructure. That is to say, $\langle u_0^2(\omega,\mathbf{x})u_0^{2\,*}(\omega,\mathbf{x}')\rangle_s$ depends weakly on z' compared to $\exp(-(z-z')^2/L_m^2)$. Therefore, the integral over the variable z' can be approximated separately; the result is described by the variable $K=\sqrt{\pi}|C(\omega)|^2\times|T_{\mathrm{WS}}|^2|T_{\mathrm{SW}}|^2\langle\delta\nu^2\rangle_m L_m$. The result is

$$\begin{aligned}P(z,h,L_s,R,F,L_m,\omega)=K\left(\frac{k_s}{2\pi z}\right)^4\int\int d^2\mathbf{r}\,d^2\mathbf{r}'\,\exp(-|\mathbf{r}-\mathbf{r}'|^2/L_m^2)\int\int\int\int d^2\mathbf{s}_1\,d^2\mathbf{s}_2\,d^2\mathbf{s}_3\,d^2\mathbf{s}_4\\ \times\exp\left[-\left(\frac{1}{R^2}+\frac{ik_s}{2F}\right)(s_1^2+s_2^2)-\left(\frac{1}{R^2}-\frac{ik_s}{2F}\right)(s_3^2+s_4^2)\right]\\ \times\exp\left(\frac{ik_s}{2z}\,(|\mathbf{s}_1-\mathbf{r}|^2+|\mathbf{s}_2-\mathbf{r}|^2-|\mathbf{s}_3-\mathbf{r}'|^2-|\mathbf{s}_4-\mathbf{r}'|^2)\right)\\ \times\langle\exp(i(\phi(\mathbf{s}_1)+\phi(\mathbf{s}_2)-\phi(\mathbf{s}_3)-\phi(\mathbf{s}_4)))\rangle_s.\end{aligned} \tag{8}$$

Here, R is the radius of the beam on the surface, while F is its focal length in the solid as measured from the surface. For a transducer immediately above the water–solid interface, R and F are the radius and focal length of the transducer. We first integrate over the variables $\mathbf{r}$ and $\mathbf{r}'$; the remaining integrals are evaluated following Ref. 11 as we now describe. Since the surface height distribution, $h(\mathbf{r})$, is described by a spatially uniform zero-mean Gaussian random process, we can write[11]

$$\langle \exp(i(\phi(\mathbf{s}_1)+\phi(\mathbf{s}_2)-\phi(\mathbf{s}_3)-\phi(\mathbf{s}_4)))\rangle_s = \exp(-\sigma^2 H(\phi(\mathbf{s}_1,\mathbf{s}_2,\mathbf{s}_3,\mathbf{s}_4))), \tag{9}$$

where H is defined by

$$H(\mathbf{s}_1,\mathbf{s}_2,\mathbf{s}_3,\mathbf{s}_4)=2+\Gamma(\mathbf{s}_1-\mathbf{s}_2)+\Gamma(\mathbf{s}_3-\mathbf{s}_4)-\Gamma(\mathbf{s}_1-\mathbf{s}_3)-\Gamma(\mathbf{s}_1-\mathbf{s}_4)-\Gamma(\mathbf{s}_2-\mathbf{s}_3)-\Gamma(\mathbf{s}_2-\mathbf{s}_4). \tag{10}$$

We make the coordinate transformation

$$\begin{aligned}\mathbf{s}_1&=(\mathbf{p}_1+\mathbf{p}_2+\mathbf{p}_3+\mathbf{p}_4)/2,\\ \mathbf{s}_2&=(-\mathbf{p}_1-\mathbf{p}_2+\mathbf{p}_3+\mathbf{p}_4)/2,\\ \mathbf{s}_3&=(\mathbf{p}_1-\mathbf{p}_2-\mathbf{p}_3+\mathbf{p}_4)/2,\\ \mathbf{s}_4&=(-\mathbf{p}_1+\mathbf{p}_2-\mathbf{p}_3+\mathbf{p}_4)/2.\end{aligned} \tag{11}$$

In the new coordinates, the integrand becomes independent of $\mathbf{p}_4$

$$H(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4)=2+\Gamma(\mathbf{p}_1+\mathbf{p}_2)+\Gamma(\mathbf{p}_1-\mathbf{p}_2)-\Gamma(\mathbf{p}_2+\mathbf{p}_3)-\Gamma(\mathbf{p}_1+\mathbf{p}_3)-\Gamma(\mathbf{p}_1-\mathbf{p}_3)-\Gamma(-\mathbf{p}_2+\mathbf{p}_3), \tag{12}$$

and the integral over that coordinate is completed trivially. Next, we obtain the series solution by: (1) expanding the exponential $\exp(-\sigma^2 H(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4))$ in a Taylor series; (2) changing the order of the resulting summations and integrations; and finally (3) evaluating the resulting integrals. After evaluating the integrals over $\mathbf{p}_3$, $\mathbf{p}_2$, $\mathbf{p}_1$, the backscattered noise power is obtained in series form

$$P(z,h,L_s,R,F,L_m,\omega) = \frac{K}{8\pi}\left(\frac{k_s}{2z}\right)^3 e^{-2\sigma^2} \times \sum_{\substack{a,b,c,\\ d,e,f=0}}^{\infty}\left(\frac{(-\sigma^2)^{a+b}(\sigma^2)^{c+d+e+f}}{a!b!c!d!e!f!}\frac{1}{c_2c_3c_5}\right), \tag{13}$$

where

$$c_1=\frac{1}{R^2}+\frac{1}{L_m^2}+\frac{k_s^2R^2}{4F^2}, \tag{14a}$$

$$c_2=c_1+\frac{c+d+e+f}{L_s^2}, \tag{14b}$$

$$c_3=\frac{1}{R^2}+\frac{a+b+c+f}{L_s^2}-\frac{(f-c)^2}{L_s^4c_2}, \tag{14c}$$

$$c_4=\frac{ik_s}{2}\left(\frac{1}{z}-\frac{1}{F}\right)+\frac{(b-a)}{L_s^2}+\frac{(e-d)(f-c)}{L_s^4c_2}, \tag{14d}$$

$$c_5=\frac{1}{R^2}+\frac{a+b+d+e}{L_s^2}-\frac{(e-d)^2}{L_s^4c_2}-\frac{c_4^2}{c_3}. \tag{14e}$$

Our purpose is to determine the changes in the backscatter power due to surface roughness. In order to isolate surface induced changes, we divide $P(z)$ by $P_0(z)$ to obtain the normalized backscatter power $Q(z)=P(z)/P_0(z)$. Here, $P_0(z)$ is the backscatter power for a smooth surface and is obtained by setting $h=0$ ($\sigma=0$) in Eq. (13).

III. ROUGHNESS-INDUCED CHANGES IN ACOUSTIC BACKSCATTER

We report roughness-induced changes in the normalized mean backscatter power, $Q=P/P_0$, as a function of the transducer radius, surface statistics, and the microstructural autocorrelation length. The predictions are illustrated with graphs calculated, unless otherwise noted, for the following parameters: transducer radius $R=10$ mm, frequency $f=10$ MHz, rms surface height $h=0.02$ mm, $\delta v=5\%$, surface autocorrelation length $L_s=1$ mm, microstructural autocorrelation length $L_m=0.1$ mm, speed of sound in water, $c_w=1500$ m/s, and the speed of longitudinal sound in a solid (aluminum), $c_s=6300$ m/s. Both focused, $F=40$ mm, and unfocused transducers are considered. For a transducer immediately above the water–solid interface, R denotes the radius of the transducer and F denotes the focal length of the transducer in the solid. The analytic series solutions defined by Eqs. (13) and (14) are found to converge relatively rapidly for the parameters mentioned above; typically we sum each index from 1 to 6. The resulting calculations are plotted as a function of depth.

One may expect qualitatively different results depending on the relative size of the microstructure's and the surface's autocorrelation lengths. Figure 2 shows the variation of Q with the microstructural autocorrelation length L_m. For unfocused probes Q is shown in Fig. 2(a) ($L_m=0.1$ mm, 1 mm, and 10 mm and $L_s=1$ mm). The normalized backscatter power at all depths depends inversely on L_m, that is, Q is largest when L_m is small and smallest when L_m is large. The calculated normalized backscatter increases up to a depth of approximately 8 mm. For greater depth the curves either decrease or remain nearly constant. All of the curves have reasonably similar values at $z=80$ mm, the greatest depth reported and appear to be approaching each other.

For the smallest microstructural autocorrelation length $L_m=0.1$ mm, there is almost no change in the normalized backscatter at $z=0$. For somewhat larger z, the normalized backscatter increases with depth, reaches a maximum that is greater than 1, and then decreases with depth. This curve is very similar to the results of Ref. 11, as should be expected since $L_m \ll L_s$ and λ (the wavelength, which is 0.63 mm), and the results in Ref. 11 were obtained under the assumption $L_m \ll L_s$, λ. The second curve is for $L_m=1$ mm, which is equal to the surface autocorrelation length and comparable to the wavelength. In this case the normalized backscattered power is approximately 0.50 for scatterers in the immediate

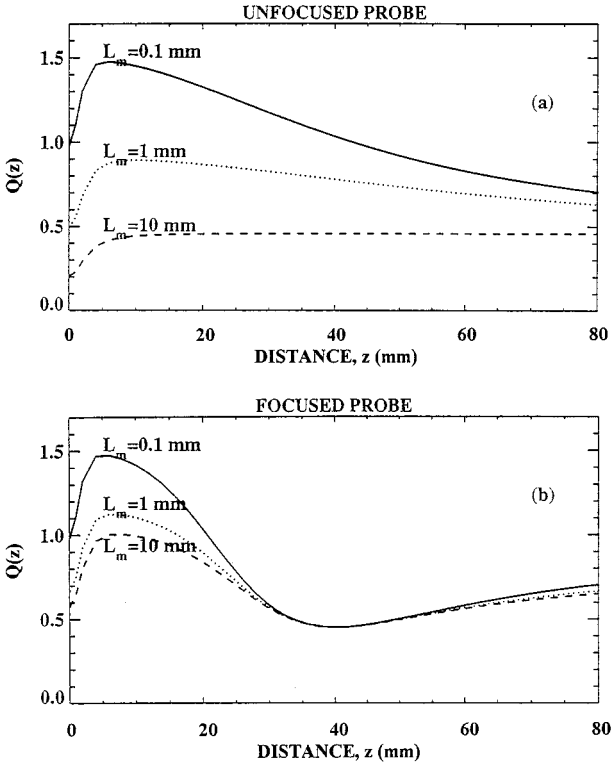


FIG. 2. Comparison of the normalized acoustic backscatter power for three different microstructure autocorrelation lengths: solid line—$L_m=0.1$ mm; dotted line—1.0 mm; dashed line—10 mm. The other parameters were $R=10$ mm, $L_s=1$ mm, $h=20$ μm, $f=10$ MHz, and $F=40$ mm.

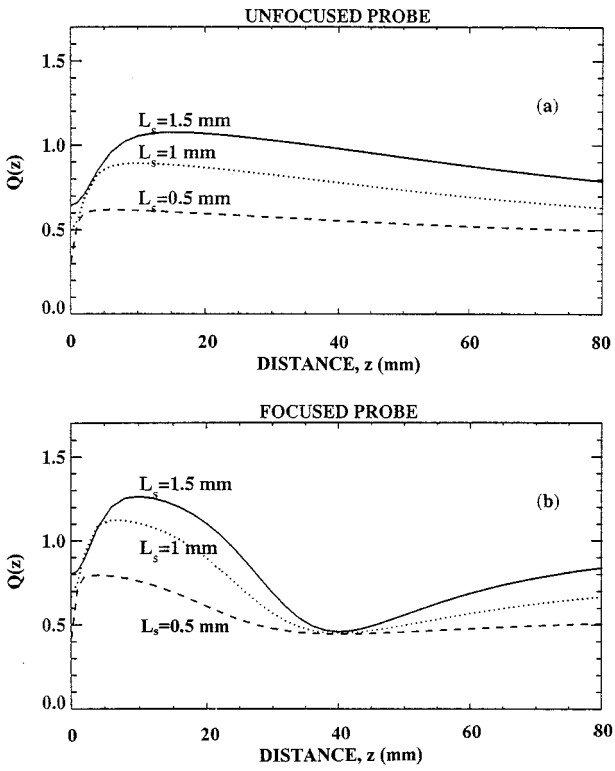


FIG. 3. Comparison of the normalized mean acoustic backscatter power for three different surface autocorrelation lengths: solid line—$L_s=1.5$ mm; dotted line—1.0 mm; dashed line—0.5 mm. The other parameters were $R=10$ mm, $L_m=1.0$ mm, $h=20$ μm, $f=10$ MHz, and $F=40$ mm.

vicinity of the surface. The curve rises to a maximum of approximately 0.90 at a depth of roughly 8 mm. The curve then falls monotonically for scatterers at greater depth. Finally, we come to the case that $L_m=10$ mm$\gg L_s$ and λ. In this case, the curve rises monotonically from a value of 0.2 at $z=0$ and reaches a plateau value of roughly 0.45 for depths greater than 10 mm.

For focused probes, the normalized acoustic backscatter power depends on L_m as shown in Fig. 2(b). All of the curves initially increase with increasing depth, reach a maximum at roughly 5 mm, and then decrease to a common minimum of approximately 0.5 at the focal depth of 40 mm. For greater depth, the curves are very nearly the same and increase slightly with z. The enhanced backscatter at $z=5$ mm is greatest for $L_m=0.1$ mm and least for $L_m=10$ mm. The minimum value at the focal depth is roughly given by $e^{-2\sigma^2}$, where $\sigma^2=\delta k^2 h^2$ is a mean-square measure of the roughness induced phase shift. This dependence on frequency (and σ) can be argued theoretically (see Sec. IV) and is based on a large number of calculations not presented graphically in this paper. Importantly, the normalized acoustic backscatter power is nearly independent of L_m for scatterers at the focal depth. As we will see in the following paragraphs, at $z=F$ it is also nearly independent of L_s and the transducer's radius R.

The dependence of the normalized acoustic backscatter power on the surface autocorrelation length L_s is reported in Fig. 3 for a relatively large microstructural autocorrelation length of 1 mm, which is comparable to the wavelength. The normalized acoustic backscatter power is reported for $L_s=0.5$, 1.0, and 1.5 mm for unfocused transducers in Fig. 3(a). The normalized backscatter power at all depths depends directly on L_s, that is, Q is large when L_s is large and small when L_s is small. In each case, the curve starts off with a relatively small value at $z=0$, rises to a maximum, and then decreases monotonically. The depth at which the maximum occurs scales as $\sim k_s L_s^2$, where k_s denotes the wave number in the solid. The normalized acoustic backscatter power for a focused probe is shown in Fig. 3(b). Each curve shares a common minimum value at the focal depth. The value at the minimum is once again approximately $e^{-2\sigma^2}$ and is nearly independent of the surface autocorrelation length.

Finally, we examined the dependence of Q on the radius of the transducer R for $L_m=1$ mm and $L_s=1.5$ mm. The results for the unfocused probe are shown in Fig. 4(a). For each radius, $R=10$, 20, and 40 mm, the curves initially increase with increasing depth, reaching a maximum at a depth of roughly 8 mm. The curves then decrease monotonically with increasing depth. The curves are common for small z. The curves became distinct for depths greater than the maximum at $z=10$ mm. The backscatter is greater for larger R at all depths. Figure 4(b) shows Q for a focused probe. The three curves have a common minimum at approximately the focal depth. The half-width of the curve about the minimum depends on the radius of transducer. The half-width is smallest for the transducer with $R=40$ mm and largest for $R=10$ mm. Finally, the common value of the curves at the minimum is approximately $e^{-2\sigma^2}$.

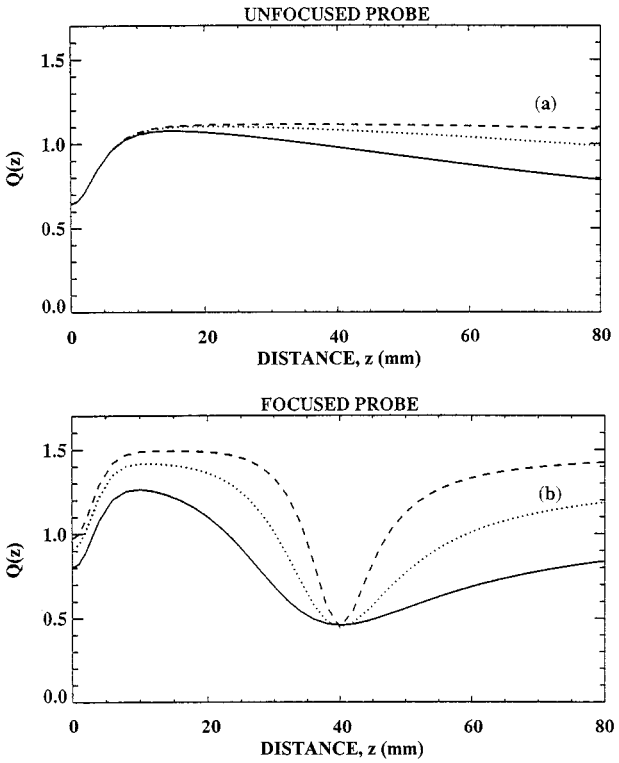


FIG. 4. Comparison of the normalized mean acoustic backscatter power for three different probe radii: solid line—$R=10$ mm; dotted line—20 mm; dashed line—40 mm. The other parameters were $L_s=1.5$ mm, $L_m=1$ mm, $h=20$ μm, $f=10$ MHz, and $F=40$ mm.

IV. DISCUSSION

In the following, we discuss the effects of surface roughness on the acoustic backscatter power as calculated from approximate analytic solutions for backscattering from a material with microstructure of finite size. The results showed certain common characteristics, such as an initial increase with depth and a peak (or plateau) at a relative depth that scales as $\sim k_s L_s^2$. For focused probes, there was a strong minimum at the focal depth and all curves showed a common minimum value of approximately $e^{-2\sigma^2}$. Qualitative explanations for many of these features were provided in Ref. 11 where we made the restrictive assumption that $L_m \ll L_s$ and λ. Here, we will discuss those features in the curves that arise from removing this restriction on the microstructural autocorrelation length.

The acoustic backscatter power, $P(z)$, is a rapidly varying function of the microstructural size; for example, $P(z)$ increases rapidly with L_m as long as $k_s L_m < 1$. We have presented results for the *normalized* backscattered power $Q(z)=P(z)/P_0$ in order to remove this strong size dependence and to isolate the effects of surface roughness. There are two striking results. First, the normalized backscattered power decreases for larger microstructures. This is probably the most important general consideration for unfocused probes. Second, for focused probes and for moderate amounts of surface roughness, the predicted normalized backscatter from the focal depth is nearly independent of the microstructural size, the surface autocorrelation length, and the transducer radius.

The decrease in the normalized acoustic backscatter power with increasing L_m can be qualitatively understood by supposing that the microstructure is made up of many independent scatterers. In the Born approximation, the signal from one of the independent scatterers is given by[11]

$$S = \text{const}\ \omega^2 \int d^3\mathbf{y}\ \nu(\mathbf{y}) u_0^2(\omega, \mathbf{y}). \tag{15}$$

Here, const is an unimportant constant of proportionality, while u_0 denotes the wave field that has passed through the rough surface but in the absence of the scatterer. The rough surface introduces fluctuations in phase and amplitude of the wavefield. These fluctuations are integrated over the volume of the scatterer according to Eq. (15), and thus tend to average out and reduce the signal. The larger the flaw, the larger the region of integration and the more completely the fluctuations average out. The prediction that the normalized backscattered signal decreases with increasing microstructure size is due to this *intrinsic* spatial averaging.

The results for the focused probe seem very odd at first blush. Inspection of the graphical results (and a large number of calculations at other frequencies and rms roughness that are not presented here) indicate that the average power backscattered from the focal depth depends primarily on $\sigma^2 = \delta k^2 h^2$, where δk is the change in wave number upon crossing the water–solid interface and h is the rms height of the surface roughness. In particular, the power backscattered from the focal depth does not seem to depend on L_m, L_s, or R. Why? Crudely, the rough surface divides the beam into two parts: an average (specular) beam that focuses and a random (diffuse) part that does not. At the focal depth, the specular part of the beam is concentrated in a small volume, while the diffuse part of the beam is widely dispersed. Hence, for moderate amounts of roughness, the signal from the focal depth is dominated by the concentrated specular part of the beam; the diffuse part contributes a negligible amount to the signal. The strength of the specular signal is given by $e^{-\sigma^2}$. Hence, the backscattered power at the focal depth is given by $e^{-2\sigma^2}$ for small and moderate degrees of roughness. Clearly, if the surface is very rough the amount of energy in the specular beam will be extremely small and our prediction breaks down; focusing has little effect for sufficiently rough surfaces.

ACKNOWLEDGMENTS

This material is based in part on work performed at the Center for NDE at Iowa State University as part of the Engine Titanium Consortium operated by the Center for NDE at Iowa State University including participants at General Electric, Pratt & Whitney, and Allied Signal, and supported by the Federal Aviation Administration Technical Administration Center, Atlantic City, New Jersey, under Grant No. 94G048.

APPENDIX

The same formalism that determined P and P_0 (which involve $|S|^2$, where S is the backscattered signal) allows us to compute the mean square of the backscattered signal, $\langle S^2 \rangle$.

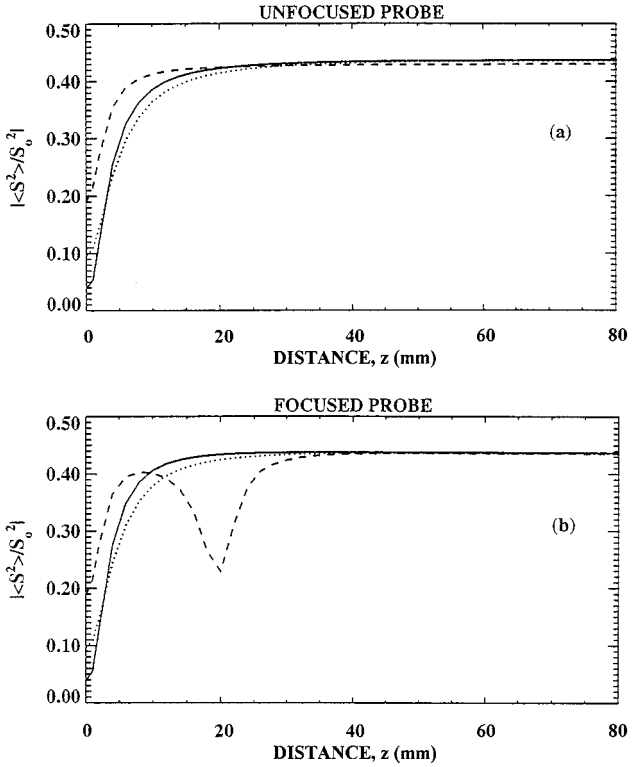


FIG. A1. Comparison of the magnitude of normalized mean-square signal for three different microstructure autocorrelation lengths: solid line—L_m=0.1 mm; dotted line—1.0 mm; dashed line—10 mm. The other parameters were R=10 mm, L_s=1 mm, h=20 μm, f=10 MHz, and F=40 mm.

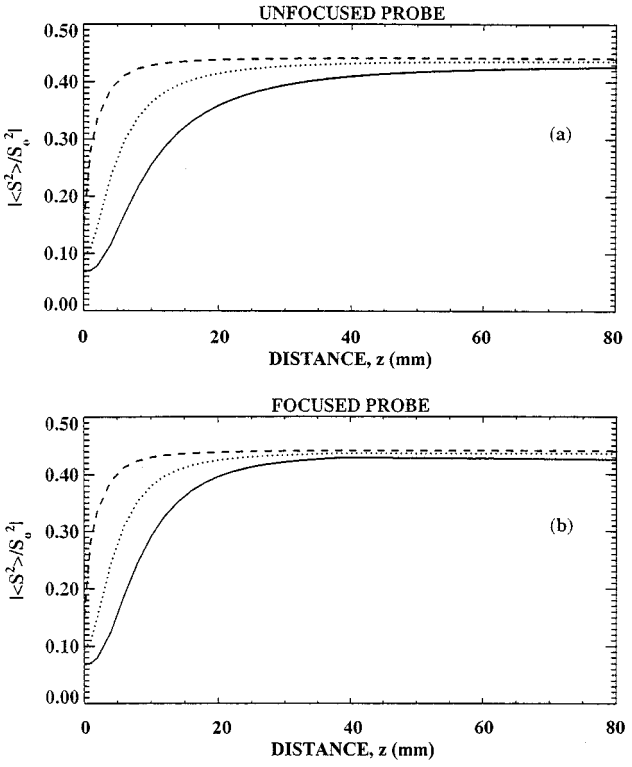


FIG. A2. Comparison of the magnitude of normalized mean-square signal for three different surface autocorrelation lengths: solid line—L_s=1.5 mm; dotted line—1.0 mm; dashed line—0.5 mm. The other parameters were R=10 mm, L_m=1.0 mm, h=20 μm, f=10 MHz, and F=40 mm.

In this Appendix, we derive a series solution for the mean square of the backscattered noise, which depends on the ensemble average

$$\langle \exp(i(\phi(\mathbf{s}_1)+\phi(\mathbf{s}_2)+\phi(\mathbf{s}_3)+\phi(\mathbf{s}_4)))\rangle = \exp(-\sigma^2 G(\phi(\mathbf{s}_1,\mathbf{s}_2,\mathbf{s}_3,\mathbf{s}_4)), \tag{A1}$$

where G is

$$G(\mathbf{s}_1,\mathbf{s}_2,\mathbf{s}_3,\mathbf{s}_4)=2+\Gamma(\mathbf{s}_1-\mathbf{s}_2)+\Gamma(\mathbf{s}_3-\mathbf{s}_4)+\Gamma(\mathbf{s}_1-\mathbf{s}_3)+\Gamma(\mathbf{s}_1-\mathbf{s}_4)+\Gamma(\mathbf{s}_2-\mathbf{s}_3)+\Gamma(\mathbf{s}_2-\mathbf{s}_4). \tag{A2}$$

The mean-square signal can be found in series form, following a procedure similar to that of Sec. II

$$\langle S^2(z,h,L_s,R,F,L_m,\omega)\rangle = \frac{-iK(k_s/2z)^3 e^{i2k_s z}}{4q_2(1/L_m^2+ik_s/2z)}\, e^{-2\sigma^2} \times \sum_{\substack{a,b,c,\\ d,e,f=0}}^{\infty} \left(\frac{(-\sigma^2)^{a+b+c+d+e+f}}{a!b!c!d!e!f!} \frac{1}{q_3 q_4 q_6} \right), \tag{A3}$$

where

$$q_1=\frac{1}{R^2}+i\,\frac{k_s}{2}\left(\frac{1}{F}-\frac{1}{z}\right), \tag{A4a}$$

$$q_2=\frac{1}{R^2}+\frac{ik_s}{2F}, \tag{A4b}$$

$$q_3=q_1+\frac{a+b+c+f}{L_s^2}, \tag{A4c}$$

$$q_4=q_1+\frac{a+b+d+e}{L_s^2}-\frac{(b-a)^2}{L_s^4 q_3}, \tag{A4d}$$

$$q_5=\frac{(e-d)}{L_s^2}+\frac{(b-a)(f-c)}{L_s^4 q_3}, \tag{A4e}$$

$$q_6=q_1+\frac{(k_s/2z)^2}{1/L_m^2+ik_s/2z}+\frac{c+d+e+f}{L_s^2}-\frac{(f-c)^2}{L_s^4 q_3}-\frac{q_5^2}{q_4}. \tag{A4f}$$

The mean-square signal in the absence of surface roughness is denoted by S_0^2 and can be found by setting σ=0 in Eqs. (A3)–(A4). The present paper concerns backscatter from microstructure. However, somewhat similar results were found for the backscatter of ultrasound from a single discrete scatterer beneath a random rough surface as discussed in Ref. 12.

The roughness-induced effects on the mean-square signal are characterized through the normalized mean-square signal $\langle S^2\rangle/S_0^2$. The magnitude of the normalized mean-square signal versus depth was calculated as a function of the microstructural autocorrelation length, the surface autocorrelation length and the probe radius.

Figure A1 shows the dependence of the normalized mean-square signal on the microstructural autocorrelation

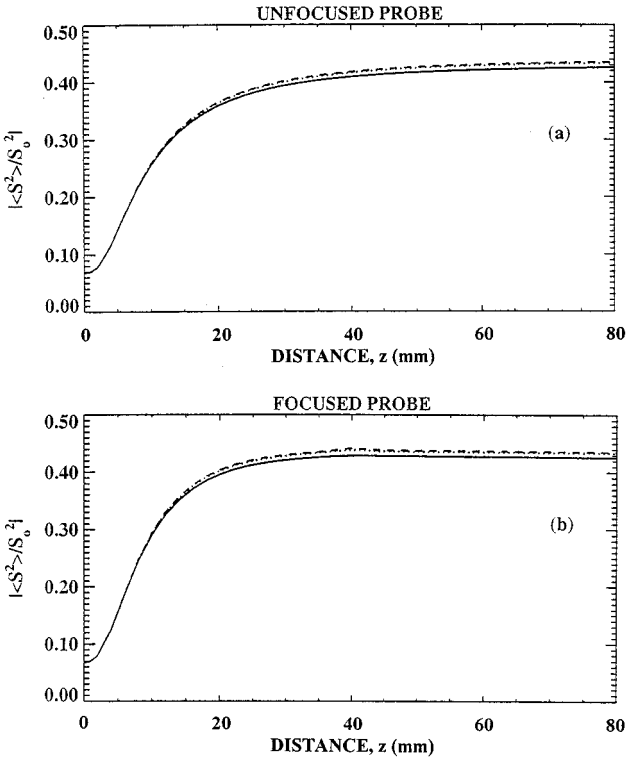


FIG. A3. Comparison of the magnitude of normalized mean-square signal for three different probe radii: solid line—$R=10$ mm; dotted line—20 mm; dashed line—40 mm. The other parameters were $L_s=1.5$ mm, $L_m=0.1$ mm, $h=20$ μm, $f=10$ MHz, and $F=40$ mm.

length. The specific values used to generate the figure are $L_m=0.1$, 1.0, and 10 mm and $L_s=1$ mm. For an unfocused transducer, $|\langle S^2\rangle/S_0^2|$ is small at $z=0$, it then rises to a common plateau of value roughly 0.45 over a characteristic distance. $|\langle S^2\rangle/S_0^2|$ is 0.04 at the surface ($z=0$) for the microstructures with the smallest L_m, while it is 0.19 for the microstructures having largest L_m. For the focused transducer, the results are shown in Fig. A1(b). The results are similar to those for the unfocused transducer for L_m equal to 0.1 or 1.0 mm. However, the results for the focused probe are quite different for $L_m=10$ mm. In this case, the normalized mean-square signal exhibits a minimum at one-half the focal length. A similar phenomena was predicted for the signal from a single defect beneath a rough surface in Ref. 12. The observed minimum occurs when $z=F/2$ and is due to highly correlated phase shifts that occur when the transmitted and reflected waves of the pulse-echo configuration focus on the same part of the rough surface.

Figure A2 shows the dependence of the normalized mean-square signal on the surface autocorrelation length L_s for $L_s=0.5$, 1.0, and 1.5 mm and $L_m=1$ mm. $|\langle S^2\rangle/S_0^2|$ is least at the surface and increases with increased depth and appears to reach a common limit of 0.45. The initial values depend strongly on the surface autocorrelation length and is smaller for larger L_s. On the other hand, for larger L_s the normalized mean-square signal increases more slowly with depth. Comparison of the results in Fig. A2(a) and (b) further indicate that $|\langle S^2\rangle/S_0^2|$ depends weakly on focusing.

The dependence of $|\langle S^2\rangle/S_0^2|$ on the transducer's radius is depicted in Fig. A3 for $R=10$, 20, and 40 mm, and for $L_m=1$ mm and $L_s=1.5$ mm. The calculations interestingly predict that the normalized mean-square signal also depends weakly on radius.

The results for the normalized mean-square signal presented in Figs. A1–A3 share certain characteristics, such as an initial increase with depth and a common plateau at increased depth. For both the focused and unfocused transducers, the normalized mean-square signal at $z=0$ is $e^{-8\sigma^2}$ for the smallest L_m and $e^{-4\sigma^2}$ for largest L_m. The transition from the value at $z=0$ to the common plateau of $e^{-2\sigma^2}$ occurs over a characteristic distance proportional to $\sim k_s L_s^2$. Thus, the larger the surface autocorrelation length the larger the transition region. The change in the size of the transducer has little effect on the normalized mean-square signal.

Finally, comparison of the normalized mean-square signal $|\langle S^2\rangle/S_0^2|$ presented in Figs. A1–A3 with the normalized power $Q(z)$ at small depths indicates an interesting fact. $|\langle S^2\rangle/S_0^2|$ increases with increase in L_m and decrease in L_s while $Q(z)$ increases with decrease in L_m and increase in L_s.

[1] J. A. Ogilvy, *Theory of Wave Scattering from Random Rough Surfaces* (Adam Hilger, Bristol, 1991).

[2] C. S. Clay and H. Medwin, *Acoustical Oceanography* (Wiley–Interscience, New York, 1977).

[3] *Ultrasonic Scattering in Biological Tissues*, edited by K. K. Shung and G. A. Thieme (CRC, Ann Arbor, 1993).

[4] F. J. Margetan, R. B. Thompson, and I. Yalda-Mooshabad, ''Backscattered microstructural noise in ultrasonic toneburst inspections,'' J. Nondestruct. Eval. **13**, 111–136 (1990).

[5] F. J. Margetan, K. Y. Han, and I. Yalda-Mooshabad, ''The practical application of grain noise models in titanium billets and forgings,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1995), Vol. 14B, pp. 2129–2136.

[6] J. H. Rose, ''Theory of ultrasonic backscatter from multiphase polycrystalline solids,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1993), Vol. 11B, pp. 1677–1684.

[7] J. R. Brewster and K. H. G. Ashbee, ''The application of wave field dislocation statistics to characterization techniques involving scattered, pulsed ultrasound,'' J. Phys. D **28**, 1811–1815 (1995).

[8] P. B. Nagy and L. Adler, ''Surface roughness induced attenuation of reflected and transmitted ultrasonic waves,'' J. Acoust. Soc. Am. **82**, 193–197 (1987).

[9] P. B. Nagy and J. H. Rose, ''Surface roughness and the ultrasonic detection of subsurface scatterers,'' J. Appl. Phys. **73**, 566–580 (1993).

[10] M. Bilgen and J. H. Rose, ''Effects of one-dimensional random rough surfaces on acoustic backscatter: Utility of phase-screen and Fresnel approximations,'' J. Acoust. Soc. Am. **96**, 2849–2855 (1994).

[11] M. Bilgen and J. H. Rose, ''Rough surface effects on incoherent scattering from random volumetric scatterers: Approximate analytic series solution,'' J. Acoust. Soc. Am. **96**, 2856–2869 (1994).

[12] M. Bilgen and J. H. Rose, ''Mean and variance of the ultrasonic signal from a scatterer beneath a rough surface,'' J. Acoust. Soc. Am. **98**, 2217–2225 (1995).

Acoustic signal-to-noise ratio for an inclusion beneath a randomly rough surface and in the presence of microstructure

Mehmet Bilgen
Department of Radiology, University of Kansas Medical Center, Kansas City, Kansas 66160

James H. Rose
Center for NDE, Iowa State University, Ames, Iowa 50011

(Received 18 March 1996; accepted for publication 19 July 1996)

The effects of surface roughness on the signal-to-noise ratio (SNR) for the phase-sensitive detection of an inclusion in a sample with microstructure is explored in a simple model. It is assumed that a phase-sensitive transducer is used in pulse-echo mode, that the sample is immersed in a water bath, and that the noise is dominated by scattering from the sample's microstructure. Since there is interest in roughness-induced *changes* in the SNR, calculations are reported of the normalized signal-to-noise ratio (NSNR), i.e., the SNR with the rough surface divided by the SNR with the smooth surface. The dependence of the NSNR on transducer parameters (radius and focal length), on surface parameters (root-mean-square height and autocorrelation length), and on material parameters (scattering strength, average crystalline size) is reported. In this paper, approximate analytic series solutions, previously obtained for the signal and the noise, are combined to estimate the NSNR as a function of the inclusion's depth beneath the surface. It is found that for unfocused transducers even moderately rough surfaces can substantially degrade the NSNR for the most typical case, a relatively large inclusion in a background of small microcrystallites. The calculations predict two other results that initially surprised us. First, surface roughness is predicted to *increase* the SNR for a relatively small strongly scattering inclusion in a sample with relatively large but weakly scattering microstructure. Second, moderate roughness is predicted to have little effect on the SNR for focused transducers at the focal depth. [S0001-4966(97)00301-9]

PACS numbers: 43.35.Zc [HEB]

INTRODUCTION

Ultrasound is widely used to detect internal flaws in parts. Typical inspections use pulse-echo phase-sensitive piezoelectric transducers with both transducer and part immersed in a water bath. The ultrasonic signal and the consequent detection and characterization of internal flaws depends on the surface finish of the part. The change in the signal-to-noise ratio (SNR) due to surface roughness is thus a practical concern, as well as a fundamental scientific problem.

A series of papers have documented the effects of surface roughness on: (1) the ultrasonic signal from an inclusion beneath a rough surface;[1,2] and (2) the changes in the acoustic backscatter signal generated by scattering from the part's microstructure.[2–5] Approximate analytic and readily evaluated series solutions for both the signal and the acoustic backscatter power were presented.[1,4,5] In the present context, we will consider the acoustic backscatter generated by the microstructure to be a source of noise that interferes with the detection of the signal. Henceforth, we refer to the acoustic backscatter as material noise. In this paper, we combine the analytical solutions for both the signal[1] and material noise[4,5] to provide a simple approximate model for the normalized signal-to-noise (NSNR) (i.e., the SNR with the rough surface divided by the SNR with the smooth surface).

The class of problems that can be simulated using our approach is large. In this paper, graphical results will be presented for a restricted set of problems of interest in nondestructive evaluation. Pulse-echo geometry is assumed with the transducer nearly touching the surface. The surface autocorrelation length is assumed to be small compared to the radius of the transducer, while the size of the inclusion and the microstructure are assumed to be smaller than or comparable to the transducer's size. We also assume that roughness is not too great and that consequently losses are not greater than 10 dB or so. Most of these restrictions are for the purpose of discussion and can be removed.

The most typical effect of surface roughness is to degrade the SNR since the signal from a large coherent scatterer is reduced more than the material noise.[2] Our analysis yielded two results that were surprising to the authors. First, surface roughness is predicted to *improve* the SNR (i.e., NSNR$>$1) if we have a small inclusion and the material noise is generated by relatively large and weakly scattering microstructure. Second, the calculations predict that for moderate roughness the deleterious effects of surface roughness on the SNR can be nearly removed at the focal depth for an appropriately focused transducer (i.e., NSNR$\approx$1).

The structure of the paper is as follows. In Sec. I we describe the problem and define the SNR and NSNR. In Sec. II we introduce the various simplifying approximations. We illustrate the results graphically in Sec. III. Section IV is a heuristic discussion of the results and a physical explanation of our two most important finding. Finally, we conclude the paper with a summary.

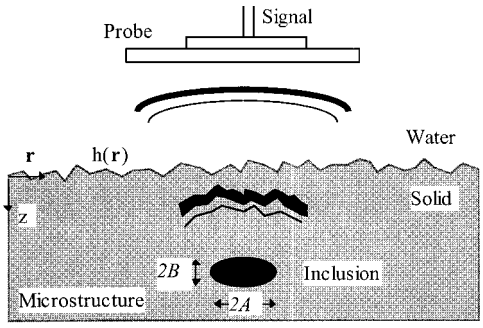


FIG. 1. Schematic geometry for model determination of the SNR. Shows the detection of a subsurface inclusion below a rough surface and embedded in an inhomogeneous background. The microstructure is ignored in the calculation of the inclusion's signal and the inclusion is ignored in the calculation of the material noise.

I. SIGNAL-TO-NOISE RATIO

The schematic geometry of the problem is shown in Fig. 1. We model the NSNR that results from the reflection of an ultrasonic beam by an inclusion in a thick solid plate that has both microstructure and a rough surface, and that is immersed in a water bath. The plate is modeled by an infinite half-space with a rough water–solid interface. The microstructure is described by random variations in the velocity.

In previous papers, we assumed that the interaction of the beam with both the inclusion and the microstructure could be described by the Born approximation, a weak scattering theory.[1,4] Basically, this amounts to ignoring interaction effects. In this paper, we continue to ignore interactions and assume that the microstructure does not change the signal generated by the inclusion, and that the inclusion does not change the material noise. That is, we calculate the signal, S, and the signal power $\langle|S|^2\rangle_s$ for an inclusion in a roughened uniform half-space without microstructure. The brackets $\langle\cdots\rangle_s$ denote an average over an ensemble of rough surfaces. We also calculate the average material noise power $\langle\langle|N|^2\rangle_s\rangle_m$ noise for an inclusion-free roughened half-space with microstructure. The angular brackets $\langle\cdots\rangle_m$ indicate an average over an ensemble of microstructures. The SNR ratio will be defined as the ratio of signal power divided by noise power

$$\mathrm{SNR}\equiv\frac{\langle|S|^2\rangle_s}{\langle\langle|N|^2\rangle_s\rangle_m}. \tag{1}$$

Our approach is suitable if the inclusion's dimension are much smaller than the beam size or if the microstructure of the inclusion is nearly the same as the microstructure of the host material.

We are interested in the question ''How does surface roughness change the SNR?'' We will quantify the change by considering the NSNR, which is defined by

$$\mathrm{NSNR}\equiv=\frac{\langle|S|^2\rangle_s/\langle\langle|N|^2\rangle_s\rangle_m}{|S_0|^2/\langle|N_0|^2\rangle_m}. \tag{2}$$

Here, S_0 and N_0 denote the signal (from an inclusion) and the material noise for a sample with a smooth surface.

We define the SNR as the ratio of the mean signal power $\langle|S|^2\rangle_s$ and the material noise power. Other choices are possible. For example, we could have plausibly chosen to define the SNR and NSNR in terms of the mean signal $\langle S\rangle_s$. This seems to be a reasonable choice and, at first glance, perhaps the preferred choice. However, there is a severe problem with defining the SNR in terms of $\langle S\rangle_s$. Namely, the surface roughness has two effects for a particular signal, S. First, it can reduce the amplitude of the signal, and second, it can introduce a random phase shift. There are common circumstances in which the amplitude of the signal is more or less unaffected by the rough surface but a large random phase shift is introduced. In this case, the experimentalist would have no trouble detecting the signal and would only lose information on its absolute phase. For typical ultrasonic inspections, the detection of the scatterer depends almost entirely on the signal's magnitude and not on its phase. Hence, we have chosen the signal power $\langle|S|^2\rangle$ as our average measure of the signal in defining the SNR and NSNR.

The following is an extreme example that exhibits the point made in the last paragraph. Suppose we have a hypothetical ensemble of samples, each of which has a surface that is flat but whose distance from the transducer is uncertain by a random function, δz. This is the same as an ensemble of rough surfaces with the surface autocorrelation length set equal to infinity. Since δz is small, on the order of a wavelength, the amplitude of the signal seen by the experimentalist is almost the same for each member of the ensemble. However, the signal's phase can be altered by large amounts since $\delta k\,\delta z\gg 1$, where δk is the change in wave vector upon crossing the surface. The result is that $\langle|S|^2\rangle_s$ is nearly unchanged, while the mean signal, $\langle S\rangle_s = \exp(-\delta k^2\langle|\delta z|^2\rangle)$, can be much less than one. Hence, $\langle|S|^2\rangle_s$ is a good indicator of the detectability of the scatterer, while $\langle S\rangle_s$ is not.

II. THEORETICAL MODEL AND APPROXIMATIONS

The signal and the noise will be determined from previously published approximate analytic series solutions.[1,4,5] Here, we briefly summarize the approximations for the rough surface, the microstructure, and the inclusion that were made to obtain these series solutions. Our approach approximates the propagation of longitudinal pressure waves in both the fluid and the solid by the scalar wave equation

$$\nabla^2 u(\omega,\mathbf{r},z)+\frac{\omega^2}{c^2(\mathbf{x})}\,u(\omega,\mathbf{r},z)=0. \tag{3}$$

Here, $c(\mathbf{x})$ denotes the velocity and ω denotes the angular frequency, while u denotes the pressure wave field. A general point in space is denoted by the vector $\mathbf{x}$, while z denotes the coordinate perpendicular to the surface and $\mathbf{r}$ the coordinate parallel to the surface, $\mathbf{x}=(\mathbf{r},z)$. We rewrite Eq. (3) in a form more suitable for scattering theory:

$$\nabla^2 u(\omega,\mathbf{r},z)+\frac{\omega^2}{c_0^2}\,u(\omega,\mathbf{r},z)-\frac{\omega^2}{c_0^2}\,\delta\nu(\mathbf{r})u(\omega,\mathbf{r},z)=0. \tag{4}$$

Here, c_0 denotes average velocity, defined more precisely below, while

$$\delta\nu(\mathbf{x}) \equiv 1 - \frac{c_0^2}{c^2(\mathbf{x})} \tag{5}$$

denotes the velocity deviation. The first two terms in Eq. (4) describe propagation in a homogeneous medium. The third term includes the effects of deviations of the velocity from the average; it is this term that leads to acoustic scattering.

The presence of a random surface requires that we treat the wave field statistically. The wave field, for a given realization of the surface, can be written as

$$u(\omega,\mathbf{r},z) \equiv \langle u(\omega,\mathbf{r},z)\rangle_s + \delta u(\omega,\mathbf{r},z). \tag{6}$$

Often, the mean field $\langle u\rangle_s$ is called the coherent field and the perturbation δu is called the incoherent field in the literature (and in some of our papers). We will avoid this terminology in this paper and call $\langle u\rangle_s$ the specular part of the field, while we will call δu the diffuse part of the field.

A. Microstructure model

The acoustic noise was assumed to arise from weak microstructural variations in the material. The material's microstructure (absent the inclusion) is assumed to be homogeneous and isotropic on the ensemble average. As noted above, the noise is computed in the absence of the inclusion. We will assume, as we have previously, that the backscattered noise results from weak spatial variations in the sound velocity and that the scattering is calculated in the Born approximation. Deviations of the velocity are calculated with respect to the average velocity c_0 that is defined by

$$\frac{1}{c_0^2} \equiv \left\langle \frac{1}{c^2(\mathbf{x})} \right\rangle_m. \tag{7}$$

Here $c(\mathbf{x})$ is the velocity distribution due to the microstructure. The effects of microstructure on the scattering is given by $\delta\nu(\mathbf{x})$, as defined by Eq. (5). We assume that the microstructural $\delta\nu(\mathbf{x})$ is generated from a spatially uniform zero-mean Gaussian random process with a normalized autocorrelation function

$$K(\mathbf{x}-\mathbf{x}') = \frac{\langle \delta\nu(\mathbf{x})\,\delta\nu(\mathbf{x}')\rangle_m}{\langle \delta\nu^2(\mathbf{x})\rangle_m}, \tag{8}$$

which is approximated as a Gaussian

$$K(\mathbf{x}-\mathbf{x}') = \exp\left(- \frac{|\mathbf{x}-\mathbf{x}'|^2}{L_m^2}\right). \tag{9}$$

Here, L_m denotes the autocorrelation length for the spatial variation of the microstructure.

B. Inclusion model

We model the scatterer as an isolated inclusion in an otherwise homogeneous and isotropic half-space, as shown in Fig. 1. The scatterer is a weakly scattering spheroidal inclusion centered at $\mathbf{x}_1=(\mathbf{r}_1,z_1)$. The change in sound velocity in the region of the inclusion gives rise to the Gaussian function $\delta\nu$

$$\delta\nu(\mathbf{x}) = 1 - \frac{c_0^2}{c^2(\mathbf{x})} = V \exp\left(- \frac{|z-z_1|^2}{B^2}\right)\exp\left(- \frac{|\mathbf{r}-\mathbf{r}_1|^2}{A^2}\right), \tag{10}$$

where V is the maximum value of the inclusion's deviation from the host. A denotes the length of the semi-axis of the inclusion parallel to the surface, while B denotes the length of the inclusion's semi-axis in the direction normal to the surface.

C. Rough surface model

We employ the rough surface model used in the previous studies.[1–3] The water–solid interface is assumed to be randomly rough and planar on the average at $z=0$. The surface's height $h(\mathbf{r})$ is described by a spatially uniform, zero-mean Gaussian random process, and assumed to have a Gaussian autocorrelation function

$$\Gamma(\mathbf{r}) = \exp(-r^2/L_s^2). \tag{11}$$

The surface is described by the rms height $h=\langle h^2(\mathbf{s})\rangle_s$ and the autocorrelation length L_s, where $h \ll L_s$.

D. Signal formation model

The signal is calculated for an inclusion in a roughened half-space but in the absence of microstructure. The transducer's signal is obtained from Auld's reciprocity relation[6] and its reduction to the scalar wave equation.[7] In the Born approximation, the signal induced by the velocity variation $\delta\nu$ is given by

$$S(\omega) = C(\omega) \int d^3\mathbf{x}\,\delta\nu(\mathbf{x}) u_0^2(\omega,\mathbf{x}). \tag{12}$$

Here, ω denotes the frequency in radians and $u_0(\omega,\mathbf{x})$ denotes the pressure wave field transmitted through the rough water–solid interface and propagating in an otherwise uniform half-space of velocity c_0. The constant $C(\omega)$ is defined to make connection to experiment and is given by $C(\omega) = i\omega \mathrm{CF}/4\rho c_0^2 E(\omega)$, where ρ denotes the density of the fluid and $E(\omega)$ is the electrical power input to the transducer. Finally, CF is the conversion factor from electrical to acoustic power.

E. Wavefield propagation and transmission through the rough surface

We imagine that the experiment is carried out with a single immersion transducer that acts as both transmitter and receiver. The transducer is assumed to be immediately above the water–solid interface at $z=0^-$; it is easy to generalize to a finite stand-off distance in the water. The pressure on the transducer's surface is prescribed to be

$$u_0(\omega,\mathbf{r},z=0^-) = \exp\left[-\left(\frac{1}{R^2} + i\,\frac{k_s}{2F}\right) r^2\right], \tag{13}$$

where $z=0^-$ indicates the wave field at the transducer face just outside the solid, R is the radius of the transducer, k_s

denotes the wave number in the solid, and F denotes the focal length in the solid.

The effects of the rough surface on $u_0(\omega,\mathbf{x})$ is calculated using the phase-screen approximation,[8] which approximates the pressure just inside the solid ($z=0^+$) by

$$u_0(\omega,\mathbf{r},z=0^+)=T_0^{w\to s}\exp(i\phi(\mathbf{r}))u_0(\omega,\mathbf{r},z=0^-), \tag{14}$$

where $T_0^{w\to s}$ is the transmission coefficient at the plane interface and the arrow indicates the wave transmission from water to solid. The phase variation at each point on the surface is defined in terms of the local height of the surface by

$$\phi(\mathbf{r})=(k_s-k_w)h(\mathbf{r})=\delta k h(\mathbf{r}). \tag{15}$$

The phase-screen approximation is most valid if both the wavelength and the rms height are much less than the autocorrelation length.

The propagation of the wave from the rough surface into the bulk of the solid plate is computed using the Fresnel approximation.[3] The result is

$$u_0(\omega,\mathbf{x})\approx-\frac{ik_sT_0^{w\to s}e^{ik_sz}}{2\pi z}\int d^2\mathbf{s}u_0(\omega,\mathbf{r},z=0^-)$$
$$\times\exp(i\phi(\mathbf{s}))\exp\left(\frac{ik_s}{2z}|\mathbf{s}-\mathbf{r}|^2\right). \tag{16}$$

III. RESULTS

The next two sections are the heart of this paper. In this section, we evaluate the NSNR in Eq. (2) which is reorganized as

$$\mathrm{NSNR}\equiv\frac{\langle|S|^2\rangle_s/|S_0|^2}{\langle\langle|N|^2\rangle_s\rangle_m/\langle|N_0|^2\rangle_m} \tag{17}$$

for a series of example problems. Parameters are chosen to be typical of inspections occasioned by nondestructive evaluation. We will plot the NSNR as a function of depth for the frequency, rms height, autocorrelation length, etc. We will calculate the normalized signal $\langle|S|^2\rangle_s/|S_0|^2$ [the numerator of Eq. (17)] using Eqs. (28) and (29a)–(29e) of Ref. 1, and the normalized noise $\langle\langle|N|^2\rangle_s\rangle_m/\langle|N_0|^2\rangle_m$ [the denominator of Eq. (17)] using Eqs. (13) and (14a)–(14e) of Ref. 5.

Results are illustrated with graphs calculated, unless otherwise noted, for the following parameters: transducer radius $R=1$ cm, frequency $f=10$ MHz, rms surface height $h=0.02$ mm, $V=5\%$, surface autocorrelation length $L_s=1.5$ mm, microstructural autocorrelation length $L_m=0.1$ mm, speed of sound in water $c_w=1500$ m/s, and the speed of longitudinal sound in the solid $c_s=6300$ m/s. Both focused, $F=40$ mm, and unfocused transducers are considered. For a transducer immediately above the water–solid interface, R denotes the radius of the transducer, while F denotes the focal length of the transducer in the solid. The analytic series solutions are found to converge relatively rapidly for the parameters mentioned above; typically, we sum each index from 1 to 10. The most important parameter in the problem is $\sigma^2=\delta k^2h^2$, which is a nondimensional estimate of the ratio of the wavelength and the rms height. Basically, σ^2 determines the division of energy between the specular and diffuse beams.

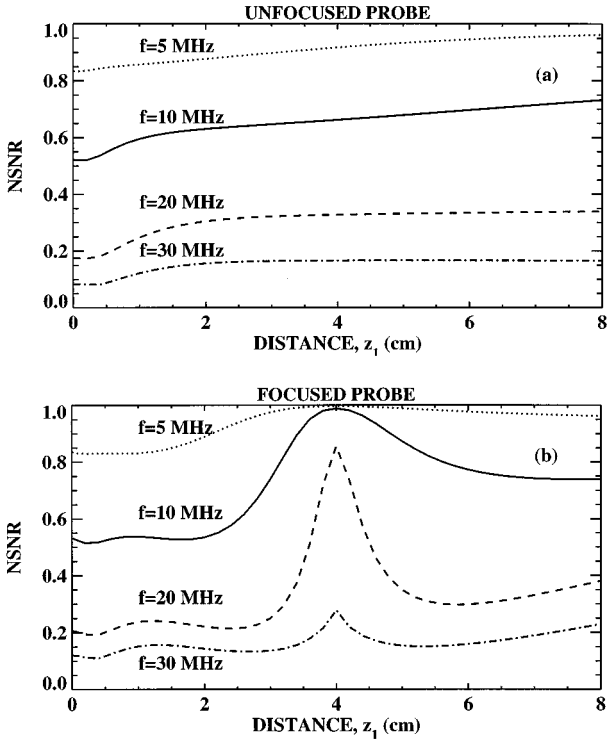


FIG. 2. Shows the NSNR as a function of depth for frequencies of 5, 10, 20, and 30 MHz for (a) an unfocused transducer and (b) a focused transducer. The other parameters are $A=1$ mm, $R=10$ mm, $L_s=1.5$ mm, $L_m=0.1$ mm, $h=0.02$ mm, and $F=40$ mm.

The frequency dependence of the NSNR is shown in Fig 2. The NSNR of an unfocused transducer is shown in Fig. 2(a) for $f=5$, 10, 20, and 30 MHz (correspondingly, $\sigma^2=0.10$, 0.41, 1.63, 3.67). For an unfocused transducer the NSNR is least for inclusions near the surface. It then increases monotonically with greater depth. Roughness has the least effect on the NSNR at low frequencies and the most at high frequencies.

A striking improvement in the NSNR is observed in Fig. 2(b), which shows the effects of using a relatively weakly focused transducer ($R=1$ cm and $F=4$ cm). As one might expect, the SNR is best at the nominal focal depth of 4 cm. For 5 MHz ($\sigma^2=0.10$) and 10 MHz ($\sigma^2=0.41$) roughness has almost no effect on the SNR for this focused transducer; i.e., the NSNR is nearly 1 at the focal depth. A large improvement is also seen for 20 MHz ($\sigma^2=1.63$), where the NSNR is increased to 0.85 compared to the value of 0.35 for the unfocused transducer. At 30 MHz ($\sigma^2=3.67$) the NSNR at the focal depth is only slightly greater than background. The improvement in NSNR due to focusing is expected to disappear altogether at even higher frequencies. Finally, the peak in NSNR is most diffuse at 5 MHz and becomes increasingly narrow as the frequency increases.

Figure 3(a) and (b) show the results of keeping the frequency fixed at 10 MHz and varying the rms height. The four curves correspond to the same values of $\sigma^2=0.10$, 0.41, 1.63, 3.67, used in Fig. 2. For unfocused transducers, the curves have their minimum value at $z_1=0$. The curves then increase relatively rapidly over a characteristic depth of 1 cm. The curves for $h=0.01$ mm, 0.02 mm, and 0.04 mm then

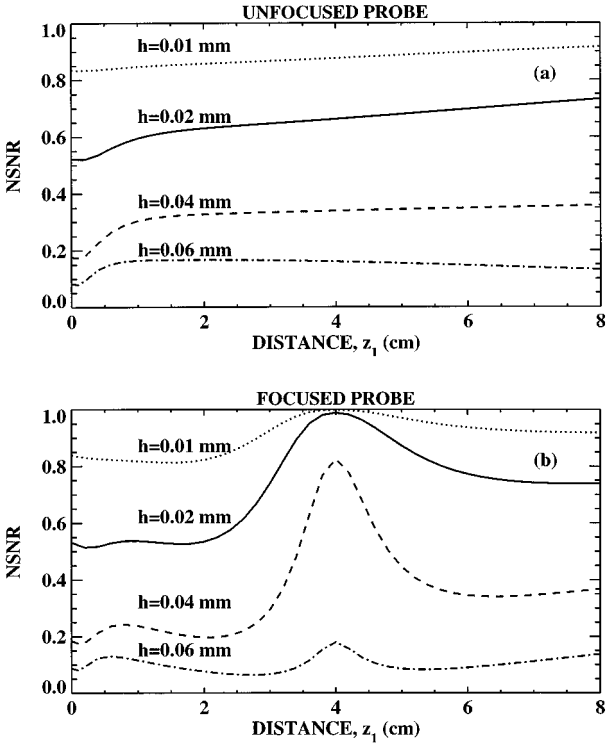


FIG. 3. Shows the NSNR as a function of depth for rms surface roughness of 0.01, 0.02, 0.04, and 0.06 mm for (a) an unfocused transducer and (b) a focused transducer. The other parameters are $A=1$ mm, $R=10$ mm, $L_s=1.5$ mm, $L_m=0.1$ mm, $f=10$ MHz, and $F=40$ mm.

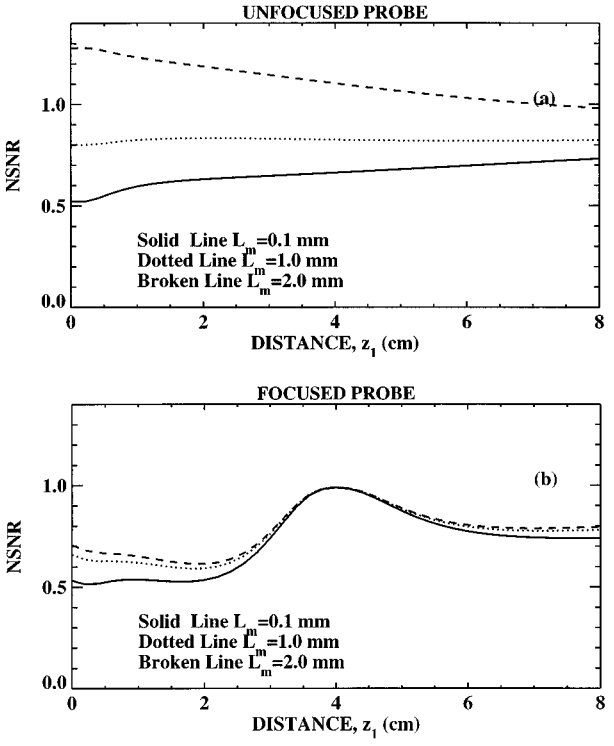


FIG. 4. Shows the NSNR as a function of depth for microstructural autocorrelation lengths of 0.1, 1.0, and 2.0 mm for (a) an unfocused transducer and (b) a focused transducer. The other parameters are $A=1.0$ mm, $R=10$ mm, $L_s=1.5$ mm, $h=0.02$ mm, $f=10$ MHz, and $F=40$ mm.

increase slowly and monotonically up the maximum distance calculated. These curves are very similar to those in Fig. 2(b). The curve calculated for $h=0.06$ mm is qualitatively different. It reaches a maximum at a depth 1 cm and then slowly decreases monotonically.

The improvement in the SNR due to focusing is apparent when the NSNR is plotted as shown in Fig. 3(b). Large improvements are found for $h=0.01$, 0.02, and 0.04 mm. Overall the curves in Figs. 2 and 3 are fairly similar for the same values of σ^2. However, variations in the rms height do not cause the width of the peak to vary significantly. Also the curves for $h=0.02$ mm ($\sigma^2=0.41$) and 0.04 mm ($\sigma^2=1.63$) have slightly lower peak values than the curves for $f=10$ MHz ($\sigma^2=0.41$) and 20 MHz ($\sigma^2=1.63$). Finally, the curve for $h=0.06$ mm ($\sigma^2=3.67$) is significantly lower ($\sim 50\%$) than the curve for 30 MHz ($\sigma^2=3.67$) over almost the entire curve.

The NSNR also depends interestingly on the relative sizes of the microstructure, the surface roughness, and the inclusion. The dependence of the NSNR on the autocorrelation length of the microstructure is shown in Fig. 4. The semi-major axis of the inclusion and the surface autocorrelation length are chosen to be 1 mm and 1.5 mm, respectively. The NSNR for the unfocused transducer is shown in Fig. 4(a). As expected, there are qualitative differences in the curves depending on whether the microstructural autocorrelation length is smaller or larger than the lateral size of the inclusion and/or the surface autocorrelation length. Perhaps most surprisingly, the NSNR is predicted to be greater than one for $L_m=2.0$ mm! That is, surface roughness is predicted to increase the SNR. For $L_m=2.0$ mm, the NSNR is predicted to be greatest at $z_1=0$ and to decrease monotonically with greater depth. The NSNR is nearly constant and less than one when $L_m=A=1$ mm. Finally, the NSNR is least, when $L_m=0.1$ mm $\ll L_s$ or A.

Focused transducers [Fig. 4(b)] yield a NSNR ratio of ~ 1.0 at the focal depth for all three microstructural autocorrelation lengths. It is interesting that for the focused transducer, the NSNR for $L_m=2.0$ mm is less than for the unfocused transducer. All of our calculations are for the NSNR, that is, the SNR for the rough surface normalized by the SNR for the smooth surface. We expect that the SNR will be higher for the focused transducer. The unfocused transducer starts out with a much lower SNR and is slightly improved by surface roughness, while the focused transducer starts out with a relatively large SNR, which is effectively unchanged by the surface roughness.

The dependence of NSNR on the surface autocorrelation length L_s is shown in Fig. 5(a) and (b). The semi-major axis of the inclusion is chosen to be $A=0.5$ mm and $L_m=0.1$ mm. The NSNR for the unfocused transducer is shown in Fig. 5(a). For small z, the NSNR is larger for larger L_s. For large z, the situation reverses and the NSNR is larger for smaller L_s. For the focused transducer all three curves show a maximum with a value of ~ 1 at the nominal focal depth of 4 cm. The NSNR orders with L_s in the same way as it does for the unfocused transducer.

The dependence of the NSNR on the lateral dimension of the inclusion is shown in Fig. 6(a) and (b). For the unfocused transducer, the results tend to the same limit at large depths. The larger the inclusion, the smaller the NSNR. The

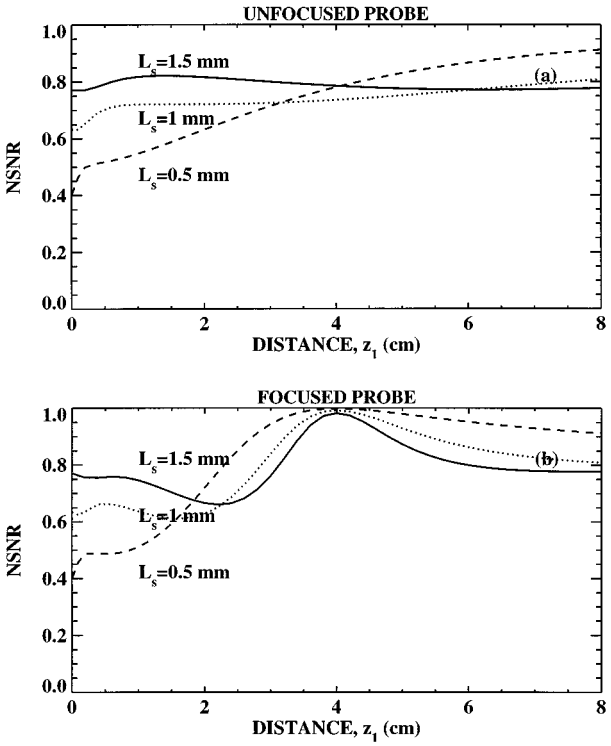


FIG. 5. Shows the NSNR as a function of depth for surface autocorrelation lengths of 0.5, 1.0, and 1.5 mm for (a) an unfocused transducer and (b) a focused transducer. The other parameters are $A=0.5$ mm, $R=10$ mm, $h=0.02$ mm, $L_m=0.1$ mm, $f=10$ MHz, and $F=40$ mm.

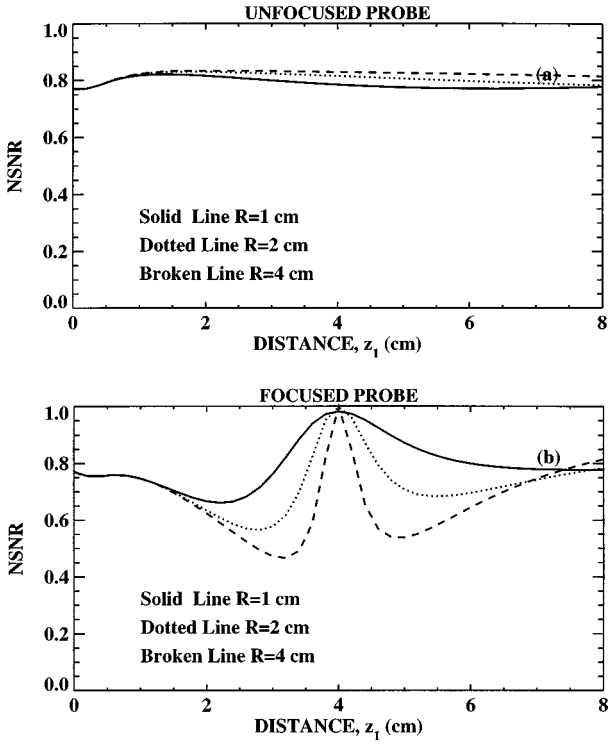


FIG. 7. Shows the NSNR as a function of depth for transducer radii of 10, 20, and 40 mm for (a) an unfocused transducer and (b) a focused transducer. The other parameters are $A=0.5$ mm, $Ls=1.5$ mm, $h=0.02$ mm, $L_m=0.1$ mm, $f=10$ MHz, and $F=40$ mm.

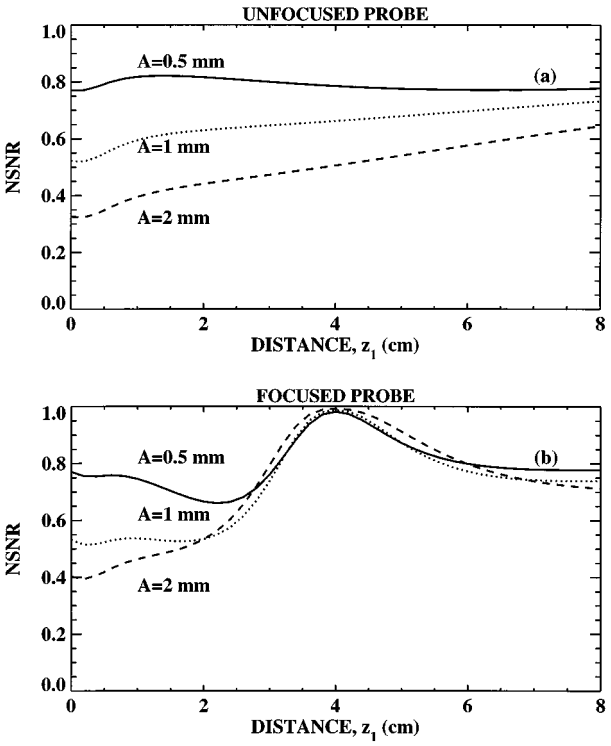


FIG. 6. Shows the NSNR as a function of depth for semi-major axis of the inclusion of 0.5, 1.0, and 2.0 mm for (a) an unfocused transducer and (b) a focused transducer. The other parameters are $L_s=1.5$ mm, $R=10$ mm, $h=0.02$ mm, $L_m=0.1$ mm, $f=10$ MHz, and $F=40$ mm.

effects of focusing is to increase the NSNR to nearly one.

Finally, Fig. 7(a) and (b) shows the dependence of the NSNR on the size of the transducer. There is little effect for the unfocused transducer. For all three radii considered, $R=1$, 2, and 4 cm, the NSNR is nearly constant with depth and has a value of roughly 0.8. The NSNR is very slightly larger for the larger radii transducers. However, varying the transducer's radius causes large changes in the NSNR curves for focused transducers. In all three cases the NSNR is nearly one at the focal depth. However, the width of the maximum varies strongly with the radius of the transducer. The width is greatest for $R=1$ cm and is least for $R=4$ cm. The curves for $R=2$ and 4 cm also have a characteristic ''butterfly'' shape.

IV. PHYSICAL INTERPRETATION

We have found two initially unexpected results. First, surface roughness can *increase* the SNR if we have a small inclusion and the material noise is generated by relatively large and weakly scattering microstructure. Second, we also found that in many cases the deleterious effects of surface roughness on the SNR can be nearly removed by the use of an appropriately focused transducer. In this section, we physically motivate these two results in turn.

A. Normalized signal-to-noise ratio greater than one

A normalized signal-to-noise ratio greater than one means that roughness suppresses material noise more than it suppresses the signal from the inclusion. This can happen if

the characteristic spatial length scale of the microstructure is larger than the size of inclusion. We can explain this result by starting with the signal S from one of the microcrystallites that constitute the microstructure. We rewrite the signal in the Born approximation, Eq. (12),

$$S = C(\omega) \int d^3\mathbf{y}\, \delta\nu(\mathbf{y}) u_0^2(\omega, \mathbf{y}). \tag{18}$$

The rough surface introduces fluctuations in phase and amplitude of the wave field. These fluctuations are integrated over the volume of the microcrystallite; this tends to average out and reduce the signal. The larger the microcrystallite, the larger the region of integration and the more completely the fluctuations are averaged out. The prediction that the material noise decreases with increasing microstructure size is due to this *intrinsic* spatial averaging. If the microstructure is large, the material noise is strongly suppressed by surface roughness. If the inclusion is relatively small, its signal is suppressed much less strongly. Hence, the NSNR will be greater than one for a small strongly scattering inclusion embedded in a relatively large and weakly scattering microstructure.

B. Improved signal-to-noise ratio due to focusing

For focused transducers, the graphical results suggest that moderate surface roughness has little effect on the SNR at the focal depth. Figures 2–7 show that the NSNR is very nearly equal to one for an inclusion found in the focal zone if σ^2 is not too large. Why? Crudely, the rough surface divides the beam into two parts: a specular part that focuses and a diffuse part that does not. At the focal depth, the specular part of the beam is concentrated in a small volume, while the diffuse part of the beam is widely dispersed. At the focal depth, the signal from the inclusion and the noise from microstructure are in fact dominated by the concentrated specular part of the beam; the diffuse part contributes negligibly. It can be shown that the specular beam is related to the beam in the absence of roughness by a frequency-dependent transmission constant. This transmission constant divides out when we compute the SNR and thus we find that the NSNR is one (supposing that we can completely exclude the contribution of the diffuse beam). Thus, we can expect NSNR ~ 1 if σ^2 is not too large. For sufficiently large σ^2 the diffuse beam dominates the scattering and we expect little improvement in the SNR due to focusing.

The separation of the beam into a specular and a diffuse component can be seen in Fig. 8 for focused (left) and unfocused (right) transducers. In contrast with the rest of the work in this paper, this figure does not show a statistical feature. Rather, it shows the calculated total field as a function of depth beneath the surface for one *particular* realization of the rough surface, calculated using the phase-screen approximation and the expansion in plane waves method as discussed in Ref. 3. Figure 8(a) and (e) show the focused and unfocused fields at depth $z = 0.02$ cm. Figure 8(b) and (f) are for $z = F/2$, etc. The unfocused transducer shows a strong diffuse component at all depths. On the other hand, one can clearly see that the focused wave field is dominated by the specular part of the beam at $z = F$. However, the specular signal is attenuated compared to the field transmitted through a smooth surface, shown by the dotted curve.

The specular part of the beam is defined as just the average $\langle u \rangle_s$ transmitted through the rough surface. The behavior of the specular part of the beam is relatively simple and has been recently investigated in Ref. 9 for plane waves and for weakly focused transducers at normal incidence. At a given frequency, the specular beam $\langle u \rangle_s$ was found to be proportional to beam in the absence of roughness, u_{Smooth}; that is,

$$\langle u \rangle_s = T_R(\omega) u_{\text{Smooth}}, \tag{19}$$

where $T_R(\omega)$ is a frequency-dependent transmission coefficient. This transmission constant is, within the phase-screen approximation,[2,9]

$$T_R(\omega) = e^{-\sigma^2/2}. \tag{20}$$

Here, σ^2 denotes the square of the change in the wave number times the rms height.

The behavior of the SNR is now clear. As argued in the first paragraph of this section, the signal and the noise both arise from the specular part of the beam if the surface is not too rough. Surface roughness only changes the specular part of the beam by an overall constant. This constant clearly divides out of the signal-to-noise ratio. Consequently, the SNR is nearly unchanged for inclusions at the focal depth for moderate degrees of roughness.

The question naturally arises ''For what ranges of parameters is the effect of surface roughness on the SNR compensated by focusing?'' This question can, of course, be answered in detail by summing the series solutions for $\langle |S|^2 \rangle_s$ and $\langle \langle |N|^2 \rangle_s \rangle_m$ given in Refs. 1, 4, and 5 as we have done for the examples given in this paper. Here, we will only give a qualitative argument that shows how the various parameters enter. We restrict ourselves to the most typical case where the microstructure is much smaller than the wavelength and the surface autocorrelation length. A key part of our argument is that the material noise should be dominated by the specular part of the beam, $\langle \langle |N_S|^2 \rangle_s \rangle_m$, and that the noise generated by the diffuse part of the beam, $\langle \langle |N_D|^2 \rangle_s \rangle_m$, should be negligible. That is,

$$\langle \langle |N_S|^2 \rangle_s \rangle_m \gg \langle \langle |N_D|^2 \rangle_s \rangle_m. \tag{21}$$

We will qualitatively estimate both the specular and the diffuse contributions to the noise next. We continue to draw on the idea that for each realization of the rough surface the beam is split into a specular and a diffuse component. The specular component focuses just as the beam would in the absence of roughness. The diffuse component does not focus, and for the purposes of a qualitative estimate we assume that it propagates in a beam whose radius is approximately that of the transducer. The material noise (backscatter power) can be estimated from[4]

$$P(\omega, z) = \langle \langle |N|^2 \rangle_s \rangle_m = \beta \int d^2\mathbf{r} |u(\omega, \mathbf{r}, z)|^4, \tag{22}$$

where β is an unimportant constant of proportionality. First, we will estimate the material noise due to the specular part of

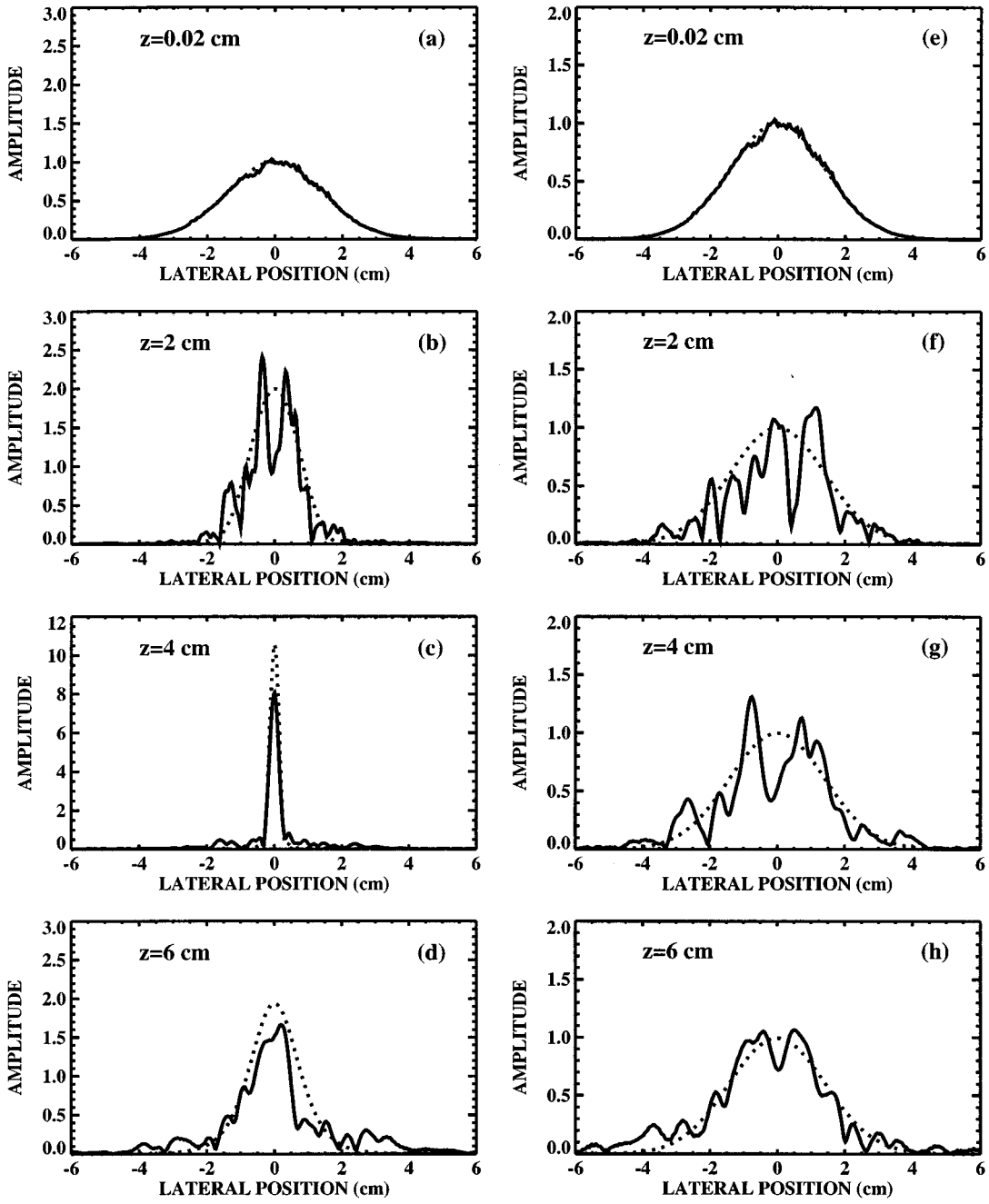


FIG. 8. Shows the field of a focused (a)–(d) and unfocused (e)–(h) beam at various depths beneath a rough surface for $f=10$ MHz, $F=4$ cm, $R=1/2$ in (1.27 cm), $L_s=1.5$ mm, $h=25$ μm. The solid and broken lines denote the field transmitted through the rough and smooth surfaces. Note the change in scale for (c) and the strong peak that arises at the focal depth despite the presence of a rough surface.

the beam. The beam's intensity at the focal spot will be increased by the effects of focusing by the ratio of area of the transducer α_T to the area of the focal spot α_F. For our qualitative estimate we assume that beam intensity is constant over the area of the focal spot and obtain (to within factors of order one)

$$\langle\langle|N_s|^2\rangle_s\rangle_m \approx \beta e^{-2\sigma^2}\left(\frac{\alpha_T}{\alpha_F}\right)^2 \alpha_F . \tag{23}$$

The material noise due to the diffuse part of the beam can be estimated in the same way as

$$\langle\langle|N_D|^2\rangle_s\rangle_m \approx \beta(1-e^{-\sigma^2})^2\alpha_T . \tag{24}$$

For a focused Gaussian beam, the radius of the beam at the focus is $F\lambda/\pi R$. Upon combining (21), (23), and (24), one obtains the following inequality:

$$e^{-\delta k^2 h^2}\left(1+\frac{\pi R^2}{F\lambda}\right) > 1. \tag{25}$$

The more this inequality is satisfied, the more the NSNR is improved by focusing the transducer. The inequality indicates, as expected, that the NSNR tends to degrade with in-

creasing rms height. On the other hand, it is improved either by increasing the transducer's radius or decreasing its focal length. For the parameters used in this paper, the inequality is satisfied for rms heights less than 50 μm at 10 MHz ($\sigma^2 \sim 2.5$), in rough agreement with the detailed results shown in Fig. 3(b).

V. SUMMARY

We have provided a method for calculating the normalized SNR for weakly scattering inclusions that are located beneath a randomly rough surface and that are embedded in a weakly scattering microstructure. We have calculated a series of example problems with parameters taken from ultrasonic nondestructive evaluation. These examples indicate that the SNR is degraded by surface roughness for the typical case of a large inclusion embedded in a relatively fine microstructure. The examples lead to the prediction that the SNR is actually increased by surface roughness if we have a small strongly scattering inclusion embedded in a relatively large and weakly scattering microstructure. Finally, for ''moderate roughness'' as defined by a simple inequality, the examples lead to the prediction that focusing can ameliorate or remove the effects of surface roughness on the SNR.

ACKNOWLEDGMENTS

This material is based in part on work performed at the Center for NDE at Iowa State University as part of the Engine Titanium Consortium operated by the Center for NDE at Iowa State University including participants at General Electric, Pratt & Whitney, and Allied Signal, and supported by the Federal Aviation Administration Technical Administration Center, New Jersey under Grant No. 94G048.

[1] M. Bilgen and J. H. Rose, ''Mean and variance of the ultrasonic signal from a scatterer beneath a rough surface,'' J. Acoust. Soc. Am. **98**, 2217–2225 (1995).
[2] P. B. Nagy and J. H. Rose, ''Surface roughness and the ultrasonic detection of subsurface scatterers,'' J. Appl. Phys. **73**, 566–580 (1993).
[3] M. Bilgen and J. H. Rose, ''Effects of one-dimensional random rough surfaces on acoustic backscatter: Utility of phase-screen and Fresnel approximations,'' J. Acoust. Soc. Am. **96**, 2849–2855 (1994).
[4] M. Bilgen and J. H. Rose, ''Rough surface effects on incoherent scattering from random volumetric scatterers: Approximate analytic series solution,'' J. Acoust. Soc. Am. **96**, 2856–2869 (1994).
[5] M. Bilgen and J. H. Rose, ''Acoustic backscatter from materials with rough surfaces and finite size microstructure: Theory,'' J. Acoust. Soc. Am. **101**, 264–271 (1997).
[6] B. A. Auld, ''General electromechanical reciprocity relations applied to the calculation of elastic wave scattering coefficients,'' Wave Motion **1**, 3–10 (1979).
[7] M. Bilgen, Ph.D. dissertation, Iowa State University, Ames, IA, 1993.
[8] P. B. Nagy and L. Adler, ''Surface roughness induced attenuation of reflected and transmitted ultrasonic waves,'' J. Acoust. Soc. Am. **82**, 193–197 (1987).
[9] K. Holman, M. Holland, J. Miller, P. Nagy, and J. Rose, ''Effective ultrasonic transmission coefficient for rough surfaces,'' J. Acoust. Soc. Am. **100**, 832 (1996).

Characterization of ultrasonic transducers by means of double-exposure holographic interferometry

Lucio N. Tello
Computalog Research, Inc., 500 Winscott Road, Fort Worth, Texas 76126

Truman D. Black and Don P. Cudmore
The University of Texas at Arlington, Department of Physics, Arlington, Texas 76019

Robert Magnusson
The University of Texas at Arlington, Department of Electrical Engineering, Arlington, Texas 76019

(Received 24 February 1994; revised 3 July 1996; accepted 3 September 1996)

Double-exposure holographic interferometry is applied to measure the surface displacements of ultrasonic transducers. Interferograms revealing detail surface displacements are produced using iron-doped lithium niobate crystals as the recording medium. Lead zirconate titanate disks poled in the longitudinal thickness mode possessing deposited, shaped electrodes are examined. Transducers with circular, strip, rectangular, and rosette electrode shapes are investigated. The displacement measurements reveal that the electrode ''drags'' the rest of the transducer surface. Thus the unmetallized portion of the transducer also contributes to the acoustic field. Pressure field calculations using the experimentally measured displacements are performed. It is shown that proper geometrical design of the electrode yields reduction of the sidelobes in the far field and smoothing of the maxima and minima in the near field, and improved focusing via reduced acoustic beam divergence. Most importantly, this work reveals that the unmetallized portion of the transducer contributes significantly to the displacement and to the resulting pressure fields. © *1997 Acoustical Society of America.* [S0001-4966(97)05212-0]

PACS numbers: 43.38.Fx, 43.38.Ar, 43.35.Yb, 43.58.Vb [SLE]

INTRODUCTION

Recent applications of holographic interferometry using photorefractive crystals such as lithium niobate ($LiNbO_3$) as recording media[1] include visualization of aerodynamic flow fields[2] and heat transfer patterns.[3] This paper reports the use of this technique for measurements of surface deformation of piezoelectric transducers. First, the basic principles of double-exposure holography using photorefractive crystals are briefly reviewed. Second, the application of double-exposure holographic interferometry for evaluation of surface displacements of ultrasonic transducers is described. Third, the results of pressure calculations for shaped electrode transducers are presented.

Surface vibration displacement patterns using holographic interferometry were recorded on film by Powell and Stetson[4] using a solenoid-driven metal can bottom. Photographs of reconstructed images were reported by Wojciechowska and Śliwinski[5] for vibrational amplitude distributions on the surface of a barium titanate ultrasonic transducer excited near resonance. These pictures were reconstructions of time-averaged holograms using films. The same type of reconstructions using photorefractive $Bi_{12}SiO_{20}$ crystals were presented by Marrakchi *et al.*[6] Their experiment consisted of a loudspeaker membrane driven at its resonance and harmonic frequencies. Other interferometric methods to measure displacements from metal plates were reported by Løkberg[7] and by Monchalin.[8]

The purpose of the present paper is to report the use of double-exposure holographic interferometry in iron-doped $LiNbO_3$ crystals to evaluate surface displacements of ultrasonic transducers with exotic electrode configurations. Special focus will be given to partially electroded transducers similar to those studied by Noorbehesht,[9] and by Martin and Breazeale[10] for rectangular and strip electrodes, and by Haselberg and Krautkrämer,[11] and Harrison and Balcer-Kubiczek,[12] for rosette or asteroidal shape. Identical lead zirconate titanate disks 2.54 cm in diameter and 0.45 cm thick provided by American Piezo Ceramics of Mackeyville, Pennsylvania[13] were mechanically polished to an optical finish (95% reflectance or better) with partial aluminum electrodes with different geometries subsequently vacuum deposited on their radiating surfaces. The measured displacements are subsequently used to calculate the resulting acoustic pressure fields. The four electrode shapes selected in this research are illustrated in Fig. 1.

I. HOLOGRAPHIC TECHNIQUE

Applications of holographic interferometry using a photorefractive crystal such as lithium niobate ($LiNbO_3$) as the storage material include visualization of air flows[1,2] by Hafiz *et al.* in 1989 and heat flow patterns[1,3] by Magnusson *et al.* in 1990. The technique has also been used for surface deformation measurements.[14] This paper will present measurements and analysis of piezoelectric transducer displacements using this technique. Double-exposure holograms using $LiNbO_3$ were written and read in every experimental phase of this research.

Double-exposure holography differs from simple holography in that two writing steps are needed instead of one. First, a Fourier transform hologram of the object is recorded

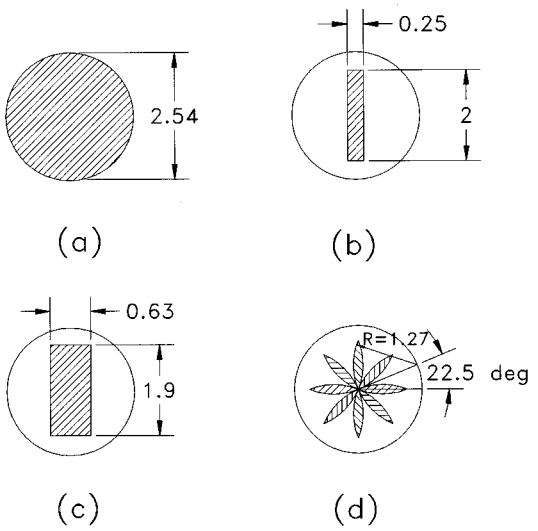


FIG. 1. Electrode shapes selected for the research (a) full circular, (b) strip, (c) rectangular, (d) rosette. All dimensions are in cm.

in the crystal. Then, the object is disturbed by some means and a second hologram is written over the first hologram. The reading is done by blocking the object beam using the reference beam to reconstruct both holograms. The result is an interference fringe pattern that reveals the difference between the first hologram and the second hologram. Therefore, double-exposure holographic interferometry refers to two written holograms that generate simultaneous interference fringes during the reading process. This procedure is useful for structural deformation and vibration analysis[4] since interferometry makes it possible to measure the relative displacement of the object surface. Additionally, the interference pattern indicates visual object flaws, hair line fractures, or places of stress. This paper demonstrates application of this technique to measure surface displacements of ultrasonic transducers.

Hologram formation in iron-doped $LiNbO_3$ occurs when exposure to light causes a distribution of the space charge

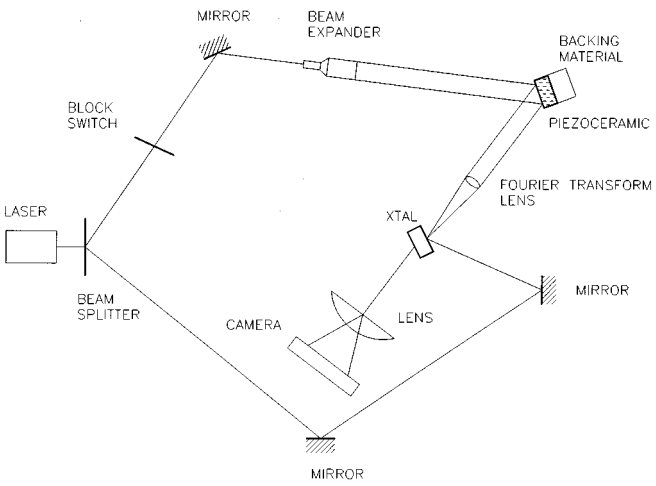
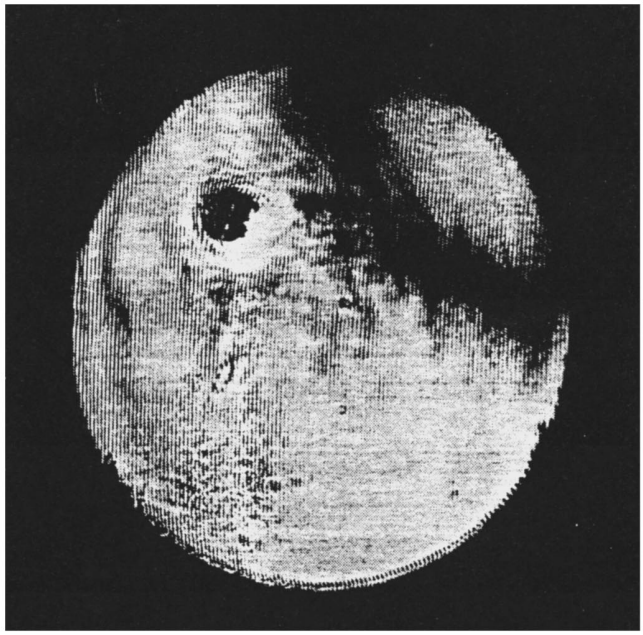
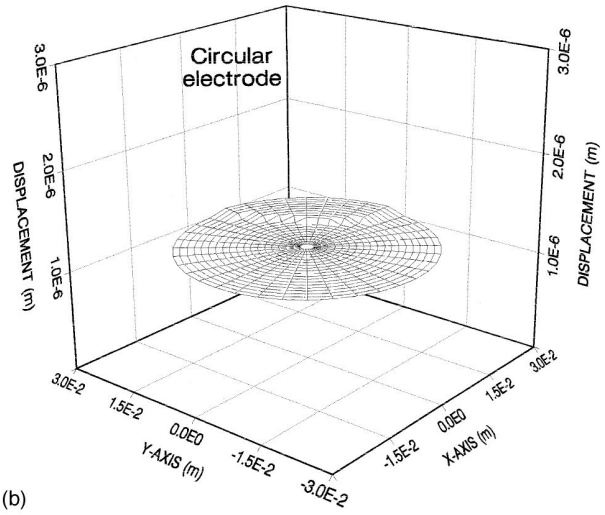


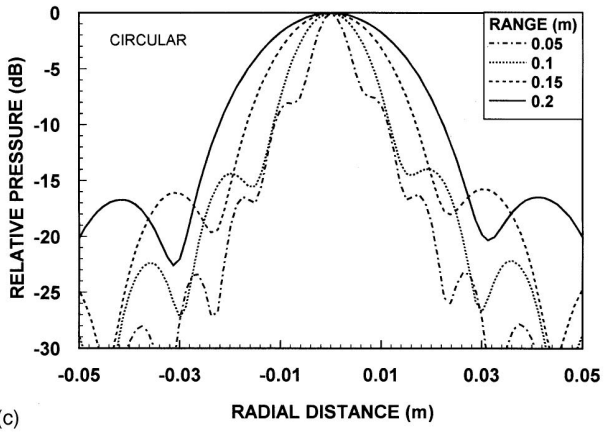
FIG. 2. Experimental system used to write and read double-exposure holograms.



(a)



(b)



(c)

FIG. 3. Full circular electrode transducer. (a) Reconstructed double-exposure hologram with 2000 V applied during the second exposure. (b) Surface displacement calculated from the double-exposure hologram. (c) Calculated pressure in dB relative to the maximum at different ranges.

density corresponding to the optical Fourier transform of the object. Because of the electro-optic effect this charge density distribution is converted into variations of the index of refraction. This process, referred to as the photorefractive effect, has been studied by several authors.[15–19] Properties such as self-development, erasability and reusability, and fast recording and readout, make photorefractive crystals excellent materials for applications in interferometry.

II. SURFACE DISPLACEMENT MEASUREMENTS AND ANALYSIS

A. Circular electrode

Since the full circular electrode (also called flat piston) has been extensively treated in the literature both theoretically[20,21] and experimentally[22,23] for decades, it will serve as the initial test object to illustrate the operation of the interferometer and to explain the variables that can be derived from the experiments.

The experimental system is schematically shown in Fig. 2. It consists of an argon-ion laser beam operating at wavelength $\lambda_0=514.5$ nm that is split into two beams by a variable beam splitter. The reference beam is used directly whereas the object beam is spatially filtered, expanded, and collimated. The collimated object beam with a diameter of 3 cm is reflected off the face of the ultrasonic transducer and then Fourier transformed by a lens which reduces the beam size to a small volume that is recorded in a $LiNbO_3$ crystal. Exact focus is avoided to prevent saturation or damage to the crystal.[1] The transducer was intimately attached to a fixed backing material of low acoustic impedance and the system mounted on a vibration-isolated optical table where bending or translations perpendicular to the transducer face can be detected within a tenth of a fringe accuracy. This accuracy translates into 25-nm random error in the displacement measurements which is about 2%.

The $LiNbO_3$ crystal ($1\times1\times0.2$ cm^3) used in the experiments was iron-doped (0.05 mole % Fe). Two exposures were recorded, the first without an applied electric field and the second with a dc voltage applied across the transducer. Several holograms were taken at different voltages starting from 250 V and up to 2000 V. The reconstructions revealed fringe patterns corresponding to the transducer surface displacement. A dark fringe was produced with 250 V, a bright fringe with 500 V, and so forth alternating from bright to dark up to a bright fringe at 2000 V. Figure 3(a) presents a photograph of a double-exposure holographic interferogram for the transducer with 2000 V across it during the second exposure. The front and back electrical connections are located at the top of the disk. The dark spot is a distortion in the displacement due to inhomogeneities in the transducer. The dark fringe, on the other hand, could be due also to rotation or bending. These anomalies are indicative of a real transducer. The photograph is an example of the flat piston approaching a position corresponding to a bright fringe, that is 4.0 complete fringes, which corresponds to $2\lambda_0$ or 1029 nm. This result can be used to calculate the scalar piezoelectric coefficient $d_{33}=514$ pm/V which compares well with the

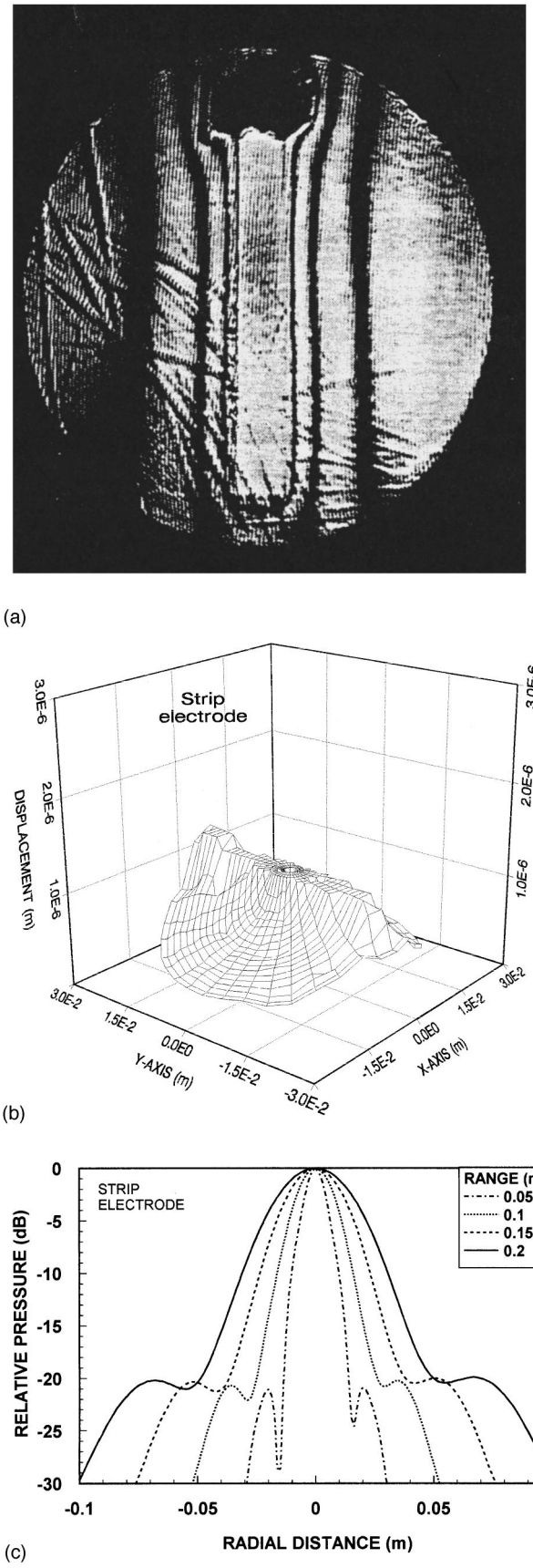
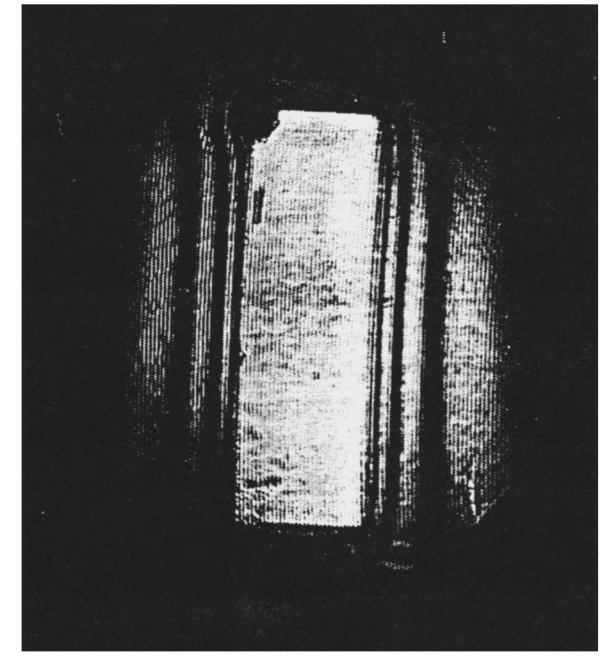
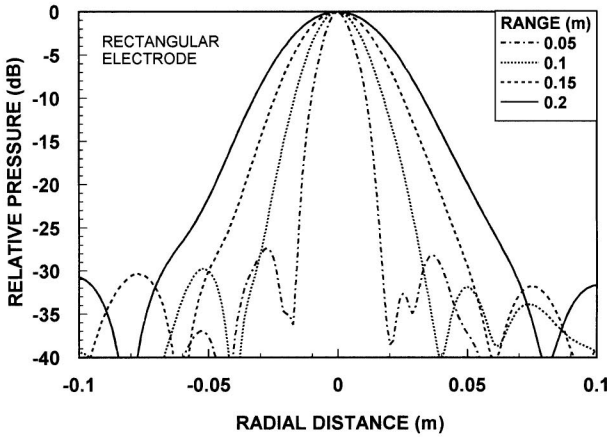


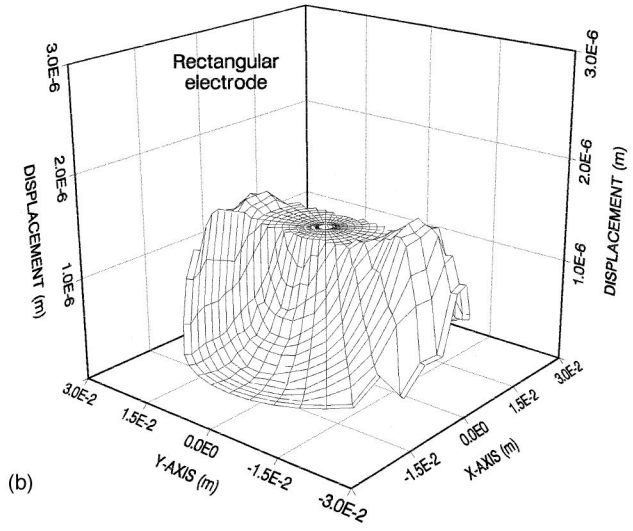
FIG. 4. Strip electrode transducer. (a) Reconstructed double-exposure hologram with 2000 V applied during the second exposure. (b) Surface displacement calculated from the double-exposure hologram. (c) Calculated pressure in dB relative to the maximum at different ranges.



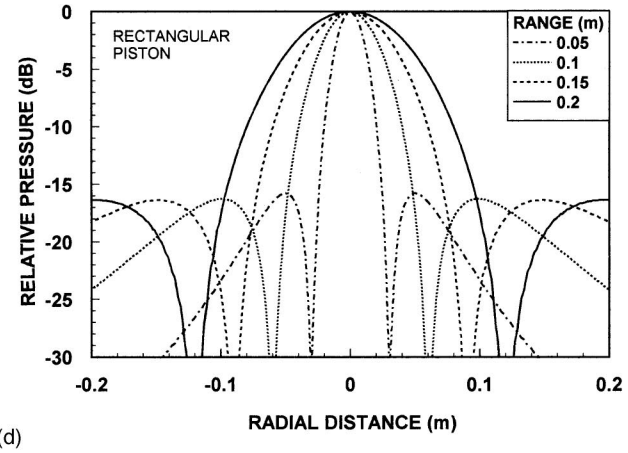
(a)



(c)



(b)



(d)

FIG. 5. Rectangular electrode transducer. (a) Reconstructed double-exposure hologram with 2000 V applied during the second exposure. (b) Surface displacement calculated from the double-exposure hologram. (c) Calculated pressure in dB relative to the maximum at different ranges. (d) Calculated pressure in dB relative to the maximum at different ranges for an ideal rectangular piston with the same dimensions of the metallized rectangular electrode.

manufacturer's specification of $d_{33}=500$ pm/V.[13] Standard interferometric fringe-counting techniques are applied to each point within the circular area. The corresponding displacement result is illustrated in Fig. 3(b). The effect of the transducer anomalies is shown by the surface displacement.

The procedure to evaluate the pressure field theoretically will be the same for all electrode configurations. The fundamental theory for these calculations is presented in the Appendix. Pressure field calculations were performed using the displacement measured by double-exposure holographic interferometry. This is a first-order approximation to the cw pressure field. The transducers had a resonance frequency of 450 kHz in their longitudinal thickness mode and a corresponding acoustic wavelength in water of $\lambda_a=0.33$ cm. The acoustic velocity in water is $\mathbf{v}_a=1.5$ km/s. The acoustic pressure amplitude is given in Pa. Each field point was computed using the displacement, ξ_n, measured experimentally, indeed Eq. (A6) in the Appendix converts to Zemanek's[20] Eq. (1) when the displacement ξ_n is a constant. The result is illustrated in Fig. 3(c) where the data are presented in terms of pressure relative to the maximum of the mainlobe. Each curve is normalized to 0 dB at the center of the beam. This figure shows the first sidelobe going from -8 dB at 0.05 m to -16.5 dB at 0.2 m and eventually approaches the far field approximation result of -18 dB at very long ranges. The asymmetry of about 0.5 dB in the sidelobes is due to the anomalies in the nonideal transducer.

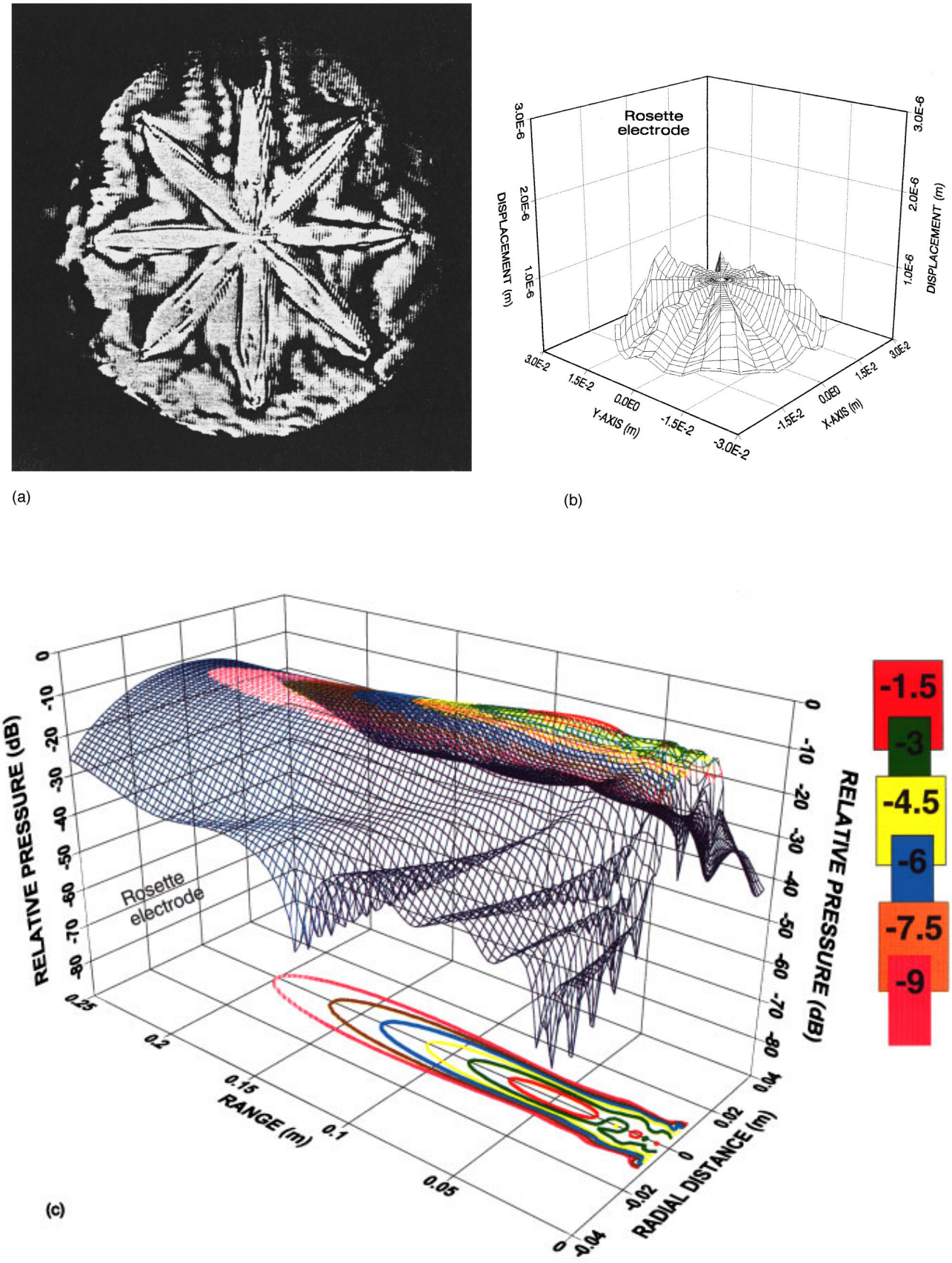


FIG. 6. Rosette electrode transducer. (a) Reconstructed double-exposure hologram with 2000 V applied during the second exposure. (b) Surface displacement calculated from the double-exposure hologram. (c) Pressure amplitude calculated using the displacement of the rosette electrode transducer. (d) Calculated pressure in dB relative to the maximum at different ranges.

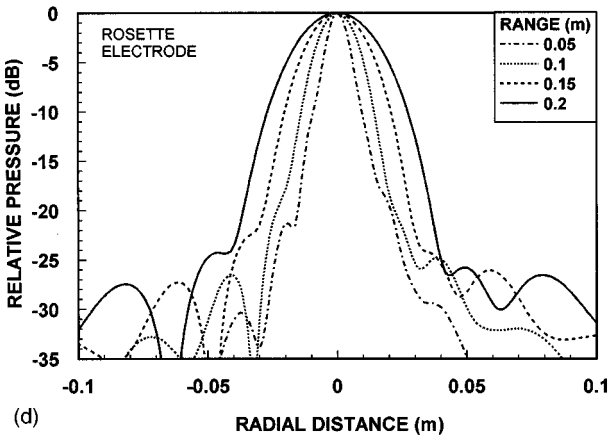


FIG. 6. (*Continued.*)

B. Strip electrode

Piezoelectric transducers with strip electrodes have been treated both theoretically and experimentally[10,11] and found to reduce diffraction lobes as compared to those emitted by piezoelectric disks. Explanations such as electrical and mechanical fringing effects have been offered to account for the sidelobe-reduction property. Figure 4(a) presents a photograph of a reconstructed image of a double-exposure hologram for a transducer with 2000 V during the second exposure. The transducer is a 2.54-cm-diam disk with a 0.254×2.03 cm^2 strip aluminum electrode vacuum deposited on its radiating surface. A series of dark and white fringes surrounding the electrode are evident as well as some scratches. The transducer is still visible and so is its electrical connection where the voltage is applied. The same standard interferometric fringe-counting techniques are applied to each point within the circular area. The corresponding displacement result is illustrated in Fig. 4(b); this figure is a typical result for this and all of the following types of electrode geometries. It can be observed that the total surface of the transducer deforms. Thus the electrode is not the only contributor to the displacement, but effects of adjacent regions need to be considered also. This conclusion will influence the calculation of the pressure field. The corresponding calculated relative pressure-amplitude results are illustrated in Fig. 4(c). Comparison with Fig. 3(c) for the full circular electrode, indicates reduction of the relative amplitude of the sidelobes. The strip electrode disk has a sidelobe below -21 dB at 0.05 m that rises slightly to -20 dB at 0.2 m. This transducer maintains relatively low sidelobes everywhere in the far field.

C. Rectangular electrode

This electrode geometry is of importance when strip electrode properties are desired with more pressure amplitude in the propagating fluid. A PZT-4 disk with a 0.63 cm wide by 1.9-cm-long aluminum electrode vacuum deposited on it was subjected to 2000 V during the second exposure. A reconstructed hologram is shown in Fig. 5(a). As before, fringes appear proportional to the applied voltage. The rectangular electrode is clearly visible along with the electrical connection in the left top corner of the electrode. The resulting surface displacement is presented in Fig. 5(b). The electrode drags the surface as in the case of the strip electrode with similar conclusions being applicable. Thus the electrode is not the only contributor in the calculation of the acoustic field because the entire surface is displaced. The results of the relative pressure calculation using this displacement are presented in Fig. 5(c). The reduction of the relative amplitudes of the sidelobes is clearly illustrated. This transducer has the first sidelobe near -28 dB at 0.05 m dropping down to -31 dB at 0.2 m, significantly lower than for the circular electrode at the same range. The maxima and minima in the near field are smoother than for the full circular electrode. A rectangular electrode transducer combines a reasonable spreading of the far field with a smooth near field behavior. Of interest is the comparison of this result with calculations of a rectangular piston with the same dimensions. Here, ξ_n in Eq. (A6) is constant and the integration is performed over the surface of the electrode only. The result is shown in Fig. 5(d). The rectangular piston shows the first sidelobe below -15 dB at 0.05 m and falls below -16 dB at 0.2 m. Besides the obvious pronounced peaks in the near field as compared to the rectangular electrode disk [Fig. 5(c)], the far field exhibits higher sidelobes with greater divergence.

D. Rosette electrode

Rosette- or asteroidal-shaped electrode transducers have also been treated in the literature. Based on work done by Roderick,[24] Haselberg and Krautkrämer[11] predicted the properties of a Gaussian velocity distribution plane transducer such as the elimination of the maxima and minima on the axis in the near field and reduction of the sidelobes in the far field. The objective is to drive a planar transducer with a radially varying displacement distribution function with fixed phase. A rosette-shaped aluminum electrode was vacuum deposited on a PZT-4 disk 2.54 cm in diameter; see Fig. 1. Figure 6(a) presents a reconstructed double-exposure hologram with 2000 V on the transducer during the second exposure. A series of star-shaped fringes surround the electrode indicating that the transducer surface is deforming according to the shape of the electrode. A transducer surface displacement plot is presented in Fig. 6(b). Again the total transducer face contributes to the acoustic field with the electrode not being the sole contributor. The fringes in between the leaves represent surface deformations corresponding to the ''cross-talk'' referred to by Harrison and Balcer-Kubiczek.[12] Using the surface displacement in the calculation of the acoustic field yields Fig. 6(c), where relatively small diffraction sidelobes are displayed. The peaks and valleys that are present in the acoustic field of the full circular electrode have been smoothed out here. The rosette electrode transducer accomplishes the purpose of reducing the diffraction sidelobes in the near field. Of particular interest is the result for the rosette electrode disk in Fig. 6(d). An asymmetric sidelobe structure below -20 dB at 0.05 m obtains; it drops to -24 dB at 0.2 m. This sidelobe reduction is achieved with geometric electrode design. For all the electrode disks, the presence of the lead connection distorts the field; this is particularly apparent at low values of the relative

TABLE I. Summary of the main features of Figs. 3(c), 4(c), 5(c), 5(d), and 6(d) at a range of 0.2 m or 60 wavelengths from the transducer face.

Transducer	Sidelobe attenuation (dB)	Sidelobe divergence (deg)	−6-dB beamwidth (m)	−6-dB beam divergence (deg)
Circular electrode	16.5	11.6	0.036	5.2
Rectangular electrode	31.1	26.9	0.051	7.1
Rectangular piston	16.3	44.0	0.126	17.5
Strip electrode	20.2	18.6	0.048	6.8
Rosette electrode	24.0	21.8	0.043	6.1

pressure amplitude. Thus the present technique can be used to predict anomalies in the acoustic field. The results presented in Figs. 3(c), 4(c), 5(c), 5(d), and 6(d) are summarized in Table I.

III. CONCLUSIONS

Double-exposure holographic interferometry with a photorefractive storage medium has been used to measure surface displacements of ultrasonic transducers. Four electrode shapes were selected in this study, i.e., the full circular electrode, as well as strip, rectangular, and rosette electrodes. Most importantly, this work reveals that the unmetallized portion of the transducer contributes significantly to the displacement and to the resulting pressure fields. Assuming constant phase over the transducer face, the pressure field calculations can be performed; in this paper standard cw mode theory was utilized. The calculations of the pressure fields resulting from transducers with partial electrodes were compared with those generated by the full circular electrode. The pressure field plots showed smoother peaks and valleys in the near field with reduction of the sidelobes in the far field.

ACKNOWLEDGMENTS

The authors thank Edwin Roberts, Gerald Alford, and Tom Blankinship of Computalog Research, Don A. Larson of Loral Vought Systems, and A. Hafiz of the Department of Electrical Engineering of the University of Texas at Arlington for their assistance during this research. We also thank the reviewers for their constructive comments. Finally we thank Dr. I. Howard of the Physics Department of the University of Texas at Arlington for her help with the field calculations.

APPENDIX: PRESSURE CALCULATION

The harmonic component of the pressure for flat pistons driven in the cw mode has been developed by several authors,[25,26]

$$p = -j\omega\rho\phi. \tag{A1}$$

The velocity potential is described by

$$\phi = \int_{S_p} \mathbf{u}_{(\omega)} \cdot \hat{n} G(\mathbf{r}, \mathbf{r}_p; \omega) dS_p, \tag{A2}$$

where $\hat{\mathbf{n}}$ is the vector normal to the transducer surface and $\mathbf{r}$ and $\mathbf{r}_p$ are the field and source coordinates, respectively. Here, $G(\mathbf{r},\mathbf{r}_p;\omega)$ is the plane Green's function given by

$$\boldsymbol{G}(\mathbf{r}, \mathbf{r}_p; \omega) = \frac{1}{2\pi} \frac{e^{-j\omega[t - R(\mathbf{r},\mathbf{r}_p)/v_a]}}{R(\mathbf{r},\mathbf{r}_p)}, \tag{A3}$$

where $\mathbf{v}_a$ is the velocity of sound in the fluid and $R(\mathbf{r},\mathbf{r}_p)$ is the distance between a field and a source point,

$$R(\mathbf{r},\mathbf{r}_p) = |\mathbf{r} - \mathbf{r}_p|. \tag{A4}$$

If the phase is assumed to be constant on the transducer face then the surface particle displacement, $\boldsymbol{\xi}$, and surface particle velocity are related by[27]

$$\mathbf{u} = \frac{\partial \boldsymbol{\xi}}{\partial t} = -j\omega\boldsymbol{\xi}. \tag{A5}$$

The pressure can be calculated by integrating the normal component of the displacement, giving

$$p = -\frac{\rho\omega^2}{2\pi} \int_{S_p} \xi_n G(\mathbf{r}, \mathbf{r}_p; \omega) dS_p, \tag{A6}$$

where the integration is performed over the entire surface of the transducer. The experimental displacement measurements ξ_n are used in the integrand of Eq. (A6) to calculate the acoustic field. Therefore the static response function is used in the dynamic theory to calculate the field. The assumption is that all the points on the transducer face oscillate in phase.

[1] R. Magnusson, X. Wang, A. Hafiz, T. D. Black, L. N. Tello, A. Haji-Sheikh, S. Konecni, and D. R. Wilson, ''Experiments with photorefractive crystals for holographic interferometry,''Opt. Eng. (Bellingham) **33**(2), 596–607 (1994).

[2] A. Hafiz, R. Magnusson, J. S. Bagby, D. R. Wilson, and T. D. Black, ''Visualization of aerodynamic flow fields using photorefractive crystals,''Appl. Opt. **28**(8), 1521–1524 (1989).

[3] R. Magnusson, A. Hafiz, J. S. Bagby, and A. Haji-Sheikh, ''Holographic interferometry using self-developing optical crystals for heat flux evaluation,'' J. Electron. Packaging **112**, 255–259 (1990).

[4] R. Powell and K. Stetson, ''Interferometric vibration analysis by wavefront reconstruction,''J. Opt. Soc. Am. **55**(12), 1593–1598 (1965).

[5] I. Wojciechowska and A. Sliwinski, ''Examination of vibration amplitude distribution of ultrasonic transducers using optical holography with a

modulated reference beam,''Ultrasonics **19**, 115–119 (1981).
[6] A. Marrakchi, J. P. Herriau, and J. P. Huignard, ''Real time holographic interferometry with photorefractive $Bi_{12}SiO_{20}$ crystals,'' Proc. SPIE, Industrial and Commercial Applications of Holography, San Diego, CA **353**, 24–25 (1982).
[7] J. Løkberg, ''ESPI—The ultimate holographic tool for vibration analysis?,''J. Acoust. Soc. Am. **75**, 1783–1791 (1984).
[8] J. Monchalin, ''Heterodyne interferometric laser probe to measure continuous ultrasonic displacements,''Rev. Sci. Instrum. **56**(4), 543–546 (1985).
[9] B. Noorbehesht, ''Radiation pattern of partially electroded piezoelectric transducers,''J. Acoust. Soc. Am. **70**, 1789–1791 (1981).
[10] F. D. Martin and M. A. Breazeale, ''A simple way to eliminate diffraction lobes emitted by ultrasonic transducers,''J. Acoust. Soc. Am. **49**, 1668–1669 (1971).
[11] K. Haselberg and J. Krauträmer, ''Ein Ultraschall-Strahler für die Werkstoffprüfung mit Verbessertem Nahfeld,''Acustica **9**, 359–364 (1959).
[12] G. H. Harrison and E. K. Balcer-Kubiczek, ''Implementation of Fourier optics in ultrasonic transducer design,'' Proc. IEEE Ultrason. Symp., Williamsburg, VA **2**, 655–658 (1986).
[13] American Piezo Ceramics Inc., APC-853 Material data sheet. Duck Run Road P.O. Box 180. Mackeyville, Pennsylvania 17750.
[14] M. Saito, J. Yamaguchi, and T. Nakajina, ''Applications of holographic interferometry to mechanical experiments,'' U.S.–Japan Seminar on Holography, Washington, DC, October 1969, published in *Applications of Holography* (Plenum, New York, 1971).
[15] A. M. Glass, ''The photorefractive effect,''Opt. Eng. (Bellingham) **17**, 470–479 (1978).
[16] J. P. Huignard and J. P. Herriau, ''Real time double-exposure interferometry with $Bi_{12}SiO_{20}$ crystal in transverse electro-optic configuration-,''Appl. Opt. **16**, 1807–1809 (1977).
[17] P. Yeh, ''Fundamental limit of the speed of photorefractive effect and its impact on device applications and materials research,''Appl. Opt. **26**, 602–604 (1987).
[18] P. Günter and J. P. Huignard (Editors), *Photorefractive Materials and Their Applications* (Springer-Verlag, Berlin, 1988), Vol. I.
[19] P. Günter and J. P. Huignard (Editors), *Photorefractive Materials and Their Applications* (Springer-Verlag, Berlin, 1988), Vol. II.
[20] J. Zemanek, ''Beam behavior within the nearfield of a vibrating piston,''J. Acoust. Soc. Am. **49**, 181–191 (1971).
[21] G. Harris, ''Transient field of a baffled planar piston having an arbitrary vibration amplitude distribution,''J. Acoust. Soc. Am. **70**, 186–204 (1981).
[22] R. Reibold and F. Holzer, ''Complete mapping of ultrasonic fields from optically measured data in a single cross-section,''Acustica **58**, 11–16 (1985).
[23] J. D. Aindow, A. Markiewicz, and C. Chivers, ''Quantitative investigation of disk ultrasonic sources,''J. Acoust. Soc. Am. **78**, 1519–1529 (1985).
[24] R. L. Roderick, ''The radiation pattern from a rotationally symmetric stress source of a semi-infinite solid,'' Ph.D. thesis, Brown University, 1951.
[25] V. Maki, S. Gianzero, R. Strickland, N. Kepple, and M. Gianzero, ''Dynamically focused transducer applied to the CAST imaging tool,'' SPWLA 32nd Annual Logging Symposium, Midland Texas **2**, 1–24 (1991).
[26] G. S. Kino, *Acoustic Waves: Devices, Imaging, and Analog Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1987).
[27] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics* (Wiley, New York, 1982), 3rd ed.

A nonlinear material model of lead magnesium niobate (PMN)

Jean C. Piquette and Stephen E. Forsythe
Naval Undersea Warfare Center, Underwater Sound Reference Detachment, P.O. Box 568337, Orlando, Florida 32856-8337

(Received 22 December 1995; revised 12 March 1996; accepted 3 September 1996)

A one-dimensional material model of lead magnesium niobate (PMN) is presented. The model includes saturation phenomenology, but excludes hysteresis and dispersion. (Constant temperature is assumed.) It is shown that the strain can be taken to respond as an exactly quadratic function in the electric displacement *D throughout the saturation region* and yet still deduce the full observed response of the material, including flattening of the curve, when the strain is expressed as a function of electric field E. The model developed here is shown to be compatible with experimental measurements previously acquired by other researchers. [S0001-4966(97)02001-8]

PACS numbers: 43.38.Ar, 43.20.Px, 43.30.Yj, 43.58.Vb [SLE]

INTRODUCTION

Lead magnesium niobate (PMN) is currently enjoying considerable interest in the transducer community because this material exhibits strains that are approximately an order of magnitude greater than those of typical piezoceramics for the same electric field strength. We present a new material model useful for determining the transduction behavior of this material. The model is developed by correlation with experiment.

In developing the present material model of PMN, we have carefully reexamined the fundamental laws of electromagnetism. Three properties, useful for determining whether a material model is compatible with electromagnetism, have been identified. The properties are used to examine an existing theory of electrostriction. A revised version of the theory, which is shown to be consistent with the required properties, is proposed. (The revision is based on incorporating a postulated constitutive equation of polarization into the existing theory.) The revised theory incorporates saturation phenomenology, and is demonstrated to be compatible with measurements.

In examining the laws of electromagnetism we have made extensive use of a textbook by Jefimenko,[1] and have adopted the notation used there. Thus the symbol ϵ is used to denote the dimensionless relative permittivity, and the symbol ϵ_0 denotes the permittivity of free space ($\epsilon_0 = 8.85 \times 10^{-12}$ F/m). In this notation, the low-field permittivity of a dielectric material is denoted by the product $\epsilon_0 \epsilon$.

In Sec. I, we present the three fundamental properties which must be obeyed by any material model if it is to be compatible with the laws of electromagnetism. The properties are justified in detail in Sec. II. In Sec. III, we present our theory for the material behavior of PMN. In this section, a specific model of polarization is postulated, and one application to experimental data is presented. Further comparisons of the model with experiments are given in Sec. IV. The model is examined for consistency with the three fundamental properties in Sec. V. A summary is given in Sec. VI.

I. THREE PROPERTIES REQUIRED OF ANY MATERIAL MODEL

In order to be compatible with the laws of electromagnetism, a material model should exhibit the following properties.

(1) The equation $D = \epsilon_0 E + P$ (where D is the electric displacement, E is the electric field, P is the polarization, and ϵ_0 is the permittivity of free space) is fundamental (in fact, this equation can be taken as the *definition* of polarization P), and any material model must be consistent with it.

(2) To the extent that fringing fields can be neglected, an electroded planar slab of material, with electrodes shorted together, should exhibit zero internal E field, even if remanent polarization is nonzero.

(3) To the extent that fringing fields can be neglected (e.g., in models which are one dimensional), the equations of the model should be consistent with the fact that an isolated specimen containing no free charge should exhibit zero D field, both internal and external to the sample, even if remanent polarization is nonzero.

II. JUSTIFICATIONS OF THE THREE PROPERTIES

A. Property 1

As stated within property 1, this equation can be taken to be the definition of polarization (see Ref. 1, p. 246), and therefore even nonlinear models must be consistent with it. In fact, according to Jefimenko,[1] this equation must be obeyed by all dielectric media, even anisotropic materials. Clearly, no further justification for imposing consistency of any material model with this property is required.

B. Property 2

The electric field E is related to the electric potential ϕ through the equation $E = -\nabla\phi$. When the electrodes of a capacitor are shorted together, no difference in potential ϕ exists between the plates and, in view of the symmetry imposed when fringing fields are neglected, no electric field can exist. (Of course, the equation $E = -\nabla\phi$ is not fundamental and, in general, an electric field contribution also arises from

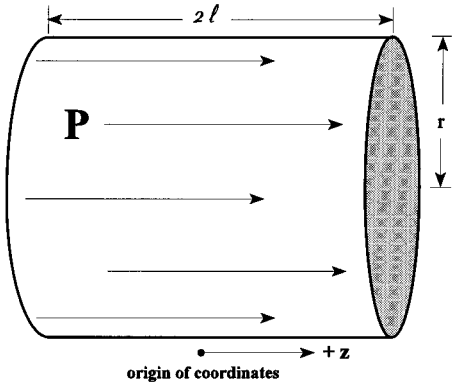


FIG. 1. A permanent electret of fixed polarization P, radius r, and length $2l$.

a time-changing magnetic vector potential A. We ignore this magnetic contribution to the electric field, as is typical at audio frequencies in transducer work.)

C. Property 3

As a guide in understanding the basis of property 3, we will consider the case of a finite cylindrical electret of length $2l$, radius r, and permanent polarization P. See Fig. 1. (An electret is the electrical analog of a magnet; i.e., it is a passive material exhibiting a fixed remanent polarization. See, e.g., Jefimenko,[1] p. 252.) Any material retaining a remanent polarization is an electret. Thus typical piezoceramics that contain a permanent pole are electrets, and the following considerations apply to these materials.

Despite the finite geometry, the static fields both internal and external to the electret can be solved exactly. The axial electric field within the electret is given by

$$E=\frac{P}{\epsilon_0}\left[-1+\frac{(z+l)}{2\sqrt{(z+l)^2+r^2}}-\frac{(z-l)}{2\sqrt{(z-l)^2+r^2}}\right], \tag{1}$$

and external to (and to the right of) the electret is given by

$$E=\frac{P}{2\epsilon_0}\left[\frac{(z+l)}{\sqrt{(z+l)^2+r^2}}-\frac{(z-l)}{\sqrt{(z-l)^2+r^2}}\right]. \tag{2}$$

Equation (2) shows that a nonzero electric field exists external to the electret. This also shows that a nonzero D field exists external to the electret, by virtue of the fundamental electromagnetic vacuum equation $D=\epsilon_0 E$, despite the fact that there is no free electric charge present.[2] However, in the limit as the disk radius $r\to\infty$, the external E field (and hence, the D field) vanishes. (Although this result has been derived using an expression for the axial field, it applies at all points external to the electret, by virtue of the symmetry of the infinite electret that results in the limit under consideration.) Since the D field must be continuous over any boundary which contains no free charge, if there is zero external D field, there must also be zero internal D field. Furthermore, Eq. (1) shows that in the limit as $r\to\infty$, there is a nonzero internal E field given by $E=-P/\epsilon_0$. (This nonzero internal electric field and zero internal D field are consistent with the equation $D=\epsilon_0 E+P$.) This shows that the presence of a nonzero D field in the case of an electret of finite radius, despite the absence of any free charge, is purely a consequence of the influence of fringing fields. Thus any one-dimensional theory must be consistent with property 3.

III. A NEW ONE-DIMENSIONAL MATERIAL MODEL FOR PMN

Lead magnesium niobate (PMN) is an electrostrictive ceramic.[3] Hence, we develop a one-dimensional model useful for analyzing and predicting the behavior of a PMN-based transducer incorporating a plate-shaped actuator by generalizing a phenomenological model of electrostriction proposed by Mason.[4] The generalization involves incorporating saturation phenomenology such as that reported by Rittenmeyer,[3] and by Blackwood and Ealey.[5] However, we ignore hysteresis and dispersion, and constant temperature is assumed. Hence, this model is general in the sense that it applies to conditions of high electric field strength, not in the sense that all the physical phenomena manifested by PMN are accommodated.

A. Generalized model for electrostriction

We develop our generalized model empirically, by making reference to experimental data. In particular, we *postulate* the following equation for accommodating stress-free polarization data acquired for PMN:

$$P=P_0+\frac{\epsilon_0(\epsilon-1)E}{\sqrt{1+aE^2}}, \tag{3}$$

where P_0 is the remanent polarization (including a constant polarization term is also considered in Ref. 6, p. 174), ϵ is the low-field relative dielectric permittivity, and a is a constant that is associated with the departure from constancy of the relative permittivity for high E fields.[7–9] The numerical value of the constant a directly determines the nature of saturation. That is, a large value of a means that a sample saturates quickly (i.e., at relatively modest electric field strengths). If $a=0$, saturation does not occur. For these reasons, we will term a the ''saturation parameter.'' Notice that for $a=0$ and $P_0=0$, Eq. (3) is consistent with elementary linear models of polarization, such as that considered by Jefimenko (see Ref. 1, p. 259). Moreover, the functional dependence exhibited by Eq. (3) is consistent with the measurements of Rittenmeyer (see Ref. 3, Fig. 7), with the exception that hysteresis is excluded. The polarizations depicted in Rittenmeyer's Fig. 7 do not pass through zero value at zero electric field. This suggests the presence of a nonzero remanent polarization P_0 in these measurements, justifying the inclusion of this term in the present model of PMN. We also remark that the need to postulate an empirical function for P in order to accommodate certain saturation phenomenology (such as flattening of the curve) has previously been pointed out by Cady (see Ref. 6, p. 588).

In Fig. 2 is shown an attempt to match our theory to Rittenmeyer's Fig. 7(c). The squares in Fig. 2 are digitized data points taken from Fig. 7(c) of Ref. 3, and the solid-line curve is the theoretical result based on Eq. (3), in which three theoretical parameters have been determined by least-squares fitting of the equation to the digitized data. One of the three

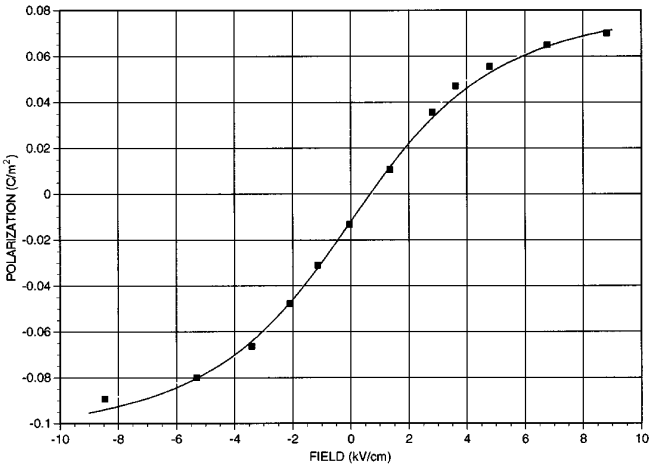


FIG. 2. Theoretical polarization (solid line) compared with digitized data taken from Fig. 7(c) of Rittenmeyer[3] (squares).

adjusted parameters is the combination of constants given by $\epsilon_0(\epsilon-1)$, while the other two parameters are the quantities P_0 and a. The numerical values used to generate the solid-line curve, rounded to three significant figures, are $\epsilon_0(\epsilon-1) = 0.0181$, $P_0 = -0.0120$, and $a = 0.0349$, where the numerical values are expressed in terms of the units of the plot. (The parameters suggest $\epsilon \approx 20500$.)

Next, we deduce the stress-free connection between polarization P and displacement D by combining the fundamental equation of property 1 with Eq. (3). This gives

$$D = \epsilon_0 E + P_0 + \frac{\epsilon_0(\epsilon-1)E}{\sqrt{1+aE^2}}. \tag{4}$$

We emphasize that Eqs. (3) and (4) apply to stress-free conditions. To derive the appropriate generalizations for conditions of nonzero stress, we refer to the equations of the phenomenological barium titanate model of Mason.[4] These equations, expressed to quadratic accuracy in the variables (and restricted to one dimension), can be written in the form

$$E_3 = \alpha_1 D_3 + \alpha_2 D_3^2 + \alpha_3 D_3 T_3, \tag{5}$$

and

$$S_3 = \beta_1 T_3 + \beta_2 D_3^2, \tag{6}$$

where the ''three'' direction represents the thickness, E_3 is the electric field, D_3 is the electric displacement field, S_3 is the strain, and T_3 is the stress. [Equation (5) is the ninth, and Eq. (6) is the third, of Mason's Eqs. 12.23, Ref. 4, pp. 298–300, with all $S_1 = S_2 = 0$, and solving for E_3 and S_3. Mason's equations have also been translated into the notation of Ref. 1.] The quantities α_1, α_2, α_3, β_1, and β_2 are constants resulting from series expansions. This notation for these constant quantities, which differs from the notation originally used by Mason, is introduced here to simplify the expressions.

It is important to note that when we state that Eqs. (5) and (6) are expressed to ''quadratic accuracy,'' we do *not* mean that these equations represent the original series expansions expressed to this accuracy, for Mason has dropped several quadratic-order terms from these series, based on the physical properties of barium titanate. Rather, what is meant here is that these equations are *equivalent to the original Mason barium titanate model*, expressed to this accuracy and restricted to one dimension. By using Eqs. (5) and (6) as the basis of the extension of our theory, we again *postulate* that these equations, as presented by Mason for barium titanate, are also applicable to PMN. This postulate is based simply on the fact that, like barium titanate, PMN is an electrostrictive ceramic.

We seek to modify Eq. (5) so as to be consistent with Eq. (4). To determine how this modification should be done, it is worthwhile to solve Eq. (4) for E. Unfortunately, as expressed in the form of Eq. (4), the desired solution appears as a root of a fourth-degree polynomial. However, a great simplification of this solution is possible for PMN materials, for which the low-field relative permittivity ϵ has a numerical value which is typically on the order of 20 000. In fact, Eq. (4) can be replaced by the approximate equation

$$D = P_0 + \frac{\epsilon_0 \epsilon E}{\sqrt{1+aE^2}}, \tag{7}$$

and produce errors in the calculation of D which are less than 1% for electric field strengths less than $\epsilon/100\ a^{1/2}$. For the parameters used in Rittenmeyer's measurements,[3] this condition translates into an electric-field strength limitation of about 1000 kV/cm. (Rittenmeyer's actual maximum electric field strength was about 10 kV/cm.) For the parameters used in the measurements of Blackwood and Ealey,[5] this condition translates into an electric-field-strength limitation of about 2500 kV/cm. (The actual maximum electric field considered by Blackwood and Ealey was about 16 kV/cm.) Hence, the approximation is extremely accurate for field strengths typically encountered in connection with PMN applications, so we will freely use Eq. (7) in lieu of Eq. (4). The solution of Eq. (7) for electric field E is

$$E_3 = \frac{D_3 - P_0}{\sqrt{\epsilon_0^2 \epsilon^2 - a(D_3 - P_0)^2}}. \tag{8}$$

Although Eq. (7) is an excellent approximation to Eq. (4) for electric field values up to $E_3 = \epsilon/1000a^{1/2}$, in view of the fact that Eq. (8) has been obtained by inverting Eq. (7), it is possible that the resulting equation is more restricted. As can be seen from examining the radical in Eq. (8), this equation clearly must be restricted to parameter values which avoid the zero in the denominator. A careful numerical comparison of the roots of Eq. (8) with the roots of Eq. (4) demonstrates that Eq. (8) gives the electric field accurate to about 1.5% for values of the parameters which satisfy $D_3 - P_0 < 0.99\epsilon_0\epsilon/(a)^{1/2}$. By making the substitution $D_3 - P_0 \to 0.99\epsilon_0\epsilon/(a)^{1/2}$ in Eq. (8), it follows that this restriction on the quantity $D_3 - P_0$ translates into an approximate restriction on the electric field of the form $E_3 < 7/(a)^{1/2}$, which is a significantly stronger restriction on the applicability of Eq. (8) than that deduced for the applicability of Eq. (7). Nonetheless, for the Rittenmeyer[3] polarization data considered here in Fig. 2, this restriction on the applicability of Eq. (8) is equivalent to restricting the electric fields to values less than about 36 kV/cm in magnitude. For the actual $E = 10$

kV/cm maximum electric field used in the Rittenmeyer data, the error in the computed E field obtained by using Eq. (8) is less than 0.03%. For the Blackwood and Ealey[5] experiment, the restriction on the applicability of Eq. (8) is for electric fields less than about 67 kV/cm. The error in the electric field produced by Eq. (8) for the actual maximum electric field strength of about 16 kV/cm used in that experiment is less than 0.02%.

Using Eq. (8) as a guide, we choose to generalize the Mason model of electrostriction so that Eqs. (5) and (6) are replaced with the equations

$$E_3=\frac{D_3-P_0}{\sqrt{\epsilon_0^2\epsilon^2-a(D_3-P_0)^2}}-2\beta_2 T_3 D_3,$$
$$\text{for } D_3-P_0<0.99\epsilon_0\epsilon/\sqrt{a}, \tag{9}$$

and

$$S_3=\beta_1 T_3+\beta_2 D_3^2, \tag{10}$$

where Eq. (6) has simply been reproduced as Eq. (10) for convenience. [The coefficients of the T_3D_3 term of Eq. (9) and the D_3^2 term of Eq. (10) are necessarily related as shown in order that energy be conserved. This is also consistent with the results of Mason. See Ref. 4, pp. 296–297, Eqs. (12.15) and (12.20).] By referring to Eq. (10), it can be seen that our β_1 is a compliance, and our β_2 is essentially equivalent to the electrostriction constant, often denoted by Q in other treatments. Our saturation parameter a is approximately related to the parameter k of Ref. 7 as $a=k^2$, as follows from appropriate expansions. The generalized Eq. (9) can be seen to be compatible with Mason's original equation, denoted here by Eq. (5), by expanding the radical in Eq. (9) and assigning suitable values to the parameters a, ϵ, and P_0, with the exception that Mason's equation excludes the constant term in the expansion. In the case in which $P_0=0$, the present model would predict no quadratic term in D_3 in Eq. (9), so that the parameter α_2 of Eq. (5) would have to be zero in this case, and an extra term in D_3^3 would have to be retained in Mason's expansion to produce the first nonzero nonlinear correction in D_3.

In Mason's original development,[4] the form of Eq. (6) arose from a series expansion, and hence this equation is an approximation to an expression containing an infinite number of additional terms. However, we *postulate* that the quadratic term in D_3 of Eq. (10) is *exact*; that is, we postulate that no additional powers of D are required in this equation. Once again, this postulate is based on experiment.[10]

Mason's original equations, represented here by Eqs. (5) and (6), are based on expansions carried out about the zero values of the variables. However, doing this has the undesirable consequence of causing the resulting model to fail to satisfy property 1 discussed above, at least in the case for which there is nonzero remanent polarization P_0. As we have seen in Rittenmeyer's data,[3] although the remanent polarization of PMN is relatively small, it is not zero, and so the possibility of remanent polarization should be accommodated in the model.

In order to understand the difficulty with samples retaining a remanent polarization, consider a laterally infinite, finite-thickness slab of electrostrictive material having a nonzero remanent polarization P_0 but *containing no free charge*. As described in the previous section, in this case the D field internal to the material is zero, but the E field is $-P_0/\epsilon_0$. Equation (5) shows that a zero D but nonzero E are incompatible with this equation in the case of nonzero P_0. Hence, it is seen that the original Mason model is not strictly compatible with nonzero remanent polarization.[11–13] On the other hand, if $\epsilon=1$ and $a=0$ are substituted into Eq. (9) (these substitutions reflect the fact that the sample is assumed to have a fixed polarization), one obtains $E=-P_0/\epsilon_0$ for $T=D=0$ [compare this result with Eq. (1), evaluated for infinite radius].

At this point, it is worthwhile noting that the present generalization not only modifies the original Mason barium titanate model in a way that renders it compatible with property 1, but also subsumes it as a special case. Consequently the applications considered by Mason, such as the derivation of the linear laws of piezoelectricity from the laws of electrostriction (see Ref. 4, pp. 298–302), also follow from the present theory.

Since the Mason model is subsumed, it follows that the present theory applies to barium titanate as well as PMN. Thus although PMN typically retains only a weak remanent pole P_0, the present theory is not limited in applicability to such cases. If P_0 is a strong remnent, as in the case of barium titanate, the present model could be linearized in the same manner as was done by Mason. A strong remnent would induce a strong residual free charge on the electrodes (i.e., a free charge that persists when zero internal E field is present, i.e., when zero voltage is applied across the electrodes). Under zero external stress, the residual free charge density would be equal in magnitude to P_0 [as can be seen by solving Eq. (9) for D_3 when $E_3=T_3=0$, and recalling that the D field can be considered to arise from free charge when fringing is neglected]. This suggests that a new variable, $\delta=D-P_0$, could be introduced (where δ represents the variable part, and P_0 represents the fixed part of D), and the equations of linear piezoelectricity could be derived in a manner similar to that originally carried out by Mason (Ref. 4, pp. 298–302).

IV. APPLICATION TO DATA OF BLACKWOOD AND EALEY,[5] AND TO DATA OF RITTENMEYER[3]

In what follows, we consider the application of the present model to data available in the literature. The calculations involve fitting to the data by adjusting the constants of the model. It is important to point out that the samples considered by Rittenmeyer[3] and by Blackwood and Ealey[5] differ in composition. Hence, it is not expected that the model parameters P_0, ϵ, and a determined for each of these cases would be similar in numerical value.

For conditions of zero stress ($T_3=0$) and zero remanent polarization ($P_0=0$), solving Eqs. (9) and (10) for the strain produces

$$S_3=\frac{\beta_2(\epsilon_0\epsilon E_3)^2}{1+aE_3^2}. \tag{11}$$

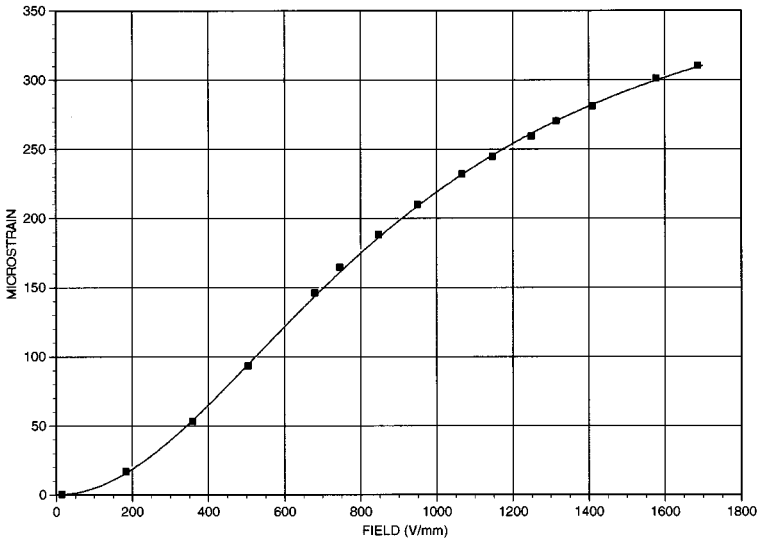


FIG. 3. Theoretical strain at zero stress (solid line) compared with digitized data taken from Fig. 1 of Blackwood and Ealey[5] (squares).

In Fig. 3 is shown an attempt to fit Eq. (11) to the data of Blackwood and Ealey.[5] In this figure, the squares represent digitized data taken from Fig. 1 of Ref. 5, while the solid line is a plot of Eq. (11), in which two parameters have been least-squares adjusted to fit the digitized data. One of the two adjusted parameters is the combination of constants $\beta_2\ (\epsilon_0\epsilon)^2$, and the second parameter is the constant a. The numerical values of the two parameters determined by least-squares adjustment (and rounded to three significant figures) are $\beta_2\ (\epsilon_0\epsilon)^2 = 4.87\times10^{-4}$, and $a = 1.22\times10^{-6}$, in the units of the graph. (The reader should not be troubled by the difference between the values of the saturation parameters a obtained for the Blackwood and Ealey data and that obtained for the Rittenmeyer data, as reported in Sec. III A in connection with the discussion of Fig. 2. Most of this difference is attributable to differences in the units used in each of the original plots, which we have retained to aid the reader in comparing our results with the original publications. When the two values of the saturation parameters are put into the same units, it is found that the Rittenmeyer value is only 2.86 times the Blackwood and Ealey value. This larger value for the Rittenmeyer material simply suggests that saturation occurs more rapidly than for the Blackwood and Ealey material.)

An attempt to match the effective piezoelectric constant reported by Blackwood and Ealey (Ref. 5, Fig. 2), defined as the derivative of S_3 with respect to E_3, is depicted here in Fig. 4. The squares in Fig. 4 are digitized data taken from Fig. 2 of Ref. 5. The solid line is the theoretical result, obtained by differentiating Eq. (11) with respect to E_3. (The same parameter values used to generate the theoretical curve of Fig. 3 were also used to produce the theoretical curve of Fig. 4, but notice that the units of the plots are different.) Although the failure of the curves to overlap near the end of Fig. 4 may represent a failure of the theory, this may also be indicative of some difficulty with the data; note the unusual

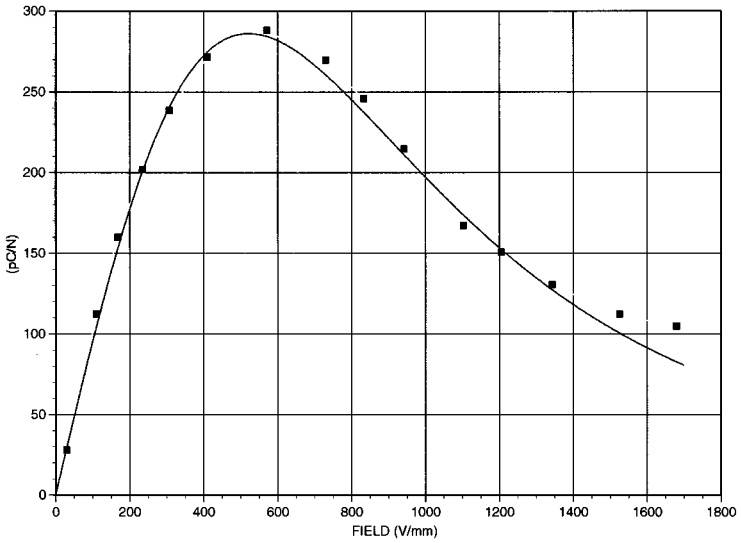


FIG. 4. Theoretical effective piezoelectric constant (solid line) compared with digitized data taken from Fig. 2 of Blackwood and Ealey[5] (squares).

abrupt change in slope near $E=1100$ V/mm in the experimental data. (The reader may wish to check the original curve, published as Fig. 2 in Ref. 5, to more clearly see the abrupt change in slope.) The close agreement of the theoretical curves, generated by adjusting only two parameters of the present theory, with the data reported in Ref. 5 demonstrates the viability of the saturation model incorporated within Eqs. (9) and (10). Since the forms of Eqs. (9) and (10) are based on Eq. (3), which was postulated for consistency with measured *polarization* data, the fact that Eq. (11) is consistent with measured *strain* data is an independent indicator of the validity of the model.

Next, if we assume the quantity aE^2 of Eq. (7) is small relative to 1, we may expand the radical in the denominator, retaining terms up to the quadratic in E, to produce

$$D \approx P_0 + \epsilon_0 \epsilon E - \tfrac{1}{2} a \epsilon_0 \epsilon E^3. \tag{12}$$

For zero remanent polarization P_0, Eq. (12) is consistent with the higher-order field dependence identified by Blackwood and Ealey. [See Ref. 5, Eq. (25). Even the minus sign in the third term on the right-hand side of our Eq. (12) is consistent with the observations of Blackwood and Ealey. See Ref. 5, p. 127; note especially the discussion following their Eq. (27).] Also, if $P_0=0$, we can calculate what Blackwood and Ealey call the *secant permittivity* (formed by the ratio of D to E), which we will denote by $\epsilon_0\epsilon_1$. We obtain

$$\epsilon_0 \epsilon_1 = D/E \approx \epsilon_0 \epsilon - \tfrac{1}{2} a \epsilon_0 \epsilon E^2. \tag{13}$$

Equation (13) suggests that the first-order nonlinear dependence of the permittivity upon electric-field strength is quadratic in character, which is clearly consistent with the observations of Rittenmeyer (see Ref. 3, Figs. 8 and 9). The higher-order tapering off of ϵ_1, which can be seen in Rittenmeyer's Fig. 9 for electric field strengths exceeding 5 kV/cm, is consistent with the behavior for the ratio D/E that would be predicted by Eq. (4) for $P_0=0$. Thus the present theory is consistent with the main features[14] of Rittenmeyer's data. A better definition of secant permittivity, when P_0 is nonzero, would be $\epsilon_0\epsilon_1=(D-P_0)/E$. This definition avoids the artificial singularity in ϵ_1 for nonzero P_0 which results from the definition introduced by Blackwood and Ealey.

Finally, we note that the postulated equation for polarization, introduced here in Eq. (3), produces a more satisfactory theoretical behavior than the behavior deducible for the polarization from the relation between D and E postulated by Blackwood and Ealey, here represented by Eq. (12). [Of course, it is the fundamental equation of property 1 which is used to deduce the polarization from D and E by forming the difference $P=D-\epsilon_0 E$.] In particular, Eq. (3) has the satisfactory theoretical behavior of predicting a limiting value for the polarization as electric field strength E increases without bound, viz., $P \to P_0 + \epsilon_0(\epsilon-1)/(a)^{1/2}$ as $E \to \infty$, whereas the magnitude of the expression deducible from Eq. (12) for the polarization increases without limit as $E \to \infty$. Of course, it would be unreasonable to expect the constitutive relation of Eq. (12) to properly deal with arbitrarily strong electric fields, since it was introduced by Blackwood and Ealey only to address the first-order nonlinear behavior of the material, and not as a complete model of (stress-free, dispersion-free) saturation (at constant temperature) as Eq. (3) is intended to be.

V. DISCUSSION

We now consider whether the present model satisfies the requirements for a good material model as outlined in Sec. I. We consider each requirement in the order originally considered in Sec. I.

A. Property 1

The model guarantees that property 1, expressed in the form of the equation $D=\epsilon_0 E+P$, is satisfied. That is, the model developed here takes the equation of property 1 to be fundamental, and the polarization is therefore simply computed by forming the difference $D-\epsilon_0 E$. This was done in developing the expression for the electric displacement D, given by Eq. (4), from the stress-free polarization, given by Eq. (3). Thus property 1 is trivially satisfied by the present model. [Of course, in the approximation represented by Eq. (9), this fundamental equation is only approximately satisfied.]

B. Property 2

If a nonzero remanent polarization P_0 is present, the fundamental equation of property 1 requires a nonzero value for the displacement field D, even when $E=0$ owing to the fact that the electrodes are shorted together. The presence of a zero E field, but nonzero D field, is compatible with Eq. (9) of the model, *even if the strength of the nonlinear terms is zero*. [Setting $a=0$ and $\beta_2=0$ produces a null nonlinearity in Eq. (9).] This is a nontrivial matter, considering the fact that the Mason model does not satisfy this property. [Note that a zero value of E and a simultaneous nonzero value of D are incompatible with Eq. (5); this fact is especially evident, however, if $\alpha_2=\alpha_3=0$, i.e., if nonlinearity is null.] (The reader should also bear in mind the fact that the external stress T_3 must be permitted to assume arbitrary values.) That is, although one might suppose that specific numerical values for α_2 and α_3 might render Eq. (5) compatible with simultaneous values of $E=0$ and $D\neq 0$, we consider it theoretically unsatisfactory to *require* that higher-order nonlinear parameters, such as α_2 and α_3, be nonzero simply to accommodate the shorting together of electrodes in a sample having a remanent polarization. A physically satisfactory theory ought to be able to accommodate *arbitrarily small* values of α_2 and α_3, while accommodating *arbitrarily large* values of P_0, and, at the same time, should accommodate arbitrary values of external stress T_3.

C. Property 3

The ''isolated specimen'' referred to in property 3 denotes a sample of material having no electrodes (since electrodes represent a source of free charge) or, if electrodes are present, they must carry zero charge. The vanishing of the D field actually must be *imposed* in the present model in order to determine the electric field E and the polarization P. That is, in order to determine E and P in the circumstance in

which no free charge is present, one first imposes the vanishing of the D field on the fundamental equation of property 1, thus obtaining $P=-\epsilon_0 E$. By combining this result with Eq. (3), one can solve the model for both the E and P fields. In the most general case, this would result in a fourth-degree polynomial; as before, for simplicity, we consider the case of null nonlinearity (i.e., $a=\beta_2=0$), although the solution to the full nonlinear problem would not alter the conclusions. The solution for the polarization in this case is $P=P_0/\epsilon$, and the solution for the electric field is $E=-P_0/\epsilon_0\epsilon$. These solutions are clearly compatible with the fundamental equation of property 1 for zero internal D field. Also, since the D field must be continuous across a boundary containing no free charge, the external D field must vanish also in the present circumstance. Thus the present model is compatible with property 3. Once again this is not a trivial matter, because the Mason model, as represented by Eq. (5), is definitely incompatible with simultaneous field values of $D=0$ and $E\neq 0$, whether or not nonlinearity is null.

It is interesting to note that the solution for the polarization in the isolated-specimen case, viz., $P=P_0/\epsilon$, shows that the sample polarization is reduced by a factor of ϵ with respect to the polarization it has when zero internal electric field is imposed, such as when the electrodes are shorted together. This result can be understood by realizing that the polarization model represented by Eq. (3) treats the material as containing an admixture of two distinct types of poles. One type of pole, characterized by P_0, is permanent; i.e., it is independent of electric field. The second type, characterized by low-field relative permittivity ϵ, is variable; i.e., it changes in response to changes in electric field. In the present calculation, a type of self-induction is occurring. That is the remanent pole P_0 creates an electric field that *induces* a polarization within the variably polarizable material. Obviously, in order to avoid a runaway increase in polarization, the induced polarization opposes the remanent pole, thus reducing the overall polarization of the specimen from the value P_0 to the value P_0/ϵ.

VI. SUMMARY AND CONCLUSION

We have carefully reexamined the laws of electromagnetism as they relate to material models. Three fundamental properties, with which any material model must be consistent if it is to be consistent with the laws of electromagnetism, were presented. A new material model of PMN was also described. The new model was shown to be compatible with the three properties, and with two sets of experimental data from the literature.

We conclude by explicitly stating the three postulates incorporated within our new material model:

(1) Equation (3) is a postulated model of polarization which suggests PMN saturation, apart from hysteresis, can be fully characterized by the parameters P_0, ϵ, and a.

(2) PMN can be treated as analogus to barium titanate, in the sense that Mason's model (Ref. 4), as modified here, can be applied to it.

(3) The D^2 term of Eq. (10) is exact, in the sense that no other terms in Mason's original series expansion for strain need be retained.

ACKNOWLEDGMENT

This work was sponsored by the Office of Naval Research, Code 321, through the Transducer Technology Project.

[1] O. D. Jefimenko, *Electricity and Magnetism* (Meredith, New York, 1966).

[2] On the face of it, the presence of a nonzero D field, even when no free charge is present, may appear rather troubling, since free charge is generally considered to be the ''source'' of D fields. However, the concept that free charge is the source of D fields is based on the Maxwell equation $\nabla\cdot D = \rho$, where ρ is the free-charge density. A nonzero D field, even when $\rho=0$, does not necessarily contradict this Maxwell equation, provided that the divergence of the D field in question vanishes. The D fields of the finite cylindrical electret do indeed have zero divergence everywhere (although a demonstration of this fact requires consideration of the full three-dimensional problem, and not simply the axial solution considered here), so no contradiction of this Maxwell equation arises for the finite electret.

[3] K. M. Rittenmeyer, ''Electrostrictive ceramics for underwater transducer applications,'' J. Acoust. Soc. Am. **95**, 849–856 (1994).

[4] W. P. Mason, *Piezoelectric Crystals and Their Application to Ultrasonics* (Van Nostrand, New York, 1950).

[5] G. H. Blackwood and M. A. Ealey, ''Electrostrictive behavior in lead magnesium niobate (PMN) actuators. Part I: materials perspective,'' Smart Mater. Struct. **2**, 124–133 (1993).

[6] W. G. Cady, *Piezoelectricity* (Dover, New York, 1964).

[7] X. D. Zhang and C. A. Rogers, ''A macroscopic phenomenological formulation for coupled electromechanical effects in piezoelectricity,'' J. Intelligent Mater. Syst. Struct. **4**, 307–316 (1993).

[8] C. L. Hom, S. M. Pilgram, N. Shankar, K. Bridger, M. Massuda, and S. R. Winzer, ''Calculation of quasi-static electromechanical coupling coefficients for electrostrictive ceramic materials,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **41**, 542–551 (1994).

[9] Phenomenological models of polarization, such as that embodied in Eq. (3), have also been postulated previously by others. Examples are those presented by Mason,[4] Cady,[6] and more recently by others (see Refs. 7 and 8). Reference 7 proposes a polarization model based on the hyperbolic tangent, and this model is incorporated into a theory in Ref. 8. We also tried a hyperbolic tangent function model, but we found it would not satisfactorily accommodate the data available to us. Consequently, we were forced to seek another functional form, and Eq. (3) is the result. It is worthwhile noting in this regard that the phenomenological model proposed in Ref. 7, and included in Ref. 8, requires the polarization approach the *same* limiting value as $E\rightarrow +\infty$ and as $E\rightarrow -\infty$ (as can be seen by examining Fig. 6 of Ref. 7), while the data presented by Rittenmeyer do not exhibit this property. (See Fig. 7, Ref. 3.) Equation (3) permits accommodation of these observations owing to the behavior of the square root function.

[10] From the atomic point of view, it actually would be more logical to postulate within Eq. (10) that the strain is proportional to the square of the polarization P^2, rather than the square of the electric displacement D^2, since electrostriction is a phenomenon in which this has been demonstrated experimentally. See R. E. Newnham, ''Tunable transducers: Nonlinear phenomena in electroceramics,'' from the National Institute of Standards and Technology Special Publication 304, Chemistry of Electronic Ceramic Materials, 1991. However, it is computationally more convenient to use D^2, and the difference between using D^2 or P^2 in Eq. (10) would not be experimentally detectable until the electric field strength within the material was such that the product $\epsilon_0 E$ becomes comparable to the polarization P. (Note property 1.) For typical PMN materials, the required electric field strength would have to approach 1000 kV/cm, a level that may not be experimentally feasible.

[11] It should not be thought that Mason's model was not intended to be applied to cases involving nonzero remanent polarization, for he applied his model to just such a case in deriving the laws of linear piezoelectricity from the equations of electrostriction. See Ref. 4, pp. 298–303.

[12] We wish to emphasize that we do not consider any of the *applications* which Mason presented for his electrostrictive model of barium titanate to be in error, despite the incompatibility of his model with property 1 discussed here.

[13] Mason also apparently considered the inconsistency of his model of elec-

trostriction with the equation $D = \epsilon_0 E + P$. See Ref. 4, p. 465. Mason concluded that the E field arising from a remanent polarization would be ''soon...neutralized by the surface and volume conductance and *in a short time* the [internal electric field]=0.'' (Emphasis added.) The amount of time required for the internal electric field to vanish owing to the effects of conductance can actually be rather large. (The process obeys an exponential time law. See, e.g., Ref. 1, p. 498, example 15-1.1.) Using typical properties for PZT-5 [see *Physical Acoustics*, edited by W. P. Mason (Academic, New York, 1964), Vol. IA, p. 202] gives a characteristic time of approximately 25 min for the electric field to reduce to $1/e$ of its initial value.

[14] However, if the asymmetry in Rittenmeyer's curves with respect to zero electric field is not an artifact of the methodology used to produce the curves, it would be necessary to modify Eq. (3) such that the electric field E is replaced by $(E - E_0)$, where E_0 is a constant.

Analysis of harmonic distortion in electroacoustic transducers under indirect drive conditions

Charles H. Sherman, John L. Butler, and Alexander L. Butler
Image Acoustics, Inc., 97 Elm Street, Cohasset, Massachusetts 02025

(Received 8 April 1996; accepted for publication 14 August 1996)

Nonlinear transducer equations for the six major types of electroacoustic transducers (electrostatic, variable reluctance, moving coil, piezoelectric, electrostrictive, and magnetostrictive) are formulated in a common way and solved approximately by perturbation analysis for indirect drive, i.e., current drive of electric field transducers and voltage drive of magnetic field transducers. The direct drive analysis, i.e., voltage drive of electric field transducers and current drive of magnetic field transducers, was given in J. Acoust. Soc. Am. **98**, 1596–1611 (1995). Expressions and graphs are given for some of the indirect drive harmonics which apply to all six transducer types. The solutions for direct and indirect drive are compared, and some of the differences are discussed with emphasis on the question of which drive gives lower harmonic distortion. No answer to this question has been found that is applicable to all nonlinear mechanisms in all transducers at all frequencies. However, it is shown that for transducers of all types with low electrical loss operated near resonance direct drive usually gives lower second harmonics. The exceptions occur when mechanical springs are the dominant nonlinearity and the electromechanical coupling is very high; then indirect drive gives lower second harmonics. It is shown how the results, although based on simple lumped parameter models, can be applied in some cases to more complex transducer structures. © *1997 Acoustical Society of America.* [S0001-4966(97)03112-3]

PACS numbers: 43.38.Ar, 43.25.Ts, 43.30.Yj [SLE]

INTRODUCTION

We recently published a method for analysis of harmonic distortion that gives results applicable to all the major types of electroacoustic transducers.[1] However, that method is restricted to voltage drive of electric field transducers and current drive of magnetic field transducers, which will be called direct drive (DD); it could also be called constant force drive. Voltage (or current) drive means that the voltage (or current) has a fixed magnitude and a sinusoidal wave form as frequency is varied. The DD cases require solution only of the mechanical transducer equation where each term represents a force. Voltage drive of magnetic field transducers and current drive of electric field transducers, which will be called indirect drive (ID), require simultaneous solution of both the mechanical and electrical transducer equations where each term of the latter represents a voltage or current. ID corresponds to driving with a frequency-dependent force. This paper develops a method for analyzing harmonic distortion in the ID cases that also gives results applicable to all transducer types.

Nonlinear mechanical and electrical equations will be developed for the six major types of electroacoustic transducers including nonlinear mechanisms in the mechanical equation as in Ref. 1 plus other nonlinear mechanisms in the electrical equation. These equations will be formulated in a way that uses the various physical and mathematical similarities to reduce the six pairs of equations to one pair which can be solved approximately by perturbation analysis. The solutions for the harmonics will be compared with those from Ref. 1 to evaluate the differences between DD and ID.

Since most power amplifiers have low internal impedance approximate voltage drive is the common case in practice. Therefore, DD is common for electric field transducers and ID for magnetic field transducers. Voltage drive and current drive, which represent the extremes of power amplifier internal impedance, cause quite different harmonics in some respects.[2] Study of these differences may suggest possibilities for reduction of harmonic distortion. For example, current drive has been reported to reduce harmonics in moving coil transducers.[3]

I. ELECTROSTATIC TRANSDUCERS

The electrical and mechanical equations for the polarized electrostatic transducer are given by Eqs. (9) and (10) of Ref. 1 (see diagram in Fig. 1 of Ref. 1):

$$I_1 = \frac{d}{dt}(C_0 V_1) + V_0 \frac{dC_0}{dY_1}\dot{Y}_1 + V_1 G_e, \tag{1}$$

$$M\ddot{Y}_1 + R\dot{Y}_1 = -k_0(Y_1)Y_1 - \frac{\epsilon A V_0 V_1}{(l_0+Y_1)^2} - \frac{\epsilon A V_1^2}{2(l_0+Y_1)^2}, \tag{2}$$

where

I_1 = alternating current,

V_0 = bias (polarizing) voltage,

V_1 = alternating voltage,

$C_0 = \epsilon A/(l_0+Y_1)$ = capacitance,

$l_0 = l + Y_0$ = bias separation of plates,

l = separation of plates for $V_0 = V_1 = 0$,

 0001-4966/97/101(1)/297/18/$10.00

TABLE I. The functions in Eqs. (3) and (4) for electrostatic, variable reluctance (without saturation), moving coil and piezoelectric transducers. The functions for electrostrictive, magnetostrictive, and variable reluctance (with saturation) transducers are in the Appendix.

Transducer	α_1	α_2	β	γ	δ_1
Electrostatic	$C_0=\frac{\epsilon A}{l_0+Y}$		$V_0\frac{dC_0}{dY}$	$V_0\frac{dC_0}{dY}$	$\frac{1}{2}\frac{dC_0}{dY}$
Variable reluctance	$L_0=\frac{\mu_0 AN^2}{4(l_0+Y)}$		$I_0\frac{dL_0}{dY}$	$I_0\frac{dL_0}{dY}$	$\frac{1}{2}\frac{dL_0}{dY}$
Moving coil	$L(Y)$		$l_C B(Y)$	$l_C B(Y)$	$-\frac{1}{2}\frac{dL}{dY}$
Piezoelectric	$\frac{A}{(l_0+Y)}\left[\epsilon_1+2e_2\frac{Y}{l_0}\right]$	$\frac{A\epsilon_2}{(l_0+Y)^2}$	$\frac{A}{l_0}\left[e_1+2c_a\frac{Y}{l_0}\right]$	$\frac{A}{(l_0+Y)}\left[e_1+2c_a\frac{Y}{l_0}\right]$	$\frac{Ae_2}{(l_0+Y)^2}$

Y_0 = static displacement of the moveable plate caused by bias voltage,

Y_1 = alternating displacement of the moveable plate,

G_e = electrical conductance,

M = mass of moveable plate plus radiation mass,

R = viscous resistance in transducer plus radiation resistance,

A = area of each plate,

ϵ = permittivity of the material between the plates,

$$k_0(Y_1)=(k_1+2Y_0k_2+3Y_0^2k_3-\epsilon AV_0^2/l_0^3)+(k_2+3Y_0k_3+3\epsilon AV_0^2/2l_0^4)Y_1+(k_3-2\epsilon AV_0^2/l_0^5)Y_1^2+\cdots,$$

$k_1+k_2Y_1+k_3Y_1^2+\cdots$ = nonlinear mechanical spring "constant."

Here, Y will be used as the symbol for displacement to distinguish ID solutions from the DD solutions in Ref. 1 where X was used.

Equation (2) can be written in notation similar to Eq. (1):

$$M\ddot{Y}_1+R\dot{Y}_1=-k_0(Y_1)Y_1+V_1V_0\frac{dC_0}{dY_1}+\frac{1}{2}\frac{dC_0}{dY_1}V_1^2. \tag{2'}$$

The equal coefficients of $\dot{Y}_1$ in Eq. (1) and V_1 in Eq. (2′) indicate the reciprocity that occurs for low-amplitude linear operation. When voltage is specified (DD for the electrostatic transducer) Eq. (2′) has only one unknown, Y_1, which can be determined and put in Eq. (1) to determine I_1. However, when current is specified (ID for the electrostatic transducer) Eqs. (1) and (2′) must be solved simultaneously for the two unknowns V_1 and Y_1. These coupled equations are nonlinear in both Y_1 and V_1; the first cannot be solved for V_1 to substitute in the second, because it involves $\dot{V}_1$, and the second cannot be solved for V_1 to substitute in the first because it involves V_1^2. Similarly, Y_1 cannot be eliminated from the equations. However, this problem can be handled approximately by perturbation analysis.

A more general form of Eqs. (1) and (2′) is

$$I_1=\frac{d}{dt}[\alpha_1(Y_1)V_1+\alpha_2(Y_1)V_1^2+\alpha_3(Y_1)V_1^3]+\beta(Y_1)\dot{Y}_1+V_1G_e, \tag{3}$$

$$M\ddot{Y}_1+R\dot{Y}_1=-k_0(Y_1)Y_1+\gamma(Y_1)V_1+\delta_1(Y_1)V_1^2+\delta_2(Y_1)V_1^3. \tag{4}$$

These equations will be seen to apply to the other electric field transducers and also to magnetic field transducers with V and I interchanged and G_e replaced by R_e, the electrical resistance. For the electrostatic transducer the functions α_1, β, γ, and δ_1 can be identified by comparison with Eqs. (1) and (2′):

$$\alpha_1=C_0,\quad \beta=\gamma=V_0\frac{dC_0}{dY_1},\quad \delta_1=\frac{1}{2}\frac{dC_0}{dY_1}.$$

The functions α_2, α_3, and δ_2 are not required for the model of the electrostatic transducer used here, but will be required for other transducer types. The α, β, γ, and δ functions for the other transducers are derived in the Appendix. The physical significance of the functions will be evident for each transducer. For example, α_1 is the capacitance of the electrostatic transducer, and α_1V_1 is the charge on the plates for small V_1. For other electric field transducers the relation between charge and voltage requires additional terms involving V_1^2, V_1^3, etc. For magnetic field transducers similar remarks apply with flux linkages in place of charge. Table I shows the simplest form of these functions for some of the transducers to illustrate relations between the functions for a given transducer and similarities in each function for different transducers. The complete functions and notation are in the Appendix. The analysis will be continued from Eqs. (3) and (4) for electric field transducers with occasional reminders of the corresponding results for magnetic field transducers.

II. PERTURBATION SOLUTIONS

Expanding the functions in Eqs. (3) and (4) in power series and specifying current drive, where

$$I_1=I_{10}\cos\omega t,$$

gives

TABLE II. Electrostatic transducer coefficients ($C_{00}=\epsilon A/l_0$).

Order	α_1	$\beta=\gamma$	δ_1
0	C_{00}	$-V_0C_{00}/l_0$	$-C_{00}/2l_0$
1	$-C_{00}/l_0$	$2V_0C_{00}/l_0^2$	C_{00}/l_0^2
2	C_{00}/l_0^2	$-3V_0C_{00}/l_0^3$	$-3C_{00}/2l_0^3$

TABLE IV. Moving coil transducer coefficients.

Order	α_1	$\beta=\gamma$	δ_1
0	L_0	l_cB_0	$-1/2L_1$
1	L_1	l_cB_1	$-L_2$
2	L_2	l_cB_2	$-3/2L_3$

$$I_{10}\cos\omega t=(\alpha_{10}+\alpha_{11}Y_1+\alpha_{12}Y_1^2\cdots)\dot{V}_1+(\alpha_{11}+2\alpha_{12}Y_1+\cdots)\dot{Y}_1V_1+2(\alpha_{20}+\alpha_{21}Y_1+\cdots)V_1\dot{V}_1+(\alpha_{21}+2\alpha_{22}Y_1+\cdots)\dot{Y}_1V_1^2+3(\alpha_{30}+\alpha_{31}Y_1+\cdots)V_1^2\dot{V}_1+(\alpha_{31}+2\alpha_{32}Y_1+\cdots)\dot{Y}_1V_1^3+(\beta_0+\beta_1Y_1+\cdots)\dot{Y}_1+V_1G_e, \tag{5}$$

$$M\ddot{Y}_1+R\dot{Y}=-(k_{10}+k_{20}Y_1+\cdots)Y_1+(\gamma_0+\gamma_1Y_1+\cdots)V_1+(\delta_{10}+\delta_{11}Y_1+\cdots)V_1^2+(\delta_{20}+\delta_{21}Y_1+\cdots)V_1^3. \tag{6}$$

The expansion coefficients in Eqs. (5) and (6) are given in Tables II–VI for each transducer type. The following analysis will give expressions for individual harmonic components each of which is proportional to one of the coefficients in these tables or to the coefficients of $k_0(Y_1)$. Since the ID solutions for the displacement depend on both the electrical and mechanical equations they necessarily involve more transducer parameters than the DD solutions. Furthermore, the γ_{nm} parameters defined in Ref. 1 for DD can not be used for ID, but approximate relationships between the ID and DD parameters will be given later.

The perturbation parameter will be defined by

$$\in\equiv Y_{01}/l_0, \tag{7}$$

as in Ref. 1. Y_{01} is the fundamental displacement amplitude in the zeroth-order (linear) approximation, and l_0 has been defined for the electrostatic transducer and is a similar characteristic length for each transducer type. $\in$ will be incorporated into Eqs. (5) and (6) such that the first power of $\in$ is associated with the lowest-order nonlinearities (i.e., quadratic nonlinearities) and $\in^2$ with cubic nonlinearities. Thus Eqs. (5) and (6) become

$$I_{10}\cos\omega t=\left[\alpha_{10}+\frac{\in l_0}{Y_{01}}\alpha_{11}Y_1+\left(\frac{\in l_0}{Y_{01}}\right)^2\alpha_{12}Y_1^2\right]\dot{V}_1+\left[\frac{\in l_0}{Y_{01}}\alpha_{11}+2\left(\frac{\in l_0}{Y_{01}}\right)^2\alpha_{12}Y_1\right]\dot{Y}_1V_1+2\left[\frac{\in l_0}{Y_{01}}\alpha_{20}+\left(\frac{\in l_0}{Y_{01}}\right)^2\alpha_{21}Y_1\right]V_1\dot{V}_1+\left[\left(\frac{\in l_0}{Y_{01}}\right)^2\alpha_{21}\right]\dot{Y}_1V_1^2+3\left(\frac{\in l_0}{X_{01}}\right)^2\alpha_{30}V_1^2\dot{V}_1+\left[\beta_0+\left(\frac{\in l_0}{Y_{01}}\right)\beta_1Y_1+\left(\frac{\in l_0}{Y_{01}}\right)^2\beta_2Y_1^2\right]\dot{Y}_1+V_1G_e, \tag{8}$$

$$M\ddot{Y}_1+R\dot{Y}_1=-\left[k_{10}+k_{20}\frac{\in l_0}{Y_{01}}Y_1+k_{30}\left(\frac{\in l_0}{Y_{01}}\right)^2Y_1^2\right]Y_1+\left[\gamma_0+\frac{\in l_0}{Y_{01}}\gamma_1Y_1+\left(\frac{\in l_0}{Y_{01}}\right)^2\gamma_2Y_1^2\right]V_1+\left[\frac{\in l_0}{Y_{01}}\delta_{10}+\left(\frac{\in l_0}{Y_{01}}\right)^2\delta_{11}Y_1\right]V_1^2+\left(\frac{\in l_0}{Y_{01}}\right)^2\delta_{20}V_1^3, \tag{9}$$

including all quadratic and cubic nonlinearities.

The solutions for the unknowns, V_1 and Y_1, are assumed in the form of perturbation series:

$$V_1(t)=v_0(t)+\in v_1(t)+\in^2v_2(t)+\cdots, \tag{10}$$

$$Y_1(t)=y_0(t)+\in y_1(t)+\in^2y_2(t)+\cdots. \tag{11}$$

Substituting these assumed solutions into Eqs. (8) and (9) and separating terms by powers of $\in$ gives the following pairs of linear equations for v_0,y_0; v_1,y_1; and v_2,y_2:

$$I_{10}\cos\omega t=\alpha_{10}\dot{v}_0+\beta_0\dot{y}_0+v_0G_e, \tag{12a}$$

$$M\ddot{y}_0+R\dot{y}_0=-k_{10}y_0+\gamma_0v_0; \tag{12b}$$

TABLE III. Variable reluctance transducer coefficients (with saturation) ($L_{00}=\mu_0AN^2/4l_0$).

Order	α_1	α_2	β	γ	δ_1	δ_2
0	$L_{00}\left(1-\frac{4I_0s}{l_0^2}\right)$	$-\frac{2L_{00}s}{l_0^2}$	$-\frac{I_0L_{00}}{l_0}\left(1-\frac{6I_0s}{l_0^2}\right)$	$-\frac{I_0L_{00}}{l_0}\left(1-\frac{9I_0s}{2l_0^2}\right)$	$-\frac{L_{00}}{2I_0}\left(1-\frac{9I_0s}{2l_0^2}\right)$	$-\frac{3L_{00}s}{2l_0^3}$
1	$-\frac{L_{00}}{l_0}\left(1-\frac{12I_0s}{l_0^2}\right)$	$\frac{6L_{00}s}{l_0^3}$	$\frac{2I_0L_{00}}{l_0^2}\left(1-\frac{12I_0s}{l_0^2}\right)$	$\frac{2I_0L_{00}}{l_0^2}\left(1-\frac{9I_0s}{l_0^2}\right)$	$\frac{L_{00}}{l_0^2}\left(1-\frac{18I_0s}{l_0^2}\right)$	$\frac{12L_{00}s}{2l_0^4}$
2	$\frac{L_{00}}{l_0^2}\left(1-\frac{24I_0s}{l_0^2}\right)$	$-\frac{12L_{00}s}{l_0^4}$	$-\frac{3I_0L_{00}}{l_0^3}\left(1-\frac{20I_0s}{l_0^2}\right)$	$-\frac{3I_0L_{00}}{l_0^3}\left(1-\frac{15I_0s}{l_0^2}\right)$	$-\frac{3L_{00}}{2l_0^3}\left(1-\frac{30I_0s}{l_0^2}\right)$	$-\frac{30L_{00}}{2l_0^5}$

TABLE V. Piezoelectric transducer coefficients.

Order	α_1	α_2	β	γ	δ_1
0	$\frac{A\epsilon_1}{l_0}$	$\frac{A\epsilon_2}{l_0^2}$	$\frac{Ae_1}{l_0}$	$\frac{Ae_1}{l_0}$	$\frac{e_2A}{l_0^2}$
1	$-\frac{A}{l_0^2}(\epsilon_1-2e_2)$	$-\frac{2A\epsilon_2}{l_0^3}$	$\frac{2Ac_a}{l_0^2}$	$-\frac{A}{l_0^2}(e_1-2c_a)$	$\frac{-2e_2A}{l_0^3}$
2	$\frac{A}{l_0^3}(\epsilon_1-2e_2)$	$\frac{3A\epsilon_2}{l_0^4}$		$\frac{A}{l_0^3}(e_1-2c_a)$	$\frac{3e_2A}{l_0^4}$

$$0=\alpha_{10}\dot{v}_1+\frac{l_0}{Y_{01}}\alpha_{11}(y_0\dot{v}_0+\dot{y}_0v_0)+2\frac{l_0}{Y_{01}}\alpha_{20}v_0\dot{v}_0+\beta_0\dot{y}_1+\frac{l_0}{Y_{01}}\beta_1y_0\dot{y}_0+v_1G_e, \tag{13a}$$

$$M\ddot{y}_1+R\dot{y}_1=-k_{10}y_1-\frac{l_0}{Y_{01}}k_{20}y_0^2+\gamma_0v_1+\frac{l_0}{Y_{01}}\gamma_1y_0v_0+\frac{l_0}{Y_{01}}\delta_{10}v_0^2; \tag{13b}$$

$$0=\alpha_{10}\dot{v}_2+\frac{l_0}{Y_{01}}\alpha_{11}(y_0\dot{v}_1+y_1\dot{v}_0)+\left(\frac{l_0}{Y_{01}}\right)^2\alpha_{12}y_0^2\dot{v}_0+\frac{l_0}{Y_{01}}\alpha_{11}(\dot{y}_0v_1+\dot{y}_1v_0)+2\left(\frac{l_0}{Y_{01}}\right)^2\alpha_{12}y_0\dot{y}_0v_0+2\frac{l_0}{Y_{01}}\alpha_{20}(v_0\dot{v}_1+\dot{v}_0v_1)+2\left(\frac{l_0}{Y_{01}}\right)^2\alpha_{21}y_0\dot{v}_0^2+\left(\frac{l_0}{Y_{01}}\right)^2\alpha_{21}\dot{y}_0v_0^2+3\left(\frac{l_0}{Y_{01}}\right)^2\alpha_{30}v_0^2\dot{v}_0+\beta_0\dot{y}_2+\frac{l_0}{Y_{01}}\beta_1(y_0\dot{y}_1+y_1\dot{y}_0)+\left(\frac{l_0}{Y_{01}}\right)^2\beta_2y_0^2\dot{y}_0+v_2G_e, \tag{14a}$$

$$M\ddot{y}_2+R\dot{y}_2=-k_{10}y_2-\frac{l_0}{Y_{01}}k_{20}2y_0y_1-\left(\frac{l_0}{Y_{01}}\right)^2k_{30}y_0^3+\gamma_0v_2+\frac{l_0\gamma_1}{Y_{01}}(y_0v_1+y_1v_0)+\left(\frac{l_0}{Y_{01}}\right)^2\gamma_2y_0^2v_0+2\left(\frac{l_0}{Y_{01}}\right)^2\delta_{10}v_0v_1+\left(\frac{l_0}{Y_{01}}\right)^2\delta_{11}y_0v_0^2+\left(\frac{l_0}{Y_{01}}\right)^2\delta_{20}v_0^3. \tag{14b}$$

These pairs of equations can be solved in succession starting with the solution of Eqs. (12a), (12b) for y_0 and v_0 which can then be put in Eqs. (13a), (13b) to solve for y_1 and v_1, etc. Eqs. (12a), (12b) can be solved by first solving Eq. (12b) for v_0, then substituting into Eq. (12a) to get the following equation for y_0:

$$\alpha_{10}\dddot{y}_0+\left(\frac{\alpha_{10}\omega_0}{Q_m}+G_e\right)\ddot{y}_0+\left(\alpha_{10}\omega_0^2+\frac{\beta_0\gamma_0}{M}+\frac{G_e\omega_0}{Q_m}\right)\dot{y}_0+\omega_0^2G_ey_0=\frac{\gamma_0I_{10}}{M}\cos\,\omega t, \tag{15}$$

where

$$\omega_0=(k_{10}/M)^{1/2},\quad Q_m=\omega_0M/R,$$

TABLE VI. Electrostrictive and magnetostrictive transducer coefficients. Maintained polarization—electrostrictive.

Order	α_1	α_2	α_3	β	γ	δ_1	δ_2
0	$\frac{A}{l_0}\epsilon'$	$\frac{A}{l_0^2}\epsilon''$	$\frac{A}{l_0^3}\epsilon'''$	$\frac{A}{l_0}e''$	β_0	$\frac{A}{2l_0^2}e'$	$\frac{A}{l_0^3}4e_4E_0$
1	$-\frac{A}{l_0^2}(\epsilon'-e')$	$-\frac{A}{l_0^3}(2\epsilon''-12e_4E_0)$	$-\frac{A}{l_0^4}(3\epsilon'''-4e_4)$	$\frac{A}{l_0^2}(6c_cE_0)$	$\beta_1-\frac{\beta_0}{l_0}$	$\frac{-2\delta_{10}}{l_0}+\frac{A}{l_0^3}3c_c$	$-\frac{3\delta_{20}}{l_0}$
2	$\frac{A}{l_0^3}(\epsilon'-e'+3c_c)$	$\frac{A}{l_0^4}(3\epsilon''-24e_4E_0)$	$\frac{A}{l_0^5}(6\epsilon'''-12e_4)$		$-\frac{\gamma_1}{l_0}$	$\frac{3\delta_{10}}{l_0^2}-\frac{A}{l_0^4}6c_c$	$\frac{6\delta_{20}}{l_0^2}$

$\epsilon'=\epsilon_1+2E_0\epsilon_2+3E_0^2\epsilon_3+(2e_2+12e_4E_0^2+3c_cS_0)S_0$
$\epsilon''=\epsilon_2+3E_0\epsilon_3+12e_4E_0S_0$
$\epsilon'''=\epsilon_3+4e_4S_0$

$e'=2e_2+12e_4E_0^2+6c_cS_0$
$e''=2e_2E_0+4e_4E_0^3+6c_cE_0S_0$

Maintained Polarization—Magnetostrictive
The magnetostrictive coefficients are the same as the electrostrictive coefficients with ϵ_1, ϵ_2, ϵ_3 replaced by μ_1, μ_2, μ_3 and E_0 replaced by H_0.
Remanent Polarization—Electrostrictive
The electrostrictive coefficients are the same as those for piezoelectric transducers in Table V with ϵ_1 replaced by $\epsilon_1+2e_2S_0$ and e_1 replaced by $e_1+2c_aS_0$.
Remanent Polarization—Magnetostrictive
The magnetostrictive coefficients are the same as the electrostrictive coefficients with ϵ_1, ϵ_2 replaced by μ_1, μ_2.

are the angular frequency of velocity resonance and the mechanical quality factor for low-amplitude DD. In Woollett's notation[4] ω_0 and Q_m are the resonance frequency and the mechanical quality factor at constant voltage (ω_r^E and Q_m^E) for electric field transducers and the resonance frequency and mechanical quality factor at constant current (ω_r^I and Q_m^I) for magnetic field transducers.

The solution of Eq. (15) can be written

$$y_0 = Y_{01}\cos(\omega t+\theta_1), \tag{16}$$

where

$$Y_{01}=\frac{\gamma_0 I_{10}}{\omega_0^2 M z_m(\nu)|Y_{em}(\omega)|}, \quad \text{electric field transducers,}$$

$$Y_{01}=\frac{\gamma_0 V_{10}}{\omega_0^2 M z_m(\nu)|Z_{em}(\omega)|}, \quad \text{magnetic field transducers,}$$

$$z_m(N\nu)=\frac{N\omega}{\omega_0^2 M}|Z_m(N\omega)|=\left[(1-N^2\nu^2)^2+\left(\frac{N\nu}{Q_m}\right)^2\right]^{1/2}$$

$$=\text{normalized mechanical impedance,}$$

$$Y_{em}(N\omega)=G_e+jN\omega\alpha_{10}+\frac{\beta_0\gamma_0}{Z_m(N\omega)}$$

$$=\text{electrical admittance,}$$

$$Z_{em}(N\omega)=R_e+jN\omega\alpha_{10}+\frac{\beta_0\gamma_0}{Z_m(N\omega)}$$

$$=\text{electrical impedance,}\quad N=1,2,3,\ldots,$$

and

$$\nu=\omega/\omega_0.$$

The last term in the electrical admittance and impedance is the motional contribution. The phase angle in Eq. (16) is

$$\tan\theta_N=\frac{N\nu[(\omega_0\alpha_{10}/G_e)(1-N^2\nu^2)+(1/Q_m)(1+(\beta_0\gamma_0/RG_e))]}{N^2\nu^2(\omega_0\alpha_{10}/G_eQ_m+1)-1}.$$

The linear solution for the voltage for electric field transducers is obtained by putting the solution for y_0 into Eq. (12b) giving

$$v_0 = V_{01}\cos(\omega t+\theta_1-\phi_1), \tag{17a}$$

$$V_{01}=\frac{I_{10}}{|Y_{em}(\omega)|},$$

$$\tan\phi_N=\frac{-N\nu}{Q_m(1-N^2\nu^2)}.$$

For magnetic field transducers the linear solution for the current is

$$i_0 = I_{01}\cos(\omega t+\theta_1-\phi_1), \tag{17b}$$

where

$$I_{01}=\frac{V_{10}}{|Z_{em}(\omega)|}.$$

To compare with DD the linear solution for the DD displacement is[1]

$$x_0 = X_{01}\cos(\omega t+\phi_1),$$

where

$$X_{01}=\frac{\gamma_{01}}{\omega_0^2 z_m(\nu)},$$

and γ_{01} (which is defined for each transducer in Tables I–III of Ref. 1) is approximately related to γ_0 by [see Eq. (27b)]

$$M\gamma_{01}/\gamma_0 \simeq V_{10}, \quad \text{electric field transducers,}$$

$$M\gamma_{01}/\gamma_0 \simeq I_{10}, \quad \text{magnetic field transducers.}$$

Thus the relationship between the linear amplitudes for ID and DD is

$$Y_{01}\simeq\frac{X_{01}}{V_{10}}\frac{I_{10}}{|Y_{em}(\omega)|}, \quad \text{electric field transducers,} \tag{17c}$$

$$Y_{01}\simeq\frac{X_{01}}{I_{10}}\frac{V_{10}}{|Z_{em}(\omega)|}, \quad \text{magnetic field transducers.} \tag{17d}$$

This relationship will facilitate comparison of harmonic amplitudes for ID and DD. It shows that Y_{01} is equal to X_{01} with the specified V_{10} or I_{10} of DD replaced by the frequency-dependent voltage, $I_{10}/|Y_{em}(\omega)|$, or current, $V_{10}/|Z_{em}(\omega)|$, respectively. This makes the frequency dependence for ID and DD different, because Y_{em} (or Z_{em}) becomes large near resonance where Z_m is small. Thus the resonant peak is smoothed in ID compared to DD.

The first step in finding harmonics for ID is solving Eqs. (13a), (13b), for y_1 and v_1 using the solutions above for y_0 and v_0. This can be done by solving Eq. (13b) for v_1 and substituting into Eq. (13a) to obtain the following equation for y_1:

$$\alpha_{10}\dddot{y}_1+\left(\frac{\alpha_{10}\omega_0}{Q_m}+G_e\right)\ddot{y}_1+\left(\alpha_{10}\omega_0^2+\frac{\beta_0\gamma_0}{M}+\frac{G_e\omega_0}{Q_m}\right)\dot{y}_1+\omega_0^2G_ey_1$$
$$=\frac{l_0\alpha_{10}}{MY_{01}}\left\{-\left(2k_{20}+\frac{\beta_1\gamma_0}{\alpha_{10}}\right)y_0\dot{y}_0-\frac{G_ek_{20}}{\alpha_{10}}y_0^2+\frac{G_e\gamma_1}{\alpha_{10}}y_0v_0+\left(\gamma_1-\gamma_0\frac{\alpha_{11}}{\alpha_{10}}\right)(y_0\dot{v}_0+\dot{y}_0v_0)\right.$$
$$\left.+2\left(\delta_{10}-\gamma_0\frac{\alpha_{20}}{\alpha_{10}}\right)v_0\dot{v}_0+\frac{G_e\delta_{10}}{\alpha_{10}}v_0^2\right\}. \tag{18}$$

When the solutions for y_0 and v_0 in Eqs. (16) and (17a) are used Eq. (18) becomes

$$\alpha_{10}\dddot{y}_1+\left(\frac{\alpha_{10}\omega_0}{Q_m}+G_e\right)\ddot{y}_1+\left(\alpha_{10}\omega_0^2+\frac{\beta_0\gamma_0}{M}+\frac{G_e\omega_0}{Q_m}\right)\dot{y}_1+\omega_0^2G_ey_1$$
$$=\frac{l_0\alpha_{10}}{MY_{01}}\left\{\left(2k_{20}+\frac{\beta_1\gamma_0}{\alpha_{10}}\right)\frac{1}{2}\omega Y_{01}^2\sin(2\omega t+2\theta_1)-\frac{G_ek_{20}}{2\alpha_{10}}Y_{01}^2[1+\cos(2\omega t+2\theta_1)]\right.$$
$$+\frac{G_e\gamma_1}{2\alpha_{10}}Y_{01}V_{01}[\cos\phi_1+\cos(2\omega t+2\theta_1-\phi_1)]-\left(\gamma_1-\gamma_0\frac{\alpha_{11}}{\alpha_{10}}\right)\omega Y_{01}V_{01}\sin(2\omega t+2\theta_1-\phi_1)$$
$$\left.-\left(\delta_{10}-\gamma_0\frac{\alpha_{20}}{\alpha_{10}}\right)\omega V_{01}^2\sin(2\omega t+2\theta_1-2\phi_1)+\frac{G_e\delta_{10}}{2\alpha_{10}}V_{01}^2[1+\cos(2\omega t+2\theta_1-2\phi_1)]\right\}. \tag{18′}$$

Three of the drive terms in Eq. (18′) are independent of time and give the following static displacement components, i.e., zeroth-order harmonics:

$$Y_{200}=-\frac{k_{20}}{2k_{10}}Y_{01}^2, \tag{19a}$$

$$Y_{110}=\frac{\gamma_1}{2k_{10}}Y_{01}V_{01}\cos\phi_1, \tag{19b}$$

$$Y_{000}=\frac{\delta_{10}}{2k_{10}}V_{01}^2, \tag{19c}$$

where n,m in the subscripts on Y_{nmN} designates the type of nonlinear term which causes the displacement component and N designates the harmonic order. For example, $n=2$, $m=0$ indicates that Y_{200} is caused by a nonlinear spring with spring parameter k_{20}. This is consistent with the Ref. 1 notation where the X_{nmN} component is caused by the term $\gamma_{nm}X^n\cos m\omega t$ in the DD equation of motion.

The second harmonic drive terms in Eq. (18′) can be rearranged into six terms such that each is proportional to one of the six independent nonlinear parameters k_{20}, γ_1, δ_{10}, β_1, α_{11}, and α_{20} as follows:

$$-\frac{\alpha_{10}}{M}k_{20}Y_{01}^2\left[-\omega\sin(2\omega t+2\theta_1)+\frac{G_e}{2\alpha_{10}}\cos(2\omega t+2\theta_1)\right]$$
$$=-k_{20}Y_{01}^2\frac{|Y_e(2\omega)|}{2M}\cos(2\omega t+2\theta_1+\psi_2),$$

$$\frac{\alpha_{10}}{M}\gamma_1Y_{01}V_{01}\left[-\omega\sin(2\omega t+2\theta_1-\phi_1)+\frac{G_e}{2\alpha_{10}}\cos(2\omega t+2\theta_1-\phi_1)\right]$$
$$=\gamma_1Y_{01}V_{01}\frac{|Y_e(2\omega)|}{2M}\cos(2\omega t+2\theta_1-\phi_1+\psi_2),$$

$$\frac{\alpha_{10}}{M}\delta_{10}V_{01}^2\left[-\omega\sin(2\omega t+2\theta_1-2\phi_1)+\frac{G_e}{2\alpha_{10}}\cos(2\omega t+2\theta_1-2\phi_1)\right]$$
$$=\delta_{10}V_{01}^2\frac{|Y_e(2\omega)|}{2M}\cos(2\omega t+2\theta_1-2\phi_1+\psi_2),$$

$$\frac{\beta_1\gamma_0}{2M}\omega Y_{01}^2\sin(2\omega t+2\theta_1),$$

$$\frac{\alpha_{11}\gamma_0}{M}\omega Y_{01}V_{01}\sin(2\omega t+2\theta_1-\phi_1),$$

and

$$\frac{\alpha_{20}\gamma_0}{M}\omega V_{01}^2\sin(2\omega t+2\theta_1-2\phi_1),$$

where

$$\tan\psi_N=\frac{N\omega\alpha_{10}}{G_e},$$

and

$$|Y_e(N\omega)|=[G_e^2+(N\omega\alpha_{10})^2]^{1/2}$$

is the magnitude of the blocked electrical admittance. The solutions for the six second harmonic components resulting from these six drive terms are

$$Y^M_{202}=\frac{-k_{20}Y^2_{01}|Y_e(2\omega)|}{2\omega_0^2 Mz_m(2\nu)|Y_{em}(2\omega)|}\cos(2\omega t+2\theta_1+\psi_2+\theta_2), \tag{20a}$$

$$Y^M_{112}=\frac{\gamma_1 Y_{01}V_{01}|Y_e(2\omega)|}{2\omega_0^2 Mz_m(2\nu)|Y_{em}(2\omega)|}\times\cos(2\omega t+2\theta_1-\phi_1+\psi_2+\theta_2), \tag{20b}$$

$$Y^M_{022}=\frac{\delta_{10}V^2_{01}|Y_e(2\omega)|}{2\omega_0^2 Mz_m(2\nu)|Y_{em}(2\omega)|}\times\cos(2\omega t+2\theta_1-2\phi_1+\psi_2+\theta_2), \tag{20c}$$

$$Y^E_{202}=\frac{\beta_1\gamma_0\omega Y^2_{01}}{2\omega_0^2 Mz_m(2\nu)|Y_{em}(2\omega)|}\sin(2\omega t+2\theta_1+\theta_2), \tag{20d}$$

$$Y^E_{112}=\frac{\alpha_{11}\gamma_0\omega Y_{01}V_{01}}{\omega_0^2 Mz_m(2\nu)|Y_{em}(2\omega)|}\sin(2\omega t+2\theta_1-\phi_1+\theta_2), \tag{20e}$$

$$Y^E_{022}=\frac{\alpha_{20}\gamma_0\omega V^2_{01}}{\omega_0^2 Mz_m(2\nu)|Y_{em}(2\omega)|}\sin(2\omega t+2\theta_1-2\phi_1+\theta_2). \tag{20f}$$

The superscripts M (or E) on Y_{nmN} indicate whether the ID displacement component depends on a nonlinear parameter from the mechanical (or the electrical) equation. The n,m values on Y^E_{nmN} are chosen to correspond to those for Y^M_{nmN} as regards dependence on Y_{01} and V_{01}; e.g., both Y^M_{202} and Y^E_{202} depend on Y^2_{01}. However, these two components are not caused by the same type of nonlinear mechanism; Y^M_{202} depends on k_{20} (nonlinear spring) while Y^E_{202} depends on β_1 (nonlinear coupling).

This completes the first-order solution for the ID displacement. For all transducer types it consists of three static components, given in Eq. (19), and six second harmonic components, given in Eq. (20). If the harmonics in the unspecified electrical variable are of interest they can be found by putting the solutions for y_0, y_1, and v_0 (or i_0) into Eq. (13b) to get v_1 (or i_1).

The second-order solution is obtained by solving Eq. (14b) for v_2, substituting into Eq. (14a) and using the zeroth- and first-order solutions to get an equation for y_2. This equation is the same on the left-hand side as Eq. (18a), but the right-hand side has more drive terms. The solutions consist of third harmonics and corrections to the fundamental of which only two examples will be given. The third harmonic drive terms proportional to Y^3_{01} in the equation for y_2 are

$$-\frac{\alpha_{10}}{M}k_{30}Y^3_{01}\left[\frac{3}{4}\omega\sin(3\omega t+3\theta_1)-\frac{1}{4}\frac{G_e}{\alpha_{10}}\cos(3\omega t+3\theta_1)\right]=-k_{30}Y^3_{01}\frac{|Y_e(3\omega)|}{4M}\cos(3\omega t+3\theta_1+\psi_3)$$

and

$$\frac{\beta_2\gamma_0\omega Y^3_{01}}{4M}\sin(3\omega t+3\theta_1).$$

These drive terms have the solutions

$$Y^M_{303}=\frac{-k_{30}Y^3_{01}|Y_e(3\omega)|}{4\omega_0^2 Mz_m(3\nu)|Y_{em}(3\omega)|}\times\cos(3\omega t+3\theta_1+\psi_3+\theta_3), \tag{21a}$$

$$Y^E_{303}=\frac{\beta_2\gamma_0\omega Y^3_{01}}{4\omega_0^2 Mz_m(3\nu)|Y_{em}(3\omega)|}\sin(3\omega t+3\theta_1+\theta_3). \tag{21b}$$

The complete second-order solution includes many other third harmonic and fundamental components which can be found in the same way. Some of these arise from interactions between nonlinear mechanisms and are proportional to products of nonlinear parameters as was found in the DD second order solution.[1] The solutions in Eqs. (19)–(21) are for electric field transducers; the corresponding results for magnetic field transducers are obtained by replacing I_{10} by V_{10}, Y_{em} by Z_{em}, Y_e by Z_e, and G_e by R_e.

The expressions for the ID harmonic components all involve the quantity $z_m(N\nu)|Y_{em}(N\omega)|$ which can be expressed in the dimensionless form:

$$w(N\nu)\equiv\frac{1}{G_e}z_m(N\nu)|Y_{em}(N\omega)|=\left\{\left(\frac{N\nu}{Q_{me}}\right)^2[1+Q_mQ_e(1-N^2\nu^2)]^2+\left[N^2\nu^2\left(1+\frac{Q_e}{Q_{me}}\right)-1\right]^2\right\}^{1/2}, \tag{22}$$

where

$$Q_{me}=\frac{\omega_0 M}{R+\beta_0\gamma_0/G_e}=\frac{Q_m}{1+\beta_0\gamma_0/RG_e}, \tag{23}$$

is a mechanical quality factor which includes the reflected electrical loss in addition to the mechanical resistance. Q_{me} plays a role in ID similar to Q_m in DD, and, since $Q_{me}<Q_m$, the ID resonant peaks are broader than those for DD. In Eq. (22) Q_e is an electrical quality factor defined as the ratio of the susceptance to the conductance associated with $Y_{em}(\omega_0)$:[4]

$$Q_e=\frac{\omega_0\alpha_{10}}{G_e+\beta_0\gamma_0/R}; \tag{24}$$

thus Q_e includes the reflected mechanical loss.

Using $w(N\nu)$ and $z_m(N\nu)$ the forgoing results for the harmonic component amplitudes can be expressed in terms of the general transducer parameters Q_m, Q_{me}, and Q_e as follows:

$$Y_{01}=\frac{\gamma_0 I_{10}}{\omega_0^2 M G_e w(\nu)},$$

$$|Y^M_{202}|=\frac{k_{20}\gamma_0^2 I_{10}^2[1+(2\nu(Q_eQ_m/Q_{me}))^2]^{1/2}}{2\omega_0^6 M^3 G_e^2 w^2(\nu)w(2\nu)},$$

$$|Y^M_{112}|=\frac{\gamma_1\gamma_0 I_{10}^2[1+(2\nu(Q_eQ_m/Q_{me}))^2]^{1/2}z_m(\nu)}{2\omega_0^4 M^2 G_e^2 w^2(\nu)w(2\nu)},$$

$$|Y^M_{022}|=\frac{\delta_{10} I_{10}^2[1+(2\nu(Q_eQ_m/Q_{me}))^2]^{1/2}z_m^2(\nu)}{2\omega_0^2 M G_e^2 w^2(\nu)w(2\nu)},$$

$$|Y^E_{202}|=\frac{\beta_1\gamma_0^3 I_{10}^2\nu}{2\omega_0^5 M^3 G_e^3 w^2(\nu)w(2\nu)}, \tag{25}$$

$$|Y^E_{112}|=\frac{\alpha_{11}\gamma_0^2 I_{10}^2\nu z_m(\nu)}{\omega_0^3 M^2 G_e^3 w^2(\nu)w(2\nu)},$$

$$|Y^E_{022}|=\frac{\alpha_{20}\gamma_0 I_{10}^2\nu z_m^2(\nu)}{\omega_0^3 M^2 G_e^3 w^2(\nu)w(2\nu)},$$

$$|Y^M_{303}|=\frac{k_{30}\gamma_0^3 I_{10}^3[1+(3\nu(Q_eQ_m/Q_{me}))^2]^{1/2}}{4\omega_0^8 M^4 G_e^3 w^3(\nu)w(3\nu)},$$

$$|Y^E_{303}|=\frac{\beta_2\gamma_0^4 I_{10}^3\nu}{4\omega_0^7 M^4 G_e^4 w^3(\nu)w(3\nu)}.$$

It will be more meaningful to plot the frequency dependence of the harmonic components after converting the displacement amplitudes to harmonic sound-pressure amplitudes assuming that the radiating surface of each transducer is a circular piston in an infinite rigid baffle and that the piston is small compared to the wavelength for frequencies up to the third harmonic. In that case the relation between displacement of the transducer surface and pressure at a distance of 1 m is

$$|p_{nmN}|=\tfrac{1}{2}(N\omega)^2\rho a^2|Y_{nmN}|,$$

where ρ is the density of the medium and a is the radius of the piston. The quantities plotted in Figs. 1–3 as a function of ν are normalized pressure amplitudes at 1 m defined as follows:

$$P_{01}=\frac{2MG_e|p_{01}|}{\rho a^2\gamma_0 I_{10}}=\frac{\nu^2}{w(\nu)},$$

$$P^M_{202}=\frac{4\omega_0^4 M^3 G_e^2|p^M_{202}|}{\rho a^2 k_{20}\gamma_0^2 I_{10}^2}=\frac{(2\nu)^2[1+(2\nu(Q_eQ_m/Q_{me}))^2]^{1/2}}{w^2(\nu)w(2\nu)},$$

$$P^M_{112}=\frac{4\omega_0^2 M^2 G_e^2|p^M_{112}|}{\rho a^2\gamma_1\gamma_0 I_{10}^2}=\frac{(2\nu)^2[1+(2\nu(Q_eQ_m/Q_{me}))^2]^{1/2}z_m(\nu)}{w^2(\nu)w(2\nu)},$$

$$P^M_{022}=\frac{4MG_e^2|p^M_{022}|}{\rho a^2\delta_{10} I_{10}^2}=\frac{(2\nu)^2[1+(2\nu(Q_eQ_m/Q_{me}))^2]^{1/2}z_m^2(\nu)}{w^2(\nu)w(2\nu)}, \tag{26}$$

$$P^E_{202}=\frac{4\omega_0^3 M^3 G_e^3|p^E_{202}|}{\rho a^2\beta_1\gamma_0^3 I_{10}^2}=\frac{4\nu^3}{w^2(\nu)w(2\nu)},$$

$$P^E_{112}=\frac{2\omega_0 M^2 G_e^3|p^E_{112}|}{\rho a^2\alpha_{11}\gamma_0^2 I_{10}^2}=\frac{4\nu^3 z_m(\nu)}{w^2(\nu)w(2\nu)},$$

$$P^E_{022}=\frac{2MG_e^3|p^E_{022}|}{\rho a^2\omega_0\alpha_{20}\gamma_0 I_{10}^2}=\frac{4\nu^3 z_m^2(\nu)}{w^2(\nu)w(2\nu)},$$

$$P^M_{303}=\frac{8\omega_0^6 M^4 G_e^3|p^M_{303}|}{\rho a^2 k_{30}\gamma_0^3 I_{10}^3}=\frac{(3\nu)^2[1+(3\nu(Q_eQ_m/Q_{me}))^2]^{1/2}}{w^3(\nu)w(3\nu)},$$

$$P^E_{303}=\frac{8\omega_0^5 M^4 G_e^4|p^E_{303}|}{\rho a^2\beta_2\gamma_0^4 I_{10}^3}=\frac{9\nu^3}{w^3(\nu)w(3\nu)}.$$

The curves in Figs. 1–3 show the frequency dependence of the fundamental pressure and some of the harmonics. Figure 1 is arranged for easy comparison with Fig. 6 of Ref. 1, but only the shapes can be compared since the magnitudes depend on different parameters. Note in Fig. 1(a) and (e) that the fundamental is independent of Q_m at $\nu=1$ in contrast to DD where Q_m has the most effect at $\nu=1$ [see Fig. 6(a) of Ref. 1]. Note also that the ID fundamental decreases as $Q_mQ_e/\nu Q_{me}=\omega_0\alpha_{10}/\nu G_e$ for high ν while the DD fundamental is independent of ν and Q_m for high ν.

The harmonic amplitudes relative to the fundamental are usually of most interest, but since such ratios depend on a nonlinear parameter and also on the drive level they are dif-

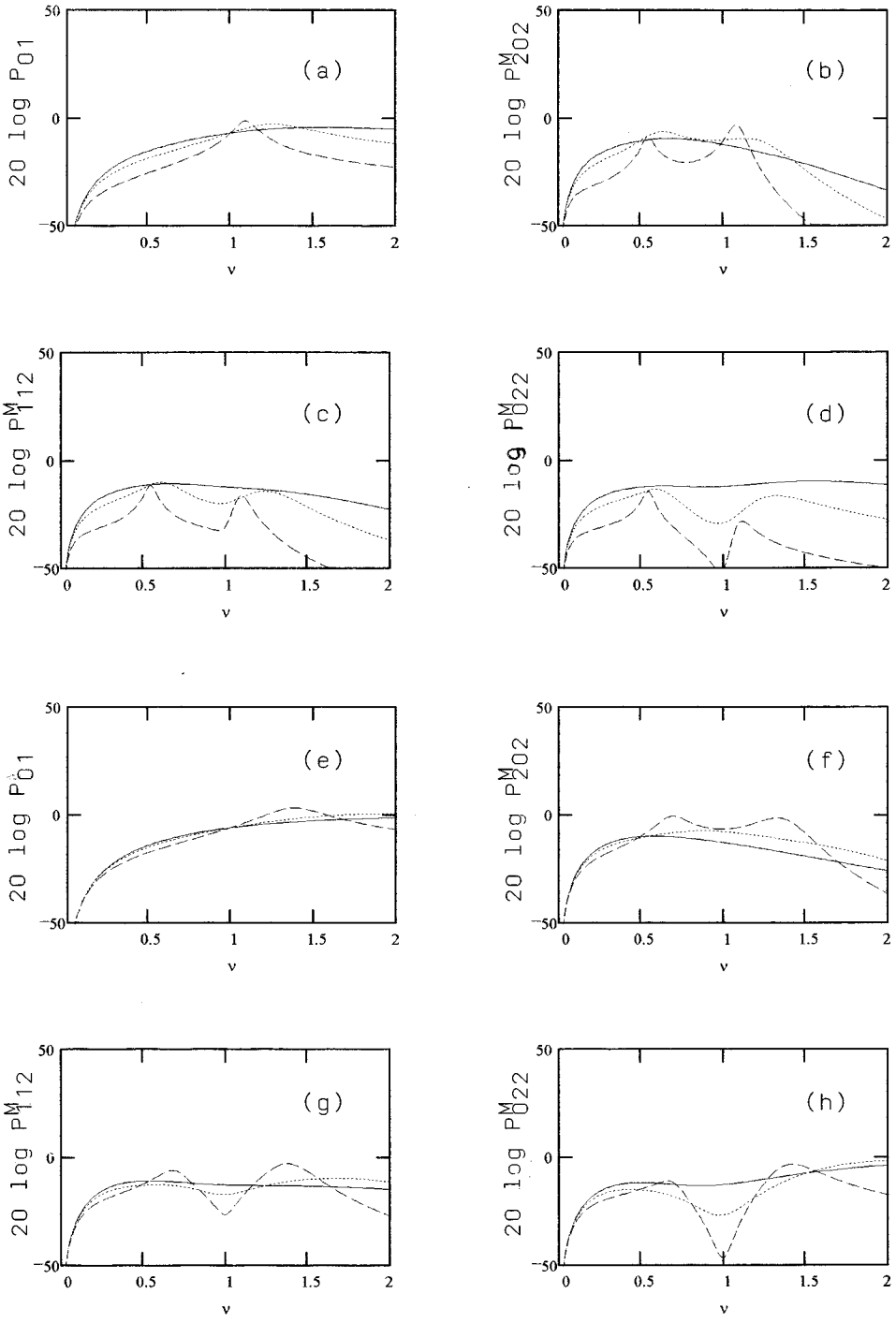


FIG. 1. Normalized pressure P (in dB) as a function of the dimensionless frequency ν for (a) the linear fundamental and second harmonics caused by nonlinear parameters in the mechanical equation; (b) k_{20}, (c) γ_1, (d) δ_{10} for $Q_{me}=0.5$, $Q_e=0.5$, and $Q_m=1$ (———), 3 (.....) and 10 (---); (e)–(h), same except for $Q_e=0.1$.

ficult to illustrate in a comprehensive way. Instead an example will be given which uses the curves in Figs. 1 and 3. Consider the moving coil transducer in Ref. 2 where springs are the dominant nonlinearity, $Q_m=16.7$, $Q_{me}=0.49$, and $Q_e=0.002$. Fitting calculations to measurements of pressure and current harmonics showed that $k_{10}=2.57\times10^5$ N/m, $k_{20}=5.96\times10^6$ N/m^2 and $k_{30}=1.32\times10^8$ N/m^3. The expressions for the pressure components in Eq. (26), using the parameter values in Ref. 2, give for ID:

$$\left|\frac{p_{202}^M}{p_{01}}\right|=\frac{k_{20}B_0l_cV_{10}P_{202}^M(\nu)}{2k_{10}^2R_eP_{01}(\nu)}=1.4\times10^{-3}V_{10}\,\frac{P_{202}^M(\nu)}{P_{01}(\nu)},$$

$$\left|\frac{p_{303}^M}{p_{01}}\right|=\frac{k_{30}(B_0l_cV_{10})^2P_{303}^M(\nu)}{4k_{10}^3R_e^2P_{01}(\nu)}$$

$$=1.8\times10^{-6}V_{10}^2\,\frac{P_{303}^M(\nu)}{P_{01}(\nu)},$$

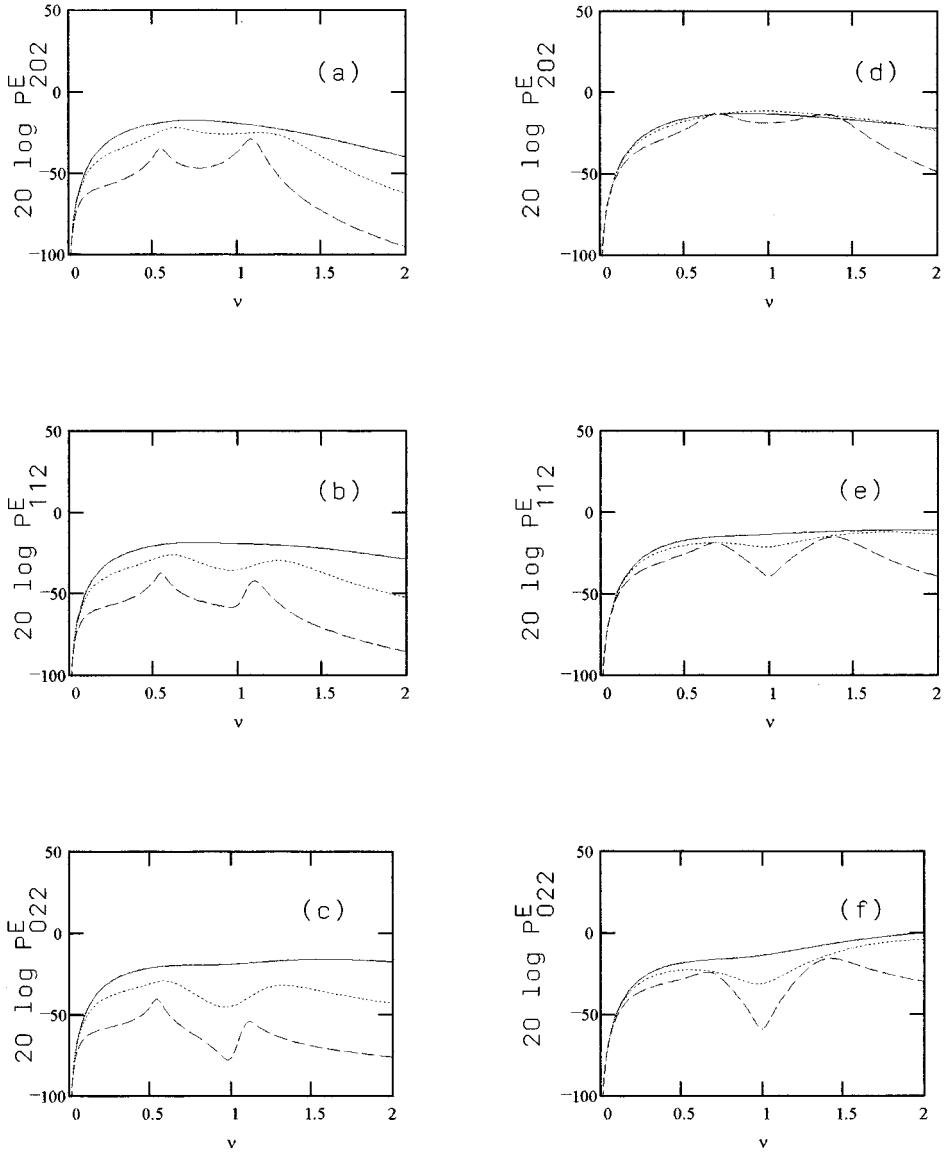


FIG. 2. Normalized pressure P (in dB) as a function of the dimensionless frequency ν for second harmonics caused by nonlinear parameters in the electrical equation; (a) β_1, (b) α_{11}, (c) α_{20} for $Q_{me}=0.5$, $Q_e=0.5$, and $Q_m=1$ (———), 3 (.....) and 10 (---), (d)–(f), same except for $Q_e=0.1$.

and for DD using Ref. 1 and Eq. (27),

$$\left|\frac{p_{202}}{p_{01}}\right| = \frac{k_{20}B_0 l_c I_{10} P_{202}(\nu)}{2k_{10}^2 P_{01}(\nu)} = 7\times 10^{-3} I_{10}\,\frac{P_{202}(\nu)}{P_{01}(\nu)},$$

$$\left|\frac{p_{303}}{p_{01}}\right| = \frac{k_{30}(B_0 l_c I_{10})^2 P_{303}(\nu)}{4k_{10}^3 P_{01}(\nu)} = 4.7\times 10^{-5} I_{10}^2\,\frac{P_{303}(\nu)}{P_{01}(\nu)}.$$

When these expressions are converted to dB Figs. 1(e), (f) and 3(c) and Figs. 6 and 7 of Ref. 1 can be used although Q_m and Q_e are not well approximated for this transducer in these figures. For example, $20\log|p^M_{202}/p_{01}| = -57 + 20\log V_{10} + 20\log P^M_{202}(\nu) - 20\log P_{01}(\nu)$. Using the drive levels in Ref. 2 of $V_{10}=56.6V$ peak and $I_{10}=1.4A$ peak the final numerical results at $\nu=1$ are

$$20\log\left|\frac{p^M_{202}}{p_{01}}\right| \simeq -25\text{ dB}, \quad 20\log\left|\frac{p^M_{303}}{p_{01}}\right| \simeq -48\text{ dB},$$

$$20\log\left|\frac{p_{202}}{p_{01}}\right| \simeq -17\text{ dB}, \quad 20\log\left|\frac{p_{303}}{p_{01}}\right| \simeq -36\text{ dB}.$$

These results agree approximately with Table I and Figs. 2 and 8 of Ref. 2. In this case ID gives lower relative harmonics than DD mainly because the ID fundamental is about 8 dB lower than the DD fundamental at resonance. Thus the ID relative second and third harmonics should be raised 8 and 16 dB, respectively, for a reasonable comparison. Then the ID and DD relative harmonics are about the same. Such comparisons will be discussed more generally in the next section.

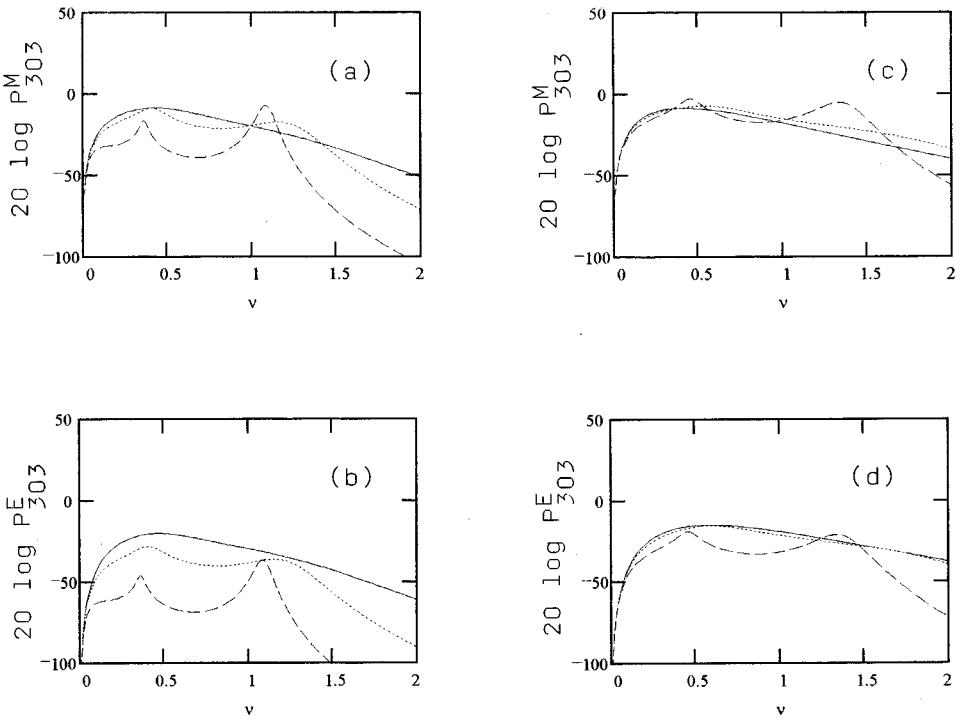


FIG. 3. Normalized pressure P (in dB) as a function of the dimensionless frequency ν for third harmonics caused by one nonlinear parameter in the mechanical equation and one in the electrical equation; (a) k_{30}, (b) β_2 for $Q_{me}=0.5$, $Q_e=0.5$, and $Q_m=1$ (———), 3 (.....), and 10 (---), (c)–(d) same except for $Q_e=0.1$.

III. COMPARISON OF ID AND DD HARMONICS

Each of the static components in Eq. (19) can be related to a DD static component from the solution of Eq. (60).[1] For example, the 200 DD component is

$$X_{200}=\frac{\gamma_{20}}{2\omega_0^2}X_{01}^2=\frac{M\gamma_{20}}{2k_{10}}X_{01}^2;$$

and from Eq. (19a)

$$Y_{200}=-X_{200}\frac{k_{20}}{M\gamma_{20}}\left[\frac{I_{10}}{V_{10}|Y_{em}(\omega)|}\right]^2,$$

for electric field transducers. Thus Y_{200} is equal to X_{200} with the DD voltage V_{10} replaced by the voltage $I_{10}/|Y_{em}(\omega)|$ and $M\gamma_{20}$ replaced by k_{20}. It can be seen by comparing Tables II–VI with Tables I–III of Ref. 1 that $M\gamma_{20}\simeq -k_{20}$ and that

$$M\gamma_{n0}\simeq -k_{n0}+\tfrac{1}{2}\delta_{1n}V_{10}^2, \quad k_{00}=0, \tag{27a}$$

$$M\gamma_{n1}\simeq \gamma_n V_{10}+\tfrac{3}{4}\delta_{2n}V_{10}^3, \tag{27b}$$

$$M\gamma_{n2}\simeq \tfrac{1}{2}\delta_{1n}V_{10}^2, \tag{27c}$$

$$M\gamma_{n3}\simeq \tfrac{1}{4}\delta_{2n}V_{10}^3, \tag{27d}$$

for electric field transducers, with V_{10} replaced by I_{10} for magnetic field transducers. Thus

$$Y_{200}\simeq X_{200}\left[\frac{I_{10}}{V_{10}|Y_{em}(\omega)|}\right]^2.$$

The same relationship holds between Y_{110} and X_{110} and between Y_{000} and X_{000}. Note that the nonlinear parameters in the electrical equation do not generate any static components in the first order approximation.

In most cases Eqs. (27a)–(27d) appear to be exact, but the γ_{nm} are defined such that they would change if more terms were included in Eq. (6). γ_{nm} represents a mathematical nonlinear mechanism ($X^n\cos m\omega t$) which usually is strongly related to one physical nonlinear mechanism such as γ_n or δ_{1n}, but generally is a mixture of two or more physical nonlinear mechanisms as indicated by Eqs. (27a) and (27b).

The Y^M_{nmN} ID components in Eqs. (20) and (21) are closely related to the DD X_{nmN} components since both are caused by nonlinear parameters in the mechanical equation; e.g., consider the DD second and third harmonic components, generated by the nonlinear spring from Eqs. (61a)[1] and (62a):[1]

$$X_{202}=\frac{\gamma_{20}X_{01}^2}{2\omega_0^2 z_m(2\nu)}\cos(2\omega t+2\phi_1+\phi_2),$$

$$X_{303}=\frac{\gamma_{30}X_{01}^3}{4\omega_0^2 z_m(3\nu)}\cos(3\omega t+3\phi_1+\phi_3).$$

Using Eqs. (27) and (17c) in Eqs. (20a) and (21a) shows that

$$|Y^M_{202}|\simeq|X_{202}|\left[\frac{I_{10}}{V_{10}|Y_{em}(\omega)|}\right]^2\left|\frac{Y_e(2\omega)}{Y_{em}(2\omega)}\right|,$$

$$|Y^M_{303}|\simeq|X_{303}|\left[\frac{I_{10}}{V_{10}|Y_{em}(\omega)|}\right]^3\left|\frac{Y_e(3\omega)}{Y_{em}(3\omega)}\right|,$$

with identical relations between $|Y^M_{112}|$ and $|X_{112}|$, etc. Here, as for the static components, $I_{10}/|Y_{em}(\omega)|$ replaces V_{10}

which smooths the peak at $\omega=\omega_0$. Other differences between the ID and DD components are caused by the factor $|Y_e(N\omega)/Y_{em}(N\omega)|$ which becomes small at $\omega=\omega_0/N$.

A major difference between ID and DD is the occurrence in ID of the additional components, Y^E_{nmN}, which are not related to DD components in general, because they arise from nonlinear parameters in the electrical equation which play no role in DD. However, although nonlinear parameters such as k_{20}, γ_1, δ_{10}, β_1, α_{11}, and α_{20} are independent in general, some are related to others for certain transducers (see Tables I–VI). For example, for the models of electrostatic and moving coil transducers used here $\beta(Y)=\gamma(Y)$ and $\beta_n=\gamma_n$ for all n. Also, for these transducers α_{11} and δ_{10} are proportional. For any transducer with a nonlinear spring DD excites X_{202} and ID excites Y^M_{202}. However, when the coupling is nonlinear, DD excites X_{112}, but ID excites both Y^M_{112} and Y^E_{202}, because the coupling is involved in both the electrical and mechanical equations. Similarly, when DD excites X_{022}, ID excites both Y^M_{022} and Y^E_{112}.

The relative magnitude of the two second harmonic components excited together by ID is

$$\left|\frac{Y^E_{202}}{Y^M_{112}}\right| = \frac{\beta_1\gamma_0\omega Y_{01}}{\gamma_1 V_{01}|Y_e(2\omega)|} = \frac{\gamma_0\beta_1\nu(1-Q_{me}/Q_m)}{\beta_0\gamma_1 Q_{me} z_m(\nu)[1+(2\nu Q_e Q_m/Q_{me})^2]^{1/2}}. \tag{28}$$

The ratio $|Y^E_{112}/Y^M_{022}|$ is the same with β_1/γ_1 replaced by $2\alpha_{11}/\delta_{10}$. At $\omega=\omega_0$

$$\left|\frac{Y^E_{202}}{Y^M_{112}}\right| = \frac{\gamma_0\beta_1(Q_m-Q_{me})}{\beta_0\gamma_1 Q_{me}\left[1+\left(2\frac{Q_eQ_m}{Q_{me}}\right)^2\right]^{1/2}}, \tag{28a}$$

and for $\omega\gg\omega_0$

$$\left|\frac{Y^E_{202}}{Y^M_{112}}\right| = \frac{\gamma_0\beta_1(1-Q_{me}/Q_m)}{2\beta_0\gamma_1\nu^2 Q_e Q_m}. \tag{28b}$$

Thus at sufficiently high frequency the Y^E components become small compared to the accompanying Y^M components. To this order of approximation there is no Y^M component associated with Y^E_{022}.

The question of which drive gives lower distortion can best be addressed by comparing the magnitude of ID and DD harmonics relative to their respective fundamentals. For example,

$$\frac{|Y^M_{202}|/Y_{01}}{|X_{202}|/X_{01}} = \frac{|P^M_{202}|/|P_{01}ID|}{|P_{202}|/|P_{01}DD|}$$

is the relevant quantity for second harmonics caused by a nonlinear spring. For other nonlinear mechanisms where Y^M and Y^E components are excited together in ID with a frequency-dependent phase difference the comparison is more complicated. However, Eq. (20) shows that for nonlinear coupling

$$\frac{|Y^M_{112}+Y^E_{202}|/Y_{01}}{|X_{112}|/X_{01}} = \frac{|Y^M_{112}|/Y_{01}}{|X_{112}|/X_{01}}\left[1+\frac{|Y^E_{202}|^2}{|Y^M_{112}|}+2\frac{|Y^E_{202}|}{|Y^M_{112}|}\sin(\psi_2-\phi_1)\right]^{1/2}, \tag{29}$$

with a similar expression for nonlinear electrical components in which Y^M_{112} is replaced by Y^M_{022}, Y^E_{202} by Y^E_{112}, and X_{112} by X_{022}. The ratios

$$\frac{Y^M_{nm2}|/Y_{01}}{|X_{nm2}|/X_{01}} \simeq \frac{I_{10}|Y_e(2\omega)|}{V_{10}|Y_{em}(2\omega)||Y_{em}(\omega)|} = \frac{I_{10}z_m(\nu)z_m(2\nu)[1+(2\nu(Q_eQ_m/Q_{me}))^2]^{1/2}}{G_eV_{10}w(\nu)w(2\nu)} \tag{30}$$

are equal for the three nonlinear mechanisms designated by $nm=20$, 11, and 02. These results apply to magnetic field transducers when I_{10} and V_{10} are interchanged and G_e is replaced by R_e. Note that Eq. (30) depends only on linear transducer parameters, because the nonlinear parameters which cause these harmonics are approximately the same for ID and DD.

Numerical comparison of ID and DD distortion requires specifying the driving current and voltage or a relationship between them. Since current (voltage) drive implies constant current (voltage) over some frequency band the current and voltage could be chosen to make the fundamentals equal at one frequency in the band of interest. This approach would give a comparison of ID and DD distortion in which the fundamentals differed except at one frequency. Another approach would, for a constant current from the current source, change the voltage of the voltage source at each frequency to keep the fundamentals equal across the band of interest. This is the approach taken by Mills and Hawksford[3] in their comparison of current and voltage drive of a moving coil transducer at two widely spaced frequencies. For this comparison

$$I_{10}=V_{10}|Y_{em}(\omega)|,$$

and Eq. (30) becomes

$$\frac{|Y^M_{nm2}|/Y_{01}}{|X_{nm2}|/X_{01}} \simeq \frac{|Y_e(2\omega)|}{|Y_{em}(2\omega)|},$$

which approaches unity both above and below resonance, but can be more or less than unity near resonance.

Since the definition of current (voltage) drive implies fixed current (voltage) the former type of comparison will be pursued here. For transducers intended for use near resonance it's natural to adjust the current and voltage to make the fundamentals equal at resonance which requires

$$I_{10}=V_{10}|Y_{em}(\omega_0)|.$$

Then Eq. (30) becomes

$$\frac{|Y^M_{nm2}|/Y_{01}}{|X_{nm2}|/X_{01}} \simeq \frac{|Y_e(2\omega)||Y_{em}(\omega_0)|}{|Y_{em}(2\omega)||Y_{em}(\omega)|},$$

which will be evaluated more fully only at resonance where it can be written

$$\frac{|Y^M_{nm2}|/Y_{01}}{|X_{nm2}|/X_{01}} \simeq \frac{|Y_e(2\omega_0)|}{|Y_{em}(2\omega_0)|} \simeq \left| \frac{((Q_{me}/Q_m)-3Q_eQ_m)+j(2Q_e+\frac{3}{2}Q_{me})}{(1-3Q_eQ_m)+j(2Q_e+\frac{3}{2}Q_{me})} \right|. \quad (31)$$

While Eqs. (29) and (31) hold for all values of the parameters consideration of the special cases of high and low electrical loss will reveal the important aspects of the relative ID to DD distortion. For high electrical loss G_e(or R_e)$\rightarrow\infty$, $Q_{me}\rightarrow Q_m$, $Q_e\rightarrow 0$, $w(N\nu)\rightarrow z_m(N\nu)$, Eq. (31)$\rightarrow$1, Eq. (28)$\rightarrow$0, Eq. (29)$\rightarrow$1 and second harmonic distortion is the same for ID and DD. Furthermore, in this limit, ID becomes the same as DD in all respects, because high G_e short circuits an electric field transducer which corresponds to voltage drive (DD) while high R_e open circuits a magnetic field transducer which corresponds to current drive (DD). However, the limit of high electrical loss is not of much practical importance—no well-designed transducer would have such high losses.

The other limit of low electrical loss has more practical importance. Then G_e(or R_e)$\ll\beta_0\gamma_0/R$, $Q_{me}\ll Q_m$ and

$$Q_eQ_m \simeq \frac{1-k^2}{k^2}, \quad (32)$$

where

$$k=\left(\frac{1}{1+\omega_0^2 M\alpha_{10}/\beta_0\gamma_0}\right)^{1/2}, \quad (33)$$

is the electromechanical coupling coefficient.[4] Since k is one of the most commonly used general transducer parameters its introduction may facilitate interpretation of the following results. Using Eq. (32) shows that Eq. (31) is less than unity for any transducer with $k\rightarrow 1$ since $Q_{me}/Q_m<1$. Thus for any transducer which has sufficiently high coupling, low electrical loss, and dominant nonlinearity in the springs the relative second harmonic at resonance is less for ID than for DD. However, the coupling has to be very high for this to occur since Eq. (31) exceeds unity when

$$\frac{1-k^2}{k^2}>\frac{1}{6}\left(1+\frac{Q_{me}}{Q_m}\right)\simeq\frac{1}{6},$$

or when $k<0.925$. Coupling this high can probably only be achieved in moving coil transducers and electrostatic and variable reluctance transducers polarized to near instability which occurs[4] at $k=1$ or in magnetostrictive metallic glass transducers.[5] The moving coil transducer of Ref. 2, discussed earlier, is a case of very high coupling ($k\simeq 0.98$). Using the parameters for this transducer Eq. (31) equals 0.64 or -4 dB, which is roughly consistent with the earlier calculations for this transducer when they are adjusted to make the fundamentals equal.

If Eq. (31) is further approximated by setting $Q_{me}=0$ it becomes

$$\frac{|Y^M_{202}|/Y_{01}}{|X_{202}|/X_{01}}=\left[\frac{(1+\frac{9}{4}Q_m^2)}{1+[(1-3Q_mQ_e)/2Q_e]^2}\right]^{1/2}. \quad (34)$$

With the same approximations for low electrical loss, with $\beta_0=\gamma_0$ (holds for all reciprocal transducers), $\beta_1=\gamma_1$ [holds for electrostatic, moving coil and variable reluctance (without saturation) transducers] and use of Eq. (28a), Eq. (29) can be written at resonance

$$\frac{|Y^M_{112}+Y^E_{202}|/Y_{01}}{|X_{112}|/X_{01}}=\left[\frac{(1+\frac{9}{4}Q_m^2)(1+1/4Q_e^2)}{1+[(1-3Q_mQ_e)/2Q_e]^2}\right]^{1/2}. \quad (35)$$

Equation (35) also holds for $|Y^M_{022}+Y^E_{112}|/Y_{01}/|X_{022}|/X_{01}$ when $2\alpha_{11}=\delta_{10}$.

The question of which drive gives lower second harmonic is answered by Eqs. (35) and (34) for transducers with low electrical loss operated at resonance. Numerical results are given in Tables VII and VIII as a function of k and Q_m [for each k and Q_m Eq. (32) determines Q_e]. Note that as $k\rightarrow 1$, $Q_e\rightarrow 0$, Eq. (35)$\rightarrow(1+\frac{9}{4}Q_m^2)^{1/2}$ and Eq. (34)$\rightarrow$0, while for $k\rightarrow 0$, $Q_e\rightarrow\infty$ and both Eqs. (35) and (34)$\rightarrow$1.

Although Tables VII and VIII hold for the ID and DD fundamentals being made equal at resonance they can be easily modified to hold when the fundamentals are unequal at resonance; e.g., if the ID fundamental is half the DD fundamental at resonance Eqs. (31), (34), and (35) and the numbers in the tables are divided by two.

The results in Tables VII and VIII show that for transducers in which nonlinear coupling or nonlinear electrical components is the dominant nonlinearity ID always generates higher second harmonics than DD at resonance, while the same holds for nonlinear springs except when $Q_mQ_e<\frac{1}{6}$; i.e., when $k>0.925$. The moving coil transducer of Ref. 2 has $k\approx 0.98$, the dominant nonlinearity is in the springs, and the ID second harmonics are less than, or similar to, the DD second harmonics at resonance. On the other hand the moving coil transducer of Ref. 3 also has $k\simeq 0.98$, but the dominant nonlinearity appears to be in the coupling, and the ID second harmonics exceed the DD second harmonics at resonance. Thus the second harmonics generated in these two transducers qualitatively illustrate the results in Tables VII and VIII, although it must be noted that, while both ID and DD harmonics were measured in Ref. 3, only ID harmonics were measured in Ref. 2.

Klippel[6] also measured harmonics for both current and voltage drive in a moving coil transducer with nonlinear springs and nonlinear coupling (Bl product variations). Ratios of relative second harmonics can be calculated from his results which are lower for voltage drive (ID) than for current drive (DD) near resonance.

The results above apply when one nonlinear mechanism is dominant. When more than one mechanism is important in the same transducer the combined harmonics can be found using the amplitudes and phases of the individual harmonic components given in Eq. (20).

A different limiting case that shows another aspect of the relationship between ID and DD occurs when the electrical component is very small ($\alpha_{10}\rightarrow 0$), which is probably only relevant to certain moving coil transducers such as

those with shorting rings and ribbon speakers. Thus for $\alpha_{10}=L_0\rightarrow 0$ Eq. (15) for magnetic field transducers becomes

$$M\ddot{y}_0+\left(R+\frac{\beta_0\gamma_0}{R_e}\right)\dot{y}_0+M\omega_0^2 y_0=\frac{\gamma_0 V_{10}}{R_e}\cos\,\omega t,$$

which is the mechanical equation for the DD fundamental with V_{10}/R_e in place of I_{10} and the mechanical resistance augmented by the reflected electrical resistance. In this limit $Q_e\rightarrow 0$, $Y_e\rightarrow G_e$, $\psi_N\rightarrow 0$ and $w(N\nu)\rightarrow z_{me}(N\nu)$, where z_{me} is the same function as z_m with Q_m replaced by Q_{me}. The Y^M_{nm2} components then become similar to the X_{nm2} components with Q_{me} in place of Q_m. However, this case is not the same as DD, because the Y^E_{nm2} components still exist and are in phase quadrature with Y^M_{nm2}.

IV. ACCURACY

The accuracy of the first- and second-order perturbation results given here depends on the ratio of the harmonic to fundamental amplitude which in turn depends on the drive level, the type of nonlinearity causing the harmonic, the magnitude of the nonlinear parameter and the drive frequency. The third-order corrections to the first-order DD second harmonic components caused by γ_{20} and γ_{11} will be discussed as illustrations (second order does not give any second harmonics). In each case the third-order contributions consist of several terms; for γ_{20} one term is dominant for $\omega\geqslant\omega_0$ and if the smaller terms are neglected the sum of the first- and third-order contributions is

$$\frac{\gamma_{20}X_{01}^2}{2\omega_0^2 z_m(2\nu)}\cos(2\omega t+2\phi_1+\phi_2)$$
$$+\frac{\gamma_{20}^3X_{01}^4}{\omega_0^6 z_m(\nu)z_m(2\nu)}\cos(2\omega t+3\phi_1+\phi_2),\quad \omega\geqslant\omega_0.$$

For γ_{11} one third-order term is dominant at $\omega=\omega_0$, while a different term is dominant for $\omega\geqslant\sqrt{2}\omega_0$. The sum of the first and third-order contributions is

$$\frac{\gamma_{11}X_{01}}{2\omega_0^2 z_m(2\nu)}\cos(2\omega t+\phi_1+\phi_2)$$
$$+\frac{\gamma_{11}^3X_{01}}{8\omega_0^6 z_m(\nu)z_m^2(2\nu)}\cos(2\omega t+2\phi_1+2\phi_2),\quad \omega=\omega_0,$$

$$\frac{\gamma_{11}X_{01}}{2\omega_0^2 z_m(2\nu)}\cos(2\omega t+\phi_1+\phi_2)$$
$$+\frac{\gamma_{11}^3X_{01}\cos\,\phi_1}{4\omega_0^6 z_m(\nu)z_m(2\nu)}\cos(2\omega t+\phi_1+\phi_2),\quad \omega\geqslant\sqrt{2}\omega_0.$$

In each of these cases the first term is the first-order approximation given in Ref. 1 and denoted by X_{202} and X_{112}, respectively. The ratio of the third-order amplitude to the first-order amplitude will be used as a measure of the accuracy of the first-order approximation:

$$R_{20231}\simeq\frac{2\gamma_{20}^2X_{01}^2}{\omega_0^4 z_m(\nu)}=\frac{8z_m^2(2\nu)}{z_m(\nu)}\left(\frac{|X_{202}|}{X_{01}}\right)^2,\quad \omega\geqslant\omega_0,$$

$$R_{11231}\simeq\frac{\gamma_{11}^2}{4\omega_0^4 z_m(\nu)z_m(2\nu)}=3Q_m\left(\frac{|X_{112}|}{X_{01}}\right)^2,\quad \omega=\omega_0,$$

$$R_{11231}\simeq\frac{\gamma_{11}^2\cos\,\phi_1}{2\omega_0^4 z_m(\nu)}=\frac{2z_m^2(2\nu)}{z_m(\nu)}\left(\frac{|X_{112}|}{X_{01}}\right)^2,\quad \omega\geqslant\sqrt{2}\omega_0.$$

The last forms of these ratios make determination of the accuracy easy since it only requires multiplying the square of the first approximation (such as $|X_{202}|/X_{01}$) by a function of frequency.

As a numerical example consider a transducer with $Q_m=3$ and $|X_{202}|/X_{01}=2\%(-34$ dB) at $\omega=\omega_0$; then $R_{20231}\simeq 9\%(-21$ dB). At $\omega=\sqrt{2}\omega_0|X_{202}|/X_{01}$ has gone down to 0.3% (-51 dB) and $R_{20231}=0.3\%(-50$ dB) which illustrates the extreme frequency dependence of the accuracy for X_{202}. The accuracy is better for the γ_{11} mechanism that generates X_{112}; for $Q_m=3$ and $|X_{112}|/X_{01}=2\%(-34$ dB) at $\omega=\omega_0$ $R_{11231}=0.35\%(-49$ dB), while at $\omega=\sqrt{2}\omega_0|X_{112}|/X_{01}|$ has gone down to $1\%(-39$ dB) and $R_{11231}=0.9\%(-41$ dB). In both cases the accuracy improves at higher frequency.

These examples show how the accuracy depends strongly on the type of nonlinear mechanism and on the frequency. The accuracy is poorer near resonance, for high Q_m and for nonlinear mechanisms which depend strongly on displacement. For example, accuracy is poorer for $\gamma_{20}X^2$ than for $\gamma_{11}X\cos\,\omega t$ because the former is nonlinear in X while the latter is linear in X but has a time-dependent coefficient. Although these examples use DD harmonics the accuracy is expected to be better for ID harmonics which have broader resonant peaks. However, it is apparent that some of the numerical results given previously for the transducer of Ref. 2 have poor accuracy and that some of the ID to DD comparisons at resonance are only valid when the harmonic to fundamental ratio is quite small.

V. DISCUSSION

Although the approach in this paper and Ref. 1 has been comprehensive in some respects certain simplifications have been made which will now be discussed. The mechanical and electrical resistances were assumed to be linear. The effects of intermediate amplifier impedance or additional electric circuit elements were not considered. Inclusion of nonlinear resistances would introduce more nonlinear parameters and lead to additional components of each harmonic. Consideration of more complex electrical circuits containing linear circuit elements would modify the harmonic components already calculated. Both of these matters can be handled by the methods of analysis used here.

Each transducer type was modeled as the simplest lumped parameter mechanical structure, but this is not as serious a limitation as it seems at first. The results obtained here can be applied to transducers with more complex structures if the active drive part of the transducer undergoes essentially one-dimensional motion and if all the other moving parts contain no significant nonlinearities. For example, many flextensional transducers with a drive part consisting

TABLE VII. Relative ID to DD second harmonic at resonance for nonlinear springs, Eq. (34).

k	$Q_m=1$	$Q_m=3$	$Q_m=10$
1	0	0	0
0.991	0.076	0.066	0.064
0.978	0.21	0.18	0.176
0.955	0.49	0.44	0.43
0.925	1	1	1
0.914	1.28	1.47	1.5
0.865	1.8	4.6	15
0.816	1.44	2.5	2.9
0.707	1.27	1.46	1.5
0.577	1.13	1.18	1.20
0.448	1.06	1.08	1.10

of a piezoceramic stack satisfy these conditions. This can be seen by noting that both the ID and DD harmonic displacement components can be expressed in terms of the fundamental displacement at the drive frequency and at the harmonic frequency. Consider, e.g., the ID second harmonic component caused by nonlinear springs; Eq. (20a) can be written, using the expression for $Y_{01}(\omega)$ in Eq. (16),

$$Y_{202}^{M} = -\frac{k_{20}}{2\gamma_0 I_{10}}|Y_e(2\omega)|Y_{01}^2(\omega)Y_{01}(2\omega) \times\cos(2\omega t+2\theta_1+\psi_2+\theta_2),$$

$$= -\frac{k_{20}}{2\gamma_0 I_{10}}\,\mathrm{Re}\{Y_e(2\omega) \times[Y_{01}(\omega)e^{j\theta_1}]^2 Y_{01}(2\omega)e^{j\theta_2}e^{j2\omega t}\},$$

where Re stands for real part of. Similar relations for DD harmonics are given in Ref. 1. Note that Y_{202}^{M} is expressed entirely in terms of linear transducer parameters except for the nonlinear parameter k_{20}. Thus using a linear model of a transducer such as a flextensional which relates displacement at the end of the drive part to acoustic pressure in the medium and is valid over the frequency range from ω to $N\omega$ the pressure at frequency $N\omega$ caused by a specific nonlinear mechanism is equal to the pressure at $N\omega$ given by the linear model multiplied by the appropriate factor for each harmonic component. For example, the magnitude of the factor for Y_{202}^{M} is $k_{20}|Y_e(2\omega)|Y_{01}^2(\omega)/2\gamma_0 I_{10}$ where $|Y_e(2\omega)|$ and $Y_{01}(\omega)$ are the blocked admittance and fundamental displacement as given by the linear model of the transducer with more complex structure. Writing Y_{202}^{M} in complex notation shows that the phases of the harmonic components can be included and that different components of each harmonic can be combined with the correct phase relations.[7] As long as the nonlinearities are confined to the drive part of the transducer, the part which has been analyzed here, transmission of the harmonic motion through the rest of the transducer structure and radiation into the medium are linear processes.

TABLE VIII. Relative ID to DD second harmonic at resonance for nonlinear coupling or nonlinear electrical components, Eq. (35).

k	$Q_m=1$	$Q_m=3$	$Q_m=10$
1	1.80	4.6	15
0.991	1.9	5.1	16
0.978	2.1	5.6	18
0.955	2.4	6.6	23
0.925	3.1	9.2	29
0.914	3.4	11.2	38
0.865	3.2	2.21	225
0.816	2	8.2	29
0.707	1.4	2.6	7.5
0.577	1.16	1.47	3.2
0.448	1.07	1.28	1.6

The calculation of more third harmonic components and corrections to the fundamental is straightforward starting from Eqs. (14a)–(14b). Higher-order harmonics can be calculated by the same methods, but it becomes increasingly tedious, and no simpler method for obtaining a first approximation to higher harmonics was found for ID such as the one found for DD.[1] Numerical methods applied to specific cases might be more effective for higher harmonics.

VI. CONCLUSION

The results presented here and in Ref. 1 apply to harmonics generated by all the common nonlinear mechanisms in all the major types of electroacoustic transducers for both voltage drive and current drive. Expressions for the harmonics are given in terms of universal nonlinear parameters (γ_{nm} in Ref. 1 for direct drive, α_{1n}, β_n, etc. here for indirect drive) which are tabulated for the six major types of electroacoustic transducers. Each tabulation represents a particular simple model of each transducer type. More complete or more accurate models could be developed and expressed in terms of the same nonlinear parameters which could then be used in the same expressions for the harmonics.

These results can be used to estimate harmonic distortion in specific transducers without further analysis by using values of the standard linear transducer parameters and values of the nonlinear parameters. The nonlinear parameter values can be calculated in some cases from linear parameters, but in most cases they depend on nonlinear properties of materials, nonlinear spring constants, etc. which can only be found by measurement. Examples are given for a specific moving coil transducer by calculating second and third harmonic components relative to the fundamental.

The practical question of which drive gives lower harmonic distortion was addressed by using the general results for indirect drive (ID) and direct drive (DD). ID generates more components of each harmonic than DD because it involves nonlinearities in both the electrical and mechanical transducer equations. In ID the linear electrical parameters change the harmonic components that are generated in DD, while the nonlinear electrical parameters generate additional components of each harmonic. Thus it would be expected that ID would generate higher harmonics than DD under many conditions, but the opposite might occur under some conditions. It is difficult to determine these conditions in a general way because of the large number of parameters involved, the dependence on drive frequency and the question of how the ID to DD comparison should be made. The limited investigation carried out here has shown that for transducers of all types with low electrical loss operated near resonance DD usually gives lower second harmonics. The

exceptions occur when mechanical springs are the dominant nonlinearity and the electromechanical coupling is very high, and then ID gives lower second harmonics. This is most likely to be observed in moving coil transducers which often have very high coupling. A general answer to the question of which drive gives lower harmonics was not found, but the question can be answered for any specific transducer (or specific values of Q_m, Q_{me}, and Q_e) in specific frequency ranges by use of Eqs. (29) and (30).

These results will facilitate estimation of harmonic distortion in most cases of transducer design and evaluation. They will allow harmonics to be estimated prior to construction and measurement. They may aid in diagnosting the causes of measured distortion and identifying the measures required to reduce excessive distortion. Although based on simple lumped parameter models, these results can be applied in some cases to more complex transducer structures.

APPENDIX

1. Variable reluctance transducers

The mechanical equation for the variable reluctance transducer, including an approximation for the effects of magnetic saturation, is given by Eq. (25),[1] and the electrical equation is obtained by putting Eq. (19)[1] into Eq. (22)[1] with $I=I_0+I_1(t)$ (see Fig. 2 of Ref. 1). When written in a form for easy comparison with Eqs. (3) and (4) these equations become:

$$V_1=\frac{\mu_0 AN^2}{4}\frac{d}{dt}\left\{\left[\frac{1}{(l_0+Y_1)}-\frac{4I_0S}{(l_0+Y_1)^3}\right]I_1-\frac{2S}{(l_0+Y_1)^3}I_1^2\right\}-\frac{\mu_0 AN^2 I_0}{4}\left\{\left[\frac{1}{(l_0+Y_1)^2}-\frac{6I_0S}{(l_0+Y_1)^4}\right]\right\}\dot{Y}_1+R_eI_1, \tag{A1}$$

$$M\ddot{Y}_1+R\dot{Y}_1=-k_0(Y_1)Y_1-\frac{\mu_0 AN^2 I_0}{4}\left[\frac{1}{(l_0+Y_1)^2}-\frac{9I_0S}{2(l_0+Y_1)^4}\right]I_1-\frac{\mu_0 AN^2}{8}\left[\frac{1}{(l_0+Y_1)^2}-\frac{9I_0S}{(l_0+Y_1)^4}\right]I_1^2+\frac{3\mu_0 AN^2S}{8(l_0+Y_1)^4}I_1^3, \tag{A2}$$

where

$S=\frac{1}{4}\mu_0^2 Nl_m\nu_2$,

$\nu_2=$ magnetic saturation parameter,[1]

$l_m=$ length of magnetic circuit,

$N=$ number of turns in coil,

$V_1=$ alternating voltage,

$I_0=$ bias (polarizing) current,

$I_1=$ alternating current,

$L_0=\mu_0 AN^2/4(l_0+Y_1)=$ inductance,

$A/2=$ cross-sectional area of magnetic circuit,

$l+Y_0=$ biased separation of magnetic poles,

$l_0=l+Y_0+\mu_0 l_m/2\mu$,

$\mu_0=$ permeability of free space,

$\mu=$ permeability of the magnetic material,

$Y_0=$ static displacement caused by bias current,

$Y_1=$ alternating displacement,

$R_e=$ electrical resistance,

$M=$ mass of moveable part plus radiation mass,

$R=$ viscous resistance in the transducer plus radiation resistance.

$$k_0(Y_1)=\left[k_1+2Y_0k_2+3Y_0^2k_3-\frac{\mu_0 AN^2}{4}\left(\frac{I_0^2}{l_0^3}-\frac{6SI_0^3}{l_0^5}\right)\right]+\left[k_2+3Y_0k_3+\frac{\mu_0 AN^2}{4}\left(\frac{3I_0^2}{2l_0^4}-\frac{15SI_0^3}{l_0^6}\right)\right]Y_1+\left[k_3-\frac{\mu_0 AN^2}{4}\left(\frac{2I_0^2}{l_0^5}-\frac{30SI_0^3}{l_0^7}\right)\right]Y_1^2$$

$k_1+k_2Y_1+k_3Y_1^2+\cdots=$ nonlinear mechanical spring ''constant.''

Comparing Eqs. (A1) and (A2) with Eqs. (3) and (4) identifies the functions α_1, α_2, β, γ, δ_1, and δ_2:

$$\alpha_1=L_0\left[1-\frac{4I_0S}{(l_0+Y)^2}\right],\quad \alpha_2=\frac{-2SL_0}{(l_0+Y)^2},$$

$$\beta=I_0\frac{dL_0}{dY}\left[1-\frac{6I_0S}{(l_0+Y)^2}\right],$$

$$\gamma=I_0\frac{dL_0}{dY}\left[1-\frac{9I_0S}{2(l_0+Y)^2}\right],$$

$$\delta_1=\frac{1}{2}\frac{dL_0}{dY}\left[1-\frac{9I_0S}{(l_0+Y)^2}\right],$$

$$\delta_2=-\frac{3S}{2(l_0+Y)^2}\frac{dL_0}{dY}.$$

2. Moving coil transducers

The equations for the moving coil transducer are given by Eqs. (29)[1] and (28)[1] (see Fig. 3 of Ref. 1):

$$V=\frac{d}{dt}(LI)+B(Y)l_c\dot{Y}+R_eI, \tag{A3}$$

$$M\ddot{Y}+R\dot{Y}=-k_0(Y)Y+B(Y)l_cI-\frac{1}{2}\frac{dL}{dY}I^2, \tag{A4}$$

where

$L=L_0+L_1Y+L_2Y^2+\cdots$

is the inductance of the moving coil,

$B=B_0+B_1Y+B_2Y^2+\cdots$

is the static magnetic flux density in which the coil moves,

$$k_0(Y)=k_1+k_2Y+k_3Y^2+\cdots$$

is the combined nonlinear mechanical stiffness of suspension and enclosure, and V=alternating voltage, I=alternating current, l_c=length of wire in the coil.

Comparing Eqs. (A3) and (A4) with Eqs. (3) and (4) identifies the functions α_1, β, γ, and δ_1 (the only ones required in this case):

$$\alpha_1=L(Y),$$

$$\beta=\gamma=l_cB(Y),$$

$$\delta_1=-\frac{1}{2}\frac{dL}{dY}.$$

3. Piezoelectric transducers

Equations (34)[1] and (38)[1] give the electrical equation and Eq. (39)[1] the mechanical equation, for piezoelectric transducers (see Fig. 4 of Ref. 1):

$$I=A\frac{d}{dt}\left\{\left[\frac{\epsilon_1+2e_2(Y/l_0)}{l_0+Y}\right]V+\frac{\epsilon_2}{(l_0+Y)^2}V^2\right\}+\frac{A}{l_0}\left(e_1+2c_a\frac{Y}{l_0}\right)\dot{Y}+VG_e, \tag{A5}$$

$$M\ddot{Y}+\dot{R}Y=-\frac{A}{l_0}\left[c_1+c_2\frac{Y}{l_0}\right]Y+A\left[\frac{e_1}{l_0+Y}+\frac{2c_a(Y/l_0)}{l_0+Y}\right]V+A\frac{e_2}{(l_0+Y)^2}V^2, \tag{A6}$$

where c_1 and c_2 are linear and nonlinear elastic constants,[1] ϵ_1 and ϵ_2 are linear and nonlinear permittivities,[1] e_1, e_2, and c_a are linear and nonlinear piezoelectric constants.[1] The functions k_0, α_1, α_2, β, γ, and δ_1 are identified by comparison with Eqs. (3) and (4):

$$\alpha_1=\frac{A}{(l_0+Y)}\left[\epsilon_1+2e_2\frac{Y}{l_0}\right],\quad \alpha_2=\frac{A\epsilon_2}{(l_0+Y)^2},$$

$$\beta=\frac{A}{l_0}\left[e_1+2c_a\frac{Y}{l_0}\right],$$

$$\gamma=\frac{A}{(l_0+Y)}\left[e_1+2c_a\frac{Y}{l_0}\right],$$

$$\delta_1=\frac{Ae_2}{(l_0+Y)^2},$$

$$k_0=\frac{A}{l_0}\left[c_1+c_2\frac{Y}{l_0}\right].$$

4. Electrostrictive transducers; maintained polarization

Transducers using electrostrictive materials with small hysteresis which require maintaining a polarizing electric field, E_0, have an electrical equation obtained by inserting Eq. (45)[1] into Eq. (38)[1] and a mechanical equation given by Eq. (46).[1] When these equations are compared with Eqs. (3) and (4) the functions are found to be (see Fig. 4 of Ref. 1):

$$\alpha_1=\frac{A}{l_0+Y}\left\{\epsilon_1+2E_0\epsilon_2+3E_0^2\epsilon_3+\left[2e_2+12e_4E_0^2+3c_c\left(S_0+\frac{Y}{l_0}\right)\right]\left(S_0+\frac{Y}{l_0}\right)\right\},$$

$$\alpha_2=\frac{A}{(l_0+Y)^2}\left\{\epsilon_2+3E_0\epsilon_3+12e_4E_0\left(S_0+\frac{Y}{l_0}\right)\right\},$$

$$\alpha_3=\frac{A}{(l_0+Y)^3}\left\{\epsilon_3+4e_4\left(S_0+\frac{Y}{l_0}\right)\right\},$$

$$\beta=\frac{A}{l_0}\left\{2e_2E_0+4e_4E_0^3+6c_cE_0\left(S_0+\frac{Y}{l_0}\right)\right\},$$

$$\gamma=\frac{A}{l_0+Y}\left\{2e_2E_0+4e_4E_0^3+6c_cE_0\left(S_0+\frac{Y}{l_0}\right)\right\},$$

$$\delta_1=\frac{A}{(l_0+Y)^2}\left\{e_2+6e_4E_0^2+3c_c\left(S_0+\frac{Y}{l_0}\right)\right\},$$

$$\delta_2=\frac{A}{(l_0+Y)^3}4e_4E_0,$$

$$k_0=\frac{A}{l_0}\left\{c_1+2c_2S_0+3c_3S_0^2-3c_cE_0^2+(c_2+3c_3S_0)\frac{Y}{l_0}+c_3\left(\frac{Y}{l_0}\right)^2\right\}.$$

These functions have been simplified by using the approximation

$$E_3=\frac{V_3-\frac{Y}{l_0}V_0}{l_0+Y}\approx\frac{V_3}{l_0+Y},$$

which emphasizes the similarities relative to the functions for piezoelectric transducers.

5. Electrostrictive transducers; remanent polarization

For electrostrictive transducers with remanent polarization the electrical equation is given by using Eq. (48)[1] in Eq. (38)[1] and the mechanical equation is Eq. (51).[1] The required functions are (see Fig. 4 of Ref. 1):

$$\alpha_1=\frac{A}{l_0+Y}\left[\epsilon_1+2e_2\left(S_0+\frac{Y}{l_0}\right)\right],\quad \alpha_2=\frac{\epsilon_2A}{(l_0+Y)^2},$$

$$\beta=\frac{A}{l_0}\left[e_1+2c_a\left(S_0+\frac{Y}{l_0}\right)\right],$$

$$\gamma=\frac{A}{l_0+Y}\left[e_1+2c_a\left(S_0+\frac{Y}{l_0}\right)\right],$$

$$\delta_1=\frac{Ae_2}{(l_0+Y)^2},$$

$$k_0=\frac{A}{l_0}\left[c_1+2S_0c_2+c_2\frac{Y}{l_0}\right].$$

These functions are identical to those for piezoelectric transducers except for the presence of the bias strain, S_0.

6. Magnetostrictive transducers

For magnetostrictive transducers the cases of maintained and remanent polarization are exactly analogous to those for electrostrictive transducers. Thus the functions for magnetostrictive transducers are the same as those for electrostrictive transducers with the bias electric field, E_0, replaced by bias magnetic field, H_0, and ϵ_1, ϵ_2, ϵ_3 replaced by μ_1, μ_2, μ_3 (see Fig. 5 of Ref. 1).

[1] C. H. Sherman and J. L. Butler, ''Analysis of harmonic distortion in electroacoustic transducers,'' J. Acoust. Soc. Am. **98**, 1596–1611 (1995).

[2] C. H. Sherman and J. L. Butler, ''Perturbation analysis of nonlinear effects in moving coil transducers,'' J. Acoust. Soc. Am. **94**, 2485–2496 (1993).

[3] P. G. L. Mills and M. O. J. Hawksford, ''Distortion reduction in moving-coil loudspeaker systems using current-drive technology,'' J. Audio Eng. Soc. **37** (3), 129–147 (1989).

[4] R. S. Woollett, ''Sonar Transducer Fundamentals,'' Naval Underwater Systems Center, Newport, RI, Scientific and Engineering Studies (undated).

[5] A. E. Clark and M. Wun-Fogle, ''A New Method of Magnetostrictivity and Magnetostriction Measurements,'' Trans. IEEE Mag. **25**, 3611–3613 (1989).

[6] W. Klippel, ''Nonlinear Large Signal Behavior of Electrodynamic Loudspeakers at Low Frequencies,'' J. Audio Eng. Soc. **40** (6), 483–496 (1992).

[7] C. H. Sherman and J. L. Butler, ''Harmonic distortion in magnetostrictive and electrostrictive transducers with application to the flextensional computer program FLEXT,'' Image Acoustics, Inc. Report on NUWC Contract No. N66609-C-0985, 30 Sept. 1994.

Field characterization of an air-coupled micromachined ultrasonic capacitance transducer

A. G. Bashford
Department of Engineering, University of Warwick, Coventry CV4 7AL, England

David W. Schindel[a)]
Department of Physics, Queen's University, Kingston, Ontario K7L 3N6, Canada

David A. Hutchins and W. M. D. Wright
Department of Engineering, University of Warwick, Coventry CV4 7AL, England

(Received 21 September 1995; accepted for publication 24 August 1996)

Investigations into the field characteristics of an air-coupled ultrasonic capacitance transducer have been performed for a range of transducer configurations. The field of a 2-MHz bandwidth silicon backplate capacitance transducer has been scanned in air using a 1-mm-diam miniature detector at frequencies of up to 1.5 MHz. The radiated peak sound-pressure field is compared to theory based on a plane piston approach for various driving signals, namely pulsed and tone burst excitation. Aperture modifications, such as an annulus and a zone plate, have also been investigated and the devices have been shown to behave as predicted by theory. © *1997 Acoustical Society of America.* [S0001-4966(97)06012-8]

PACS numbers: 43.38.Bs, 43.58.Vb [SLE]

INTRODUCTION

The transduction of ultrasonic signals in air has gained increased interest recently, due to applications such as noncontact imaging for materials evaluation, distance ranging, and robotics. Such applications require a wide range of ultrasonic frequencies in air, and for materials evaluation, in particular, it is of distinct advantage to operate in the MHz range, where wavelengths in air become less than 1 mm. There are several designs of the ultrasonic transducer that can be used in air, but those useful at high frequencies tend to be based on either piezoelectric or capacitance (electrostatic) effects.

The major problem with conventional piezoelectric ceramic elements (such as PZT) is the large impedance mismatch at the boundary between the piezoelectric element and the surrounding air boundary. The efficiency of these devices for coupling into air can be increased by several methods, such as using a metal diaphragm[1] to act as a vibrating membrane, or by attaching a quarter wavelength impedance matching the layer to the front face. These impedance matching layers need to have a much lower acoustic impedance than the piezoelectric element, and only a small selection of materials can be used, such as silicon rubber,[2] epoxy resin-quartz,[3] and aerogels.[4] Further developments have reduced the impedance mismatch with the use of a 1–3 connectivity piezocomposite active element containing piezoceramic pillars in a polymer matrix.[5] The performance can be optimized for any particular application by changing the piezoceramic pillar fraction, shape, and distribution. In general, however, it is found that such devices are optimized for narrow bandwidth operation, and recent applications have used them in such a mode.[6] Other work has been reported in thick aerospace structures.[7]

The alternative method, which is investigated here, is to use a capacitance design, where a thin membrane is attached to a contoured, conducting rigid backplate. Such devices have been studied extensively in past work, which has investigated their response in relation to the backplate topography for metal backplates. The surface of a backplate for a capacitance transducer can either be textured (as in a roughened surface) or machined (as in a V-groove backplate). A variety of techniques for producing textured backplates have been investigated, including chemical etching, mechanical machining, and shot peening.[8] Some recent research has produced backplates by either depositing polyamide ridges,[9] machining, or etching to form regular patterns such as grooves and holes on or into a polished backplate. The results shown by Refs. 10 and 11 for V-grooved transducers suggest that the resonance frequency can quantitatively be predicted from the dimensions of a V-groove with a 90% confidence limit. It was also shown that a lumped parameter approximation could be used to describe the transducer by an equivalent electrical circuit. Both the resonance frequency and sensitivity could accurately be predicted.

Although such devices are useful, a more extended frequency response is available from new designs of transducer, based on silicon micromachining technology.[12] Various configurations have been described, including devices machined with an *in situ* front membrane.[13] The design to be studied here uses a silicon backplate into which miniature cylindrical pits have been etched chemically, leading to a typical construction as shown in Fig. 1.[14] The details of how such devices are constructed and characterized have been given elsewhere,[14] and some applications described.[15–18] Briefly, the backplate contains a series of uniformly spaced holes,

[a)]Corresponding author. Present address: Institute for Aerospace Research, National Research Council, Montreal Road, Building M-14, Ottawa, Ontario K1A 0R6, Canada. Fax +1-613-952-7136; Electronic mail: david.schindel@nrc.ca

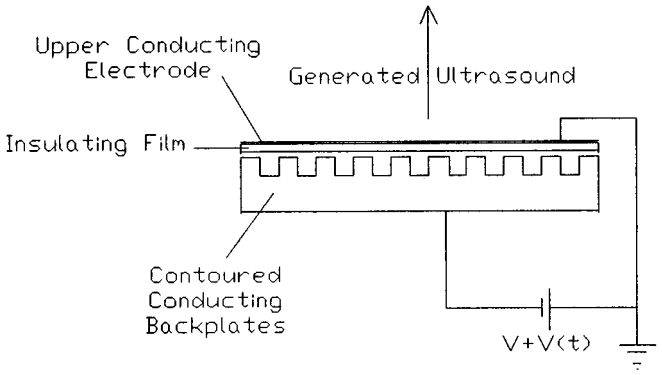


FIG. 1. Schematic diagram of a silicon backplate air-coupled transducer.

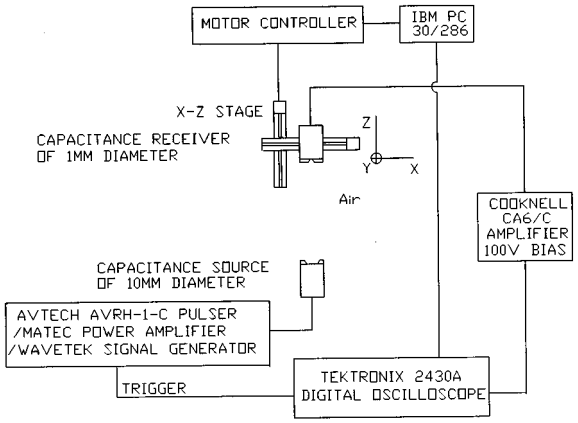


FIG. 2. Schematic diagram of the scanning system used.

etched into the top surface of a flat rigid silicon wafer. Gold is then evaporated onto the etched surface to produce a conducting backplate. A 5-μm Kapton polymer membrane (with a conducting top surface) is placed on top of the etched silicon backplate for both the source and receiver, trapping pockets of air. Applying a transient voltage between the backplate and the grounded front metallized surface of the membrane forces the membrane into motion, thus generating ultrasound in air. Detection can be achieved using an identical device, but a dc polarization voltage is required so that dynamic charges are induced on the backplate by motion of the membrane. Advantages of the above devices include the ease of manufacture, and the ability to reproduce field characteristics by introducing controlled surface roughness. Using the construction of the backplate outlined above, the transducer was found to have a good response into the MHz region, well-damped and with excellent sensitivity. As will be shown, with additional modifications to the aperture it is possible to produce transducers with increased axial intensities, thus making these devices ideal for a range of imaging applications.

As micromachined devices can be constructed with reproducible responses,[14] it is possible to study the radiated fields and waveforms for different drive waveforms and external apertures. The aim of the present paper is thus to compare the peak sound-pressure field of a micromachined silicon backplate capacitance transducer to that predicted by theory for a plane piston, and to extend this to transducer configurations (an annulus and a zone plate) which increase axial responses in air, for different voltage drive signals.

I. APPARATUS AND EXPERIMENT

The pressure waveforms throughout the field of micromachined capacitance transducers were measured in air using a scanned miniature hydrophone. The source transducers under test were derived from a device with a polished flat silicon backplate containing 40-μm-diam etched cylindrical pits of 40-μm depth[14] fitted with a 5-μm Kapton polymer membrane, and which had a 10-mm-diam active area. They were fitted with a grounded brass case, as described in Ref. 14, where additional detail of construction can be found. The scanned receiver was made in the same way, but was fitted with a 1-mm-diam aperture to define a limited detection area. Although bandwidths in excess of 2 MHz were available for both source and receiver, source driving frequencies were restricted to an upper limit of 700 kHz, where the wavelength in air is approximately 0.47 mm. Hence, at the upper frequency limit, the detector is approximately two wavelengths in diameter. In the experiment illustrated in Fig. 2, the apertured receiver was scanned over a two-dimensional horizontal plane throughout the field of the radiating transducer (the x-z plane), with the source radiating horizontally along the z axis. Scanning was achieved using a computer-controlled system which consisted of an IBM model 30 286 PC, controlling a Tektronix 2430A digital oscilloscope and a Modulynx motor controller, the latter driving two DAEDAL linear stages (one for each axis). Each stage was fitted with a synchronous stepping motor having 400 steps per revolution giving an absolute resolution of 0.005 mm and a positional accuracy of ± 0.004 mm/50 mm of travel for the system. The scans were performed with a 0.5-mm spatial resolution in the x (radial) direction and a 1-mm resolution in the z (axial) direction. A typical x-z scan area was 20 mm in the x direction and 50 mm in the z direction, with the field starting at 15 mm from the front face of the transducer membrane. All the scans were performed at atmospheric pressure, room temperature (20 °C) and low relative humidity (30%).

The source was excited with both wide bandwidth transients and tone bursts. A Matec 5100 gated power amplifier and a Wavetek 191 signal generator was used to drive the source with 350-V tone bursts of approximately 20 cycles at various frequencies in the 300- to 700-kHz range, whereas an Avtec AVRH-1-C pulser unit was used to provide a transient pulsed excitation with a width adjustable from 0.1–10 μs. The signal produced by the 1-mm apertured receiver was sent to a Cooknell CA6/C charge sensitive amplifier with a gain of 250 mV/pC, which also supplied a 100-V bias between the backplate and the membrane of the receiver. The nominal capacitance of the device was 650 pF. The signal waveform at each field position was digitized using a Tektronix 2430A oscilloscope, and stored on the PC. The peak sound-pressure variations were obtained from stored wave-

forms by taking the maximum peak-to-peak amplitude of the signal at each designated point.

II. RESULTS AND COMPARISON TO THEORY

A. Fields of a plane piston

Field variations in front of a transducer can be predicted by considering the interference of two components: plane wave from the source, and an inverted wave from the edge (referred to as the ''edge'' wave). Edge waves are diffracted from the aperture or at the boundaries of the source, and travel toward the axis of the transducer, thus interfering with the plane wave.[19–21]

The peak pressure variations in air for any size of aperture can be predicted, using established theory for the radiated field of a plane piston transducer.[19–21] The mathematical model assumes that the pressure at any point across and away from the transducer face can be computed from the interference of the plane and edge components, and also assumes that the front face is vibrating with a uniform amplitude and phase (although variations in these parameters can be included if required). The relevant expression for the time-dependent velocity potential at any given observation point in a half-space can be derived, and further details can be found in the quoted publications. Briefly, the velocity potential impulse response $\Phi(M,t)$ of a plane piston can be written as a combination of plane and edge wave components:

$$\Phi(M,t)=c|A|\underset{\text{(Plane wave)}}{\mathbf{H}(t-t_0)}+(c/2\pi)\Omega(ct)\underset{\text{(Edge wave)}}{\mathbf{H}(t-t_1)\mathbf{H}(t_2-t)}, \tag{1}$$

where M is the observation point, c is the longitudinal wave velocity, t is time, t_1 and t_2 are as defined below, and $\mathbf{H}$ is the Heaviside step function. The value of $|A|$ in the plane-wave component depends on the field position. For M opposite the front face the value is 1. For a field position such that $x=a$, the value drops to 1/2, and $|A|$ is zero at greater radial (x) distances, i.e., beyond the physical edge of the transducer. The time taken for an edge wave to reach the observation point from the closest and furthest edge of the piston is given by t_1 and t_2, respectively. In the case of a circular piston, $\Omega(ct)$ represents the angle of equidistant arc which is an analytical expression derived for the edge wave diffraction impulse response,[19–21] and is given by

$$\Omega(ct)=2\cos^{-1}\{[(ct)^2-z^2+x^2-a^2]/2x[(ct)^2-z^2]^{1/2}\}, \tag{2}$$

where a is the piston radius, x is the radial distance from the center, and z is the axial distance from the piston. The arrival times t_0 (for the plane wave), t_1 and t_2 are given by

$$t_0=z/c, \tag{3}$$

$$t_1=\{(a-x)^2+z^2\}/c, \tag{4}$$

$$t_2=\{(a+x)^2+z^2\}/c. \tag{5}$$

The pressure impulse response $\mathbf{S}(M,t)$ for an arbitrary point M is defined as the differential of the velocity potential impulse response, i.e.,

$$\mathbf{S}(M,t)=\rho_0(\delta\Phi/\delta t)(M,t), \tag{6}$$

where ρ_0 is the density of the medium. Equation (6) is only valid for values of t which are greater than or equal to t_0 and less than or equal to t_2; elsewhere $\mathbf{S}(M,t)=0$. The actual pressure waveform $\mathbf{P}(M,t)$ is then obtained by convolving the simulated motion of the piston [i.e., the velocity waveform $\nu(t)$] with the pressure impulse response for each spatial point:

$$\mathbf{P}(M,t)=\nu(t)\cdot\mathbf{S}(M,t), \tag{7}$$

where $\mathbf{P}(M,t)$ is the pressure response.

An additional factor is attenuation, which causes amplitude degradation and phase shifting. This can be high in air at ultrasonic frequencies, and so this factor needs to be incorporated into the model. The effect can be predicted using previous studies[22,23] and recent work has produced an experimentally derived expression[24] used here, which is

$$\sigma=15.895\times10^{-11}(T/T_0)^{1/2}f^2/(P/P_0), \tag{8}$$

where σ is the peak sound-pressure attenuation per meter (dB/m). The numerical constant in Eq. (8) is an experimentally measured value, determined at a constant temperature of 25 °C (Ref. 24) and f is the frequency (Hz). Here, T is the measured temperature, T_0 is the reference temperature, P is the measured pressure, and P_0 is the reference pressure. In the present work, T and P are taken as the standard reference values (25 °C and 101.325 kpa, respectively), and thus both T/T_0 and P/P_0 are unity. The attenuation was calculated at each field position by taking the Fourier transform of the predicted waveform, to give a frequency spectrum of the signal. This was then modified according to the expression for the calculated attenuation σ, so that the higher frequencies were suppressed more than the lower frequencies. An inverse Fourier transform then gave the resultant attenuated waveform. The resultant spatial variations were found as before by taking the peak-to-peak amplitude of the signal. Note that low humidity levels (30%) were maintained throughout the experiments; while water vapour is known to affect attenuation, its effect is small at high frequencies.[25]

To obtain the required theoretical velocity waveform of the vibrating membrane of the capacitance transducer [$\nu(t)$ in Eq. (7)] it was assumed that a small dc polarized component is always present on the membrane; thus we can assume the devices have a linear response[26] (i.e., the force is proportional to the drive voltage and hence the displacement). The Avtec pulser was used to provide a square wave with a 0.15-μs rise time and a width of 0.7 μs as the driving signal. This was filtered between 100 and 800 kHz (typical bandwidth for a source with a small dc bias), and the velocity waveform hence obtained by the differential with respect to time of the drive waveform. The predicted velocity waveform is shown in Fig. 3(a), and was used as $\nu(t)$ in the above theoretical predictions. The predicted theoretical sound-pressure waveform at an axial distance of 15 mm is shown in Fig. 3(b). This may be compared to the experimental waveform of Fig. 3(c), detected by the miniature scanned detector at the same position, where reasonable agreement with theory is evident.

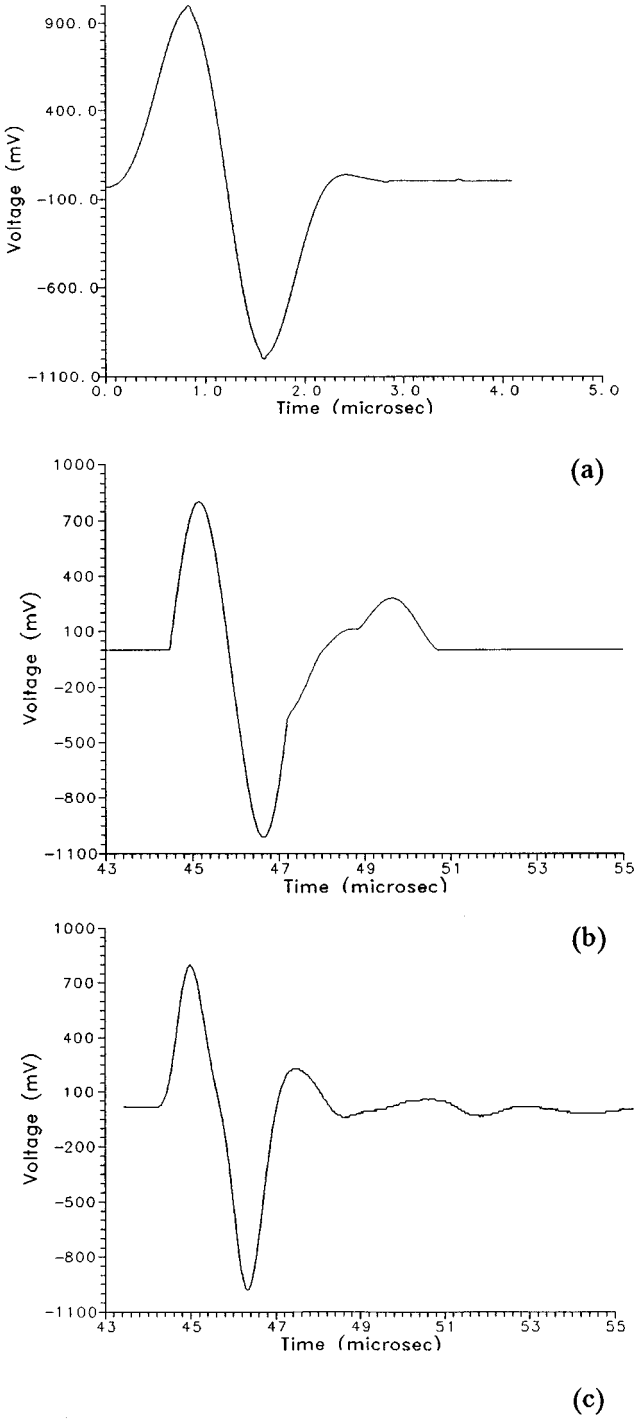


FIG. 3. (a) Predicted velocity waveform of the membrane. (b) Theoretical axial pressure waveform at $z=15$ mm, and (c) corresponding experimental waveform.

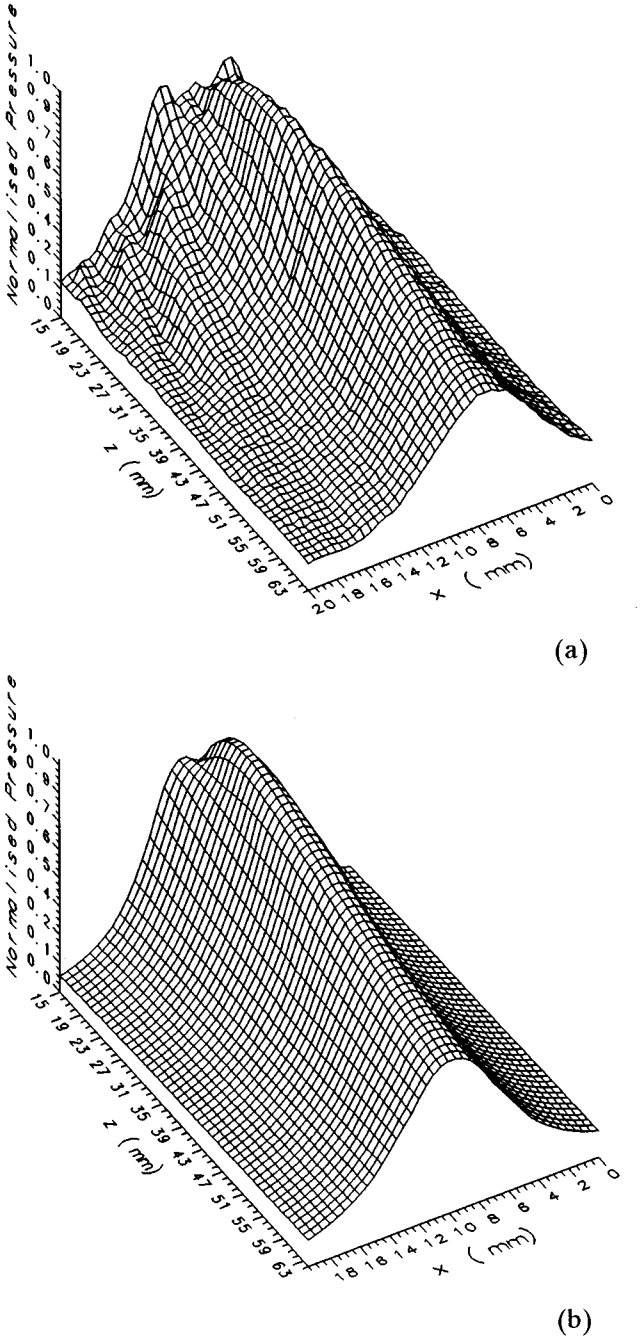
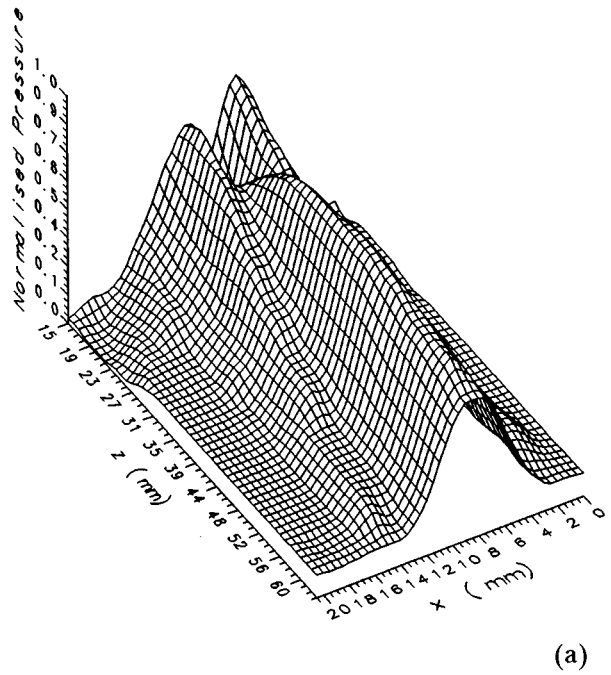


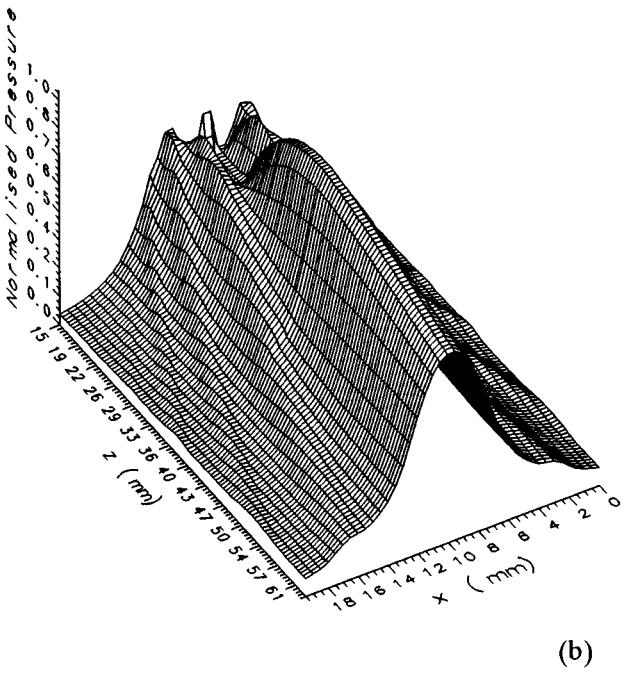
FIG. 4. (a) Experimental peak sound pressure variations, (b) theoretical peak sound-pressure variations, for a plane piston driven by a pulse centered at 316 kHz.

Peak sound-pressure spatial variations were now measured for a 10-mm-diam plane piston, using a 350-V drive pulse from the Avtec pulser unit. The driving pulse width was varied to give a plane-wave pressure component close to the transducer face in the form of a bipolar square wave pulse of variable width, and hence with a predetermined center frequency. The experimental and theoretical variations in peak sound-pressure waveform amplitude throughout the whole scanned field are presented in Fig. 4(a) for a drive excitation centred at 318 kHz. Remembering that this started 15 mm from the front face of the vibrating membrane, the remnants of the edge wave component can be seen clearly at the top left [i.e., at small axial (z) values], converging onto the axis to form the axial maximum by interference with the plane-wave component.

The predicted theoretical peak sound-pressure variations in air for a plane piston were now calculated using Eqs. (1)–(8), to compare to the experiment, the results being shown in Fig. 4(b) for the same scan area of 40 mm×20 mm. Good agreement is evident, with the near-field/far-field boundary positioned at the axial distance of approximately

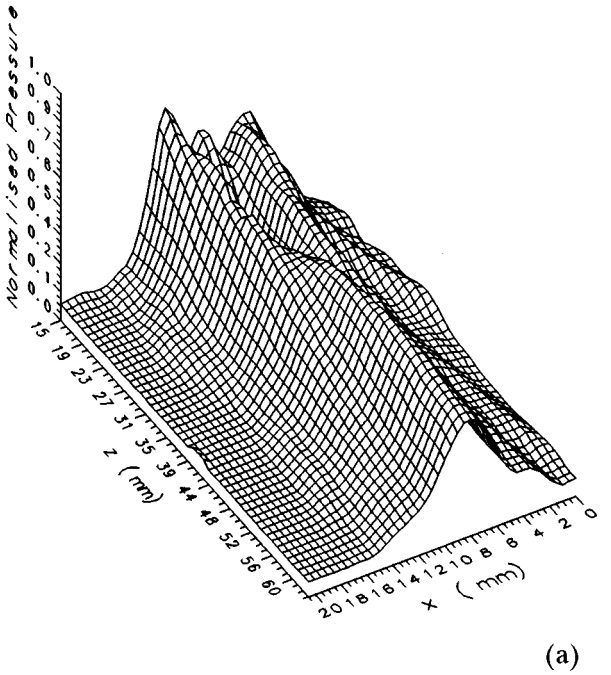


(a)

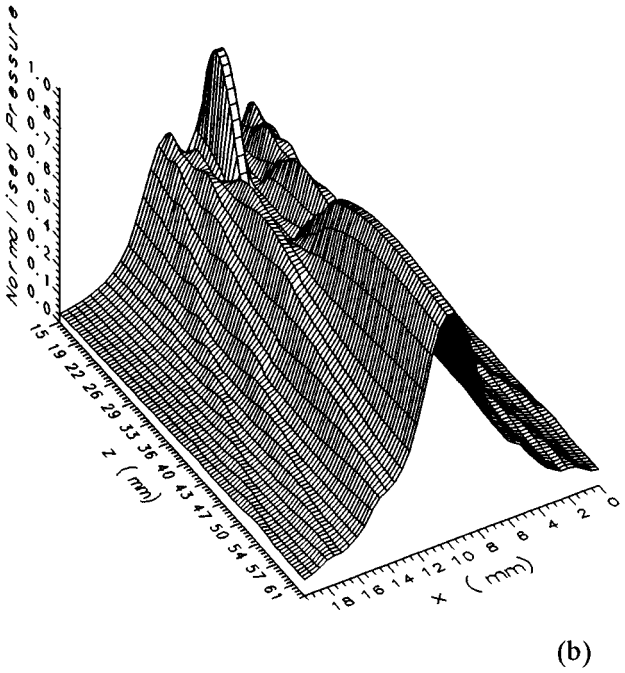


(b)

FIG. 5. (a) Experimental peak sound pressure variations, (b) theoretical peak sound-pressure variations, for a plane piston driven by a 500-kHz tone burst.



(a)



(b)

FIG. 6. (a) Experimental peak sound-pressure variations, (b) theoretical peak sound-pressure variations, for a plane piston driven by a 700-kHz tone burst.

24 mm, as expected from the quoted center frequency.

The above field plotting experiments were now repeated for tone burst excitation, using the Matec gated power amplifier to drive the transducers with a tone voltage signal containing 20 cycles. Figures 5(a) and 6(a) show the resulting spatial variations in peak sound-pressure amplitude in air, for excitation frequencies of 500 and 700 kHz, respectively. The corresponding theoretical plots for a plane piston, again including the correction for attenuation, are also presented in Figs. 5(b) and 6(b). For tone burst generation, the fields are more complicated theoretically, with marked theoretical axial variations in near-field amplitude, and the edge waves converging onto the axis. While some variations are present experimentally, these are not so evident. This is due to the finite aperture used in the receiver. Note, however, that the general features of near-field complexity and increased sidelobe levels are present, as expected for an excitation with a finite number of cycles (i.e., a narrower frequency bandwidth).

For completeness, an additional plot at 1.5 MHz is

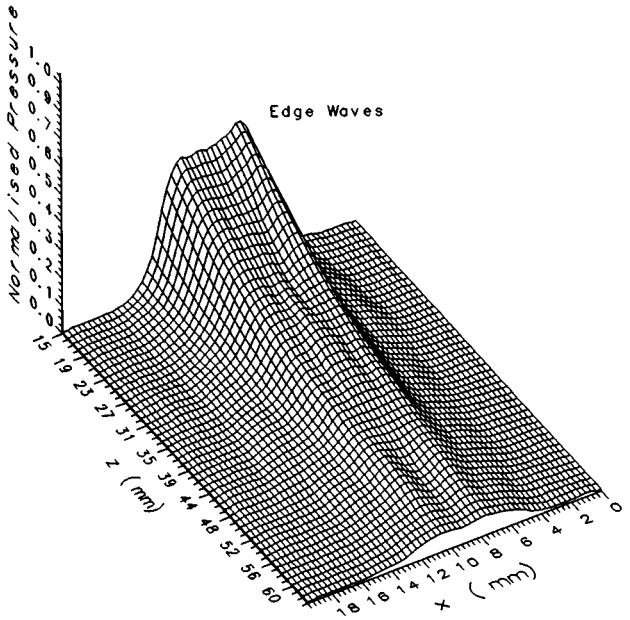


FIG. 7. Experimental peak sound-pressure variations for a plane piston driven by a 1.5-MHz tone burst.

shown in Fig. 7. This has a much faster amplitude decay with distance, but still has the recognizable features of a converging edge wave, and near-field interference effects. This rapid amplitude decay is due to several factors, one being the increased attenuation in air. However, because of the fact that the radiated wavelength is approximately 0.22 mm, and we are using a detector aperture of 1 mm, the scanned hydrophone cannot be assumed to give a wholly representative pressure plot. This increased directivity could cause the field plot to decay artificially, although attenuation at this frequency is probably the major factor.

The well-behaved nature of the emitted beam can be illustrated further by scanning in a plane perpendicular to the beam axis (i.e., in the X-Y plane). This is done in the near-field region for pulsed excitation centered at 500 kHz, 7 mm axially from the source [Fig. 8(a)], and at a distance of 50 mm [Fig. 8(b)] in the far field. Note the relatively smooth response in both cases, with the expected beam spreading in the far field.

Comparison of the above experimental scans with theoretical predictions show good qualitative agreement. However, the miniature detector does not have an absolute calibration, and probably has some directivity effects of its own at high frequencies. In addition, only relative peak sound-pressure amplitudes were plotted, and hence a quantitative comparison was not attempted.

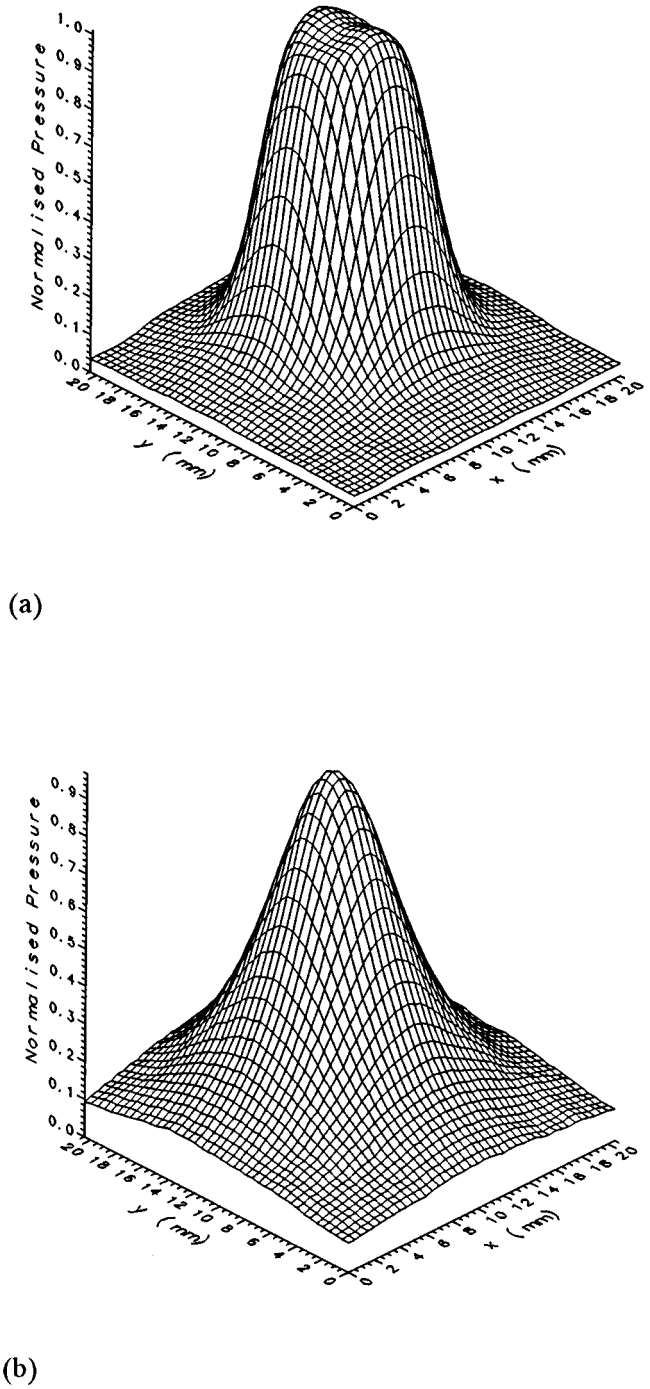
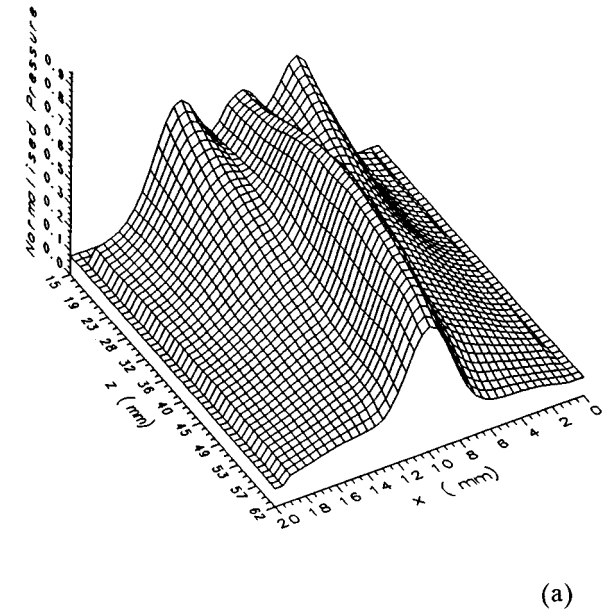


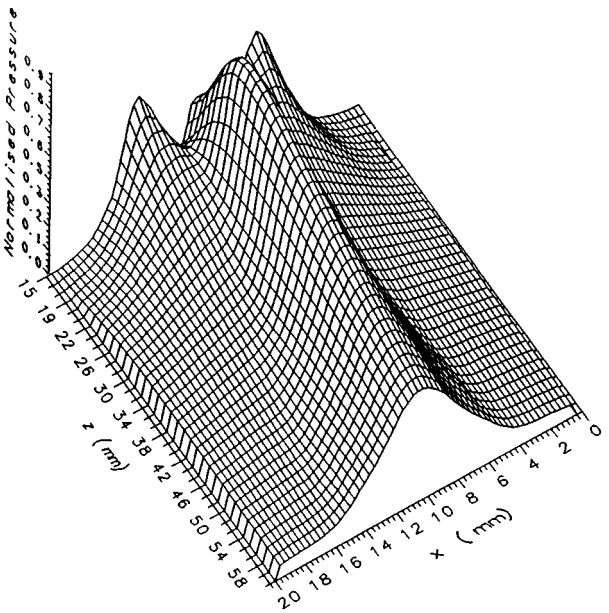
FIG. 8. Experimental peak sound-pressure variations in the X-Y plane for a plane piston transducer driven by a pulse centered at 500 kHz: (a) at 7 mm from the source; and (b) at 50 mm from the source.

B. Annular and zone plate transducers

An air-coupled annular transducer was produced by attaching a concentric 5-mm-diam paper disk to the center of the 10-mm-diam membrane. This absorbed radiated output over the 5-mm-diam central area, producing an annulus with an active inner diameter (d) of 5 mm and an outer diameter (D) of 10 mm. The peak sound-pressure field for this annulus was determined experimentally (as described above) for pulsed excitation with the Avtec pulser, giving a signal with a center frequency of 318 kHz. The results are shown in Fig. 9(a). Note the tendency for axial focusing, with a gradual increase in axial amplitude toward the maximum, accompanied by a smooth decrease in the direct radiation from the active area of the annulus. Similar features are seen in the theoretical plot, Fig. 9(b). This was produced by predicting the waveforms produced at each field point by two separate plane pistons, one of 10 mm diameter and the other of 5 mm



(a)



(b)

FIG. 9. (a) Experimental peak sound-pressure variations for an annulus, (b) theoretical pressure field plot for an annulus, using a pulse centered at 316 kHz ($d=5$ mm $D=10$ mm).

diameter. At each spatial point, the waveforms are then subtracted to predict the waveform of the annulus. Note that in the theoretical plot, the features are more sharp and less attenuated than those observed experimentally. This is again thought to be due, in part, to the finite receiver aperture. There also seems to be greater attenuation than that predicted by theory, a phenomenon that merits further investigation.

A zone plate consists of a series of active concentric rings (or zones), separated by areas that do not radiate. The zones are positioned such as to cause the radiated contributions from each zone to interfere constructively on-axis at the focal position, thus increasing the pressure intensity at the focus. The radial extent of each zone (R_M) is determined by

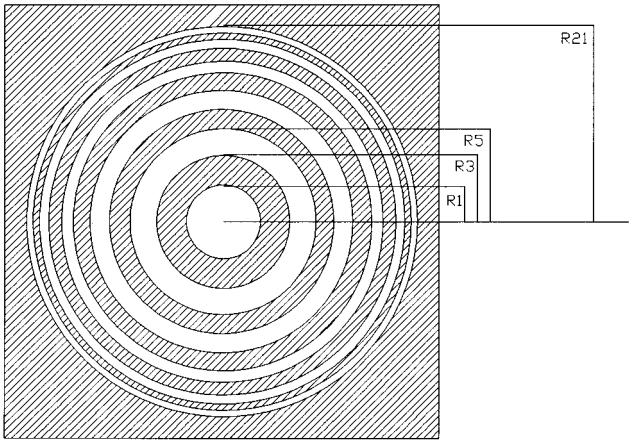


FIG. 10. Schematic diagram of a zone plate.

selecting the required axial focusing position Z_0 from the face of the zone plate (i.e., the focal length), and then positioning the edge of each zone to be at a distance $Z_0+M\lambda$ from position Z_0 where λ is the acoustic wavelength, and $M=1,3,5,7...$. As shown by Ref. 27, an approximate expression for the radial position of the edges of each zone (R_M) can be written as

$$R_m=\sqrt{\frac{M\lambda}{2}\left(Z_0+\frac{M\lambda}{8}\right)}. \tag{9}$$

Figure 10 shows a schematic diagram of a zone plate with six zones, with a predicted focal length from Eq. (9) of 7.5 mm when driven at a frequency of 500 kHz in air. The radii R_1, R_3, etc. are shown on the figure, with the unshaded areas representing regions where there is radiation; elsewhere, it is assumed that no emission occurs. In the present example, a central disk and five concentric annuli of emission were chosen, with the outer radius of the last zone (R_{21}) being 8 mm. This was thought to be a good compromise between focusing efficiency and ease of manufacture. The zone plate was etched chemically from sheet stainless steel of 0.13 mm thickness, using a computer-derived mask and photolithography. Concentric rings were held together by

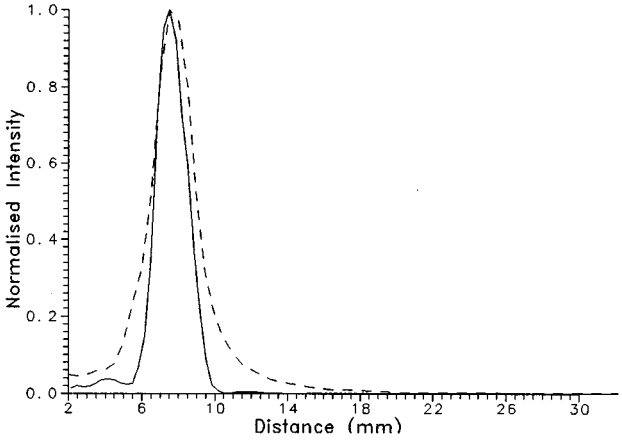


FIG. 11. Experimental (solid line) theoretical (dashed line) acoustic intensity axial plot, for a plane piston transducer with the zone plate.

thin radial bridges included in the original mask (not shown). This zone plate was placed across the transducer aperture, at a distance of 5 mm from the radiating membrane. Thus when a 500-kHz tone burst was applied to a transducer fitted with the zone plate, the axial pressure would be expected to increase to a maximum on axis at 7.5 mm from the plate (or 12.5 mm from the transducer face).

Figure 11 shows the experimental axial acoustic intensity field from the transducer, driven with a 500-kHz tone burst and fitted with the zone plate of Fig. 10. Note that now 2-D intensity variations (the square of the pressure amplitude) along the axis were plotted, to illustrate more graphically the focusing properties of these devices; due to the high noise floor on the experimental plot, pressure field variations off-axis did not correspond well to the theoretical result (this needs to be further investigated). It is evident, however that the transducer assembly has focused in air at the expected distance of 7.5 mm from the zone plate. Note the sharp focus, and the rapid axial decay of pressure amplitude at larger axial distances. Theoretically, the response of the zone plate can be simulated using the response of annuli as above. Here, the zone plate can be considered as a set of concentric annuli surrounding a central radiating disk; hence, the response can be found by adding the separate contributions of the central disk and each annulus in turn, for each field point. When this is done, the pressure field expected theoretically for 500-kHz tone burst generation in air is also shown in Fig. 11. In agreement with the predictions of Eq. (9), and the experimental observations of Fig. 11, the pressure amplitude peaks at 7.5 mm, and exhibits a rapid decay in amplitude at greater axial distances.

III. CONCLUSIONS

The experimental spatial variations in peak sound-pressure from an air-coupled micromachined capacitance transducer have been demonstrated to show many of the characteristics of a plane piston radiator. This was seen to be the case for both transient and tone burst voltage excitation. The theoretical model was modified to include frequency-dependent attenuation, and this was seen to improve agreement between theory and experiment. With additional modifications to the plane piston model, it was possible to predict the output from both annular and zone plate transducers, and this was again compared to experiment with reasonable agreement, indicating that axial pressures could be increased by both methods.

ACKNOWLEDGMENT

This work was funded by the Engineering and Physical Sciences Research Council (EPSRC).

[1] M. Babic, ''A 200-kHz Ultrasonic transducer coupled to the air with a radiating membrane,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **UFFC-38**(3), 252–255 (1990).
[2] J. D. Fox and B. T. Khuri-Yakub, ''High frequency wave measurements in air,'' Proc. IEEE Ultrason. Symp. 581–592 (1983).
[3] P. Kleinschmidt and V. Magori, ''Ultrasonic remote sensors for non-contact object direction,'' Siemens Forsch. Entwicklungsber. **10** (2), (1981).
[4] O. Krauβ, R. Gerlach, and J. Fricke, ''Experimental and theoretical investigations of SiO_2-aerogel matched piezo-transducers,'' Ultrasonics **32**(3), 217–222 (1994).
[5] D. Reilly and G. Hayward, ''Through air transmission for ultrasonic non-destructive testing,'' Proc. 1991 IEEE Ultrason. Symp. 763–766 (1991).
[6] R. Farlow and G. Hayward, ''Real-time ultrasonic techniques suitable for implementing non-contact NDT systems employing piezoceramic composite transducers,'' Insight **36**, 926–935 (1994).
[7] C. M. Fortunko, J. O. Strycek, and W. A. Grandia, ''Nondestructive testing of ''thick'' aerospace honeycomb structures using through-transmitted ultrasonic guided waves,'' Rev. Prog. Quant. NDE **8B**, 1643–1650 (1990).
[8] W. S. H. Munro and C. Wykes ''Arrays for air-borne 100kHz ultrasound,'' Ultrasonics **32**(1), 57–63 (1994).
[9] M. J. Anderson, J. A. Hill, C. M. Fortunko, N. S. Dogan, and R. D. Moore ''Broadband electrostatic transducers; modelling and experiments,'' J. Acoust. Soc. Am. **97,** 262–272 (1995).
[10] H. Carr and C. Wykes, ''Diagnostic Measurements in capacitance transducers,'' Ultrasonics **31,** 13–20 (1993).
[11] M. Rafiq and C. Wykes, ''The performance of capacitive ultrasonic transducers using v-grooved backplates,'' Meas. Sci. Technol. **2,** 168–174 (1991).
[12] K. Suzuki, K. Higuchi, and H. Tanigawa, ''A silicon electrostatic ultrasonic transducer,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **36,** 620–627 (1989).
[13] M. I. Haller and B. T. Khuri-Yakub, ''A surface micromachined electrostatic ultrasonic air transducer,'' Proc. 1994 Ultrason. Symp. 1241–1243 (1994).
[14] D. W. Schindel, D. A. Hutchins, L. Zou, and M. Sayer, ''The design and characterization of micromachined air-coupled capacitance transducers,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42,** 42–50 (1995).
[15] D. W. Schindel and D. A. Hutchins, ''Applications of micromachined capacitance transducers in air-coupled ultrasonics and nondestructive evaluation,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42,** 51–58 (1995).
[16] D. W. Schindel and D. A. Hutchins, ''Through-thickness characterization of solids by wideband air-coupled ultrasound,'' Ultrasonics **31,** 11–17 (1995).
[17] D. A. Hutchins, W. M. D. Wright, and D. W. Schindel, ''Ultrasonic measurements in polymeric materials using air-coupled transducers,'' J. Acoust. Soc. Am. **96,** 1634–1642 (1994).
[18] B. Hosten, D. A. Hutchins, and D. W. Schindel, ''Measurement of elastic constants in composite materials using air-coupled ultrasonic bulk waves,'' J. Acoust. Soc. Am. **99,** 2116–2123 (1996).
[19] P. R. Stepanishen, ''Transient Radiation from Pistons in an Infinite Planar Baffle,'' J. Acoust. Soc. Am. **49,** 1630–1638 (1971).
[20] J. C. Lockwood and J. D. Willette, ''High-Speed methods for computing the exact solution for the pressure variations in the nearfield of a baffled piston,'' J. Acoust. Soc. Am. **53,** 735–741 (1973).
[21] D. A. Hutchins and G. Hayward, ''The radiated field of ultrasonic transducers,'' in *Physical Acoustics, Vol. XIX*, edited by R. N. Thurston and A. D. Pierce (Academic New York, 1990), pp. 2–134.
[22] R. Hickling and S. P. Marin, ''The use of ultrasonics for gauging and proximity sensing in air,'' J. Acoust. Soc. Am. **79,** 1151–1159 (1986).
[23] H. E. Bass, L. C. Sutherland, and L. Evans, ''Atmospheric absorption of sound: Theoretical predictions,'' J. Acoust. Soc. Am. **51,** 1565–1572 (1972).
[24] L. J. Bond, C. Chiang, and C. M. Fortunko ''Absorption of ultrasonic waves in air at high frequencies (10–20 MHz),'' J. Acoust. Soc. Am. **92,** 2006–2015 (1992).
[25] H. E. Bass, L. C. Sutherland, and A. J. Zuckerwar, ''Atmospheric absorption of sound: Update,'' J. Acoust. Soc. Am. **88,** 2019–2020 (1990).
[26] D. W. Schindel and D. A. Hutchins, ''The capacitance transducer as a standard ultrasonic source in solids,'' J. Acoust. Soc. Am. **97,** 1650–1659 (1995).
[27] M. Z. Sleva, W. D. Hunt, and R. D. Briggs, ''Focusing performance of epoxy and air-backed polyvinylidene fluoride Fresnel zone plates,'' J. Acoust. Soc. Am. **96,** 1627–1633 (1994).

Study on the vibrational characteristics of ultrasonic transducers using tapered piezoelectric ceramic elements

Rupa Mitra and T. K. Saksena
Ultrasonics Section, National Physical Laboratory, Dr. K. S. Krishnan Road, New Delhi-110 012, India

(Received 19 November 1994; revised 29 July 1996; accepted 3 September 1996)

Any nonuniformity in the flatness of a crystal element used in an ultrasonic transducer may lead to significant variations in transducer response. An accurate estimation of such variations could obviously be a measure of the nonuniformity involved, whatever its extent may be. Studies are reported on the effect of nonparallelism between two major plane surfaces of the crystal element on transducer response by studying the normal displacement amplitude pattern and its distribution over one of the vibrating (plane) surfaces at resonance. Measurements are carried out in air using a phase-locked laser interferometry technique where the effect of tapering in disk thickness on the transducer performance is manifested through the variation in vibrational patterns. Results of these measurements are compared with those obtained from electrical (conductance) response to highlight the importance of the former. Other comparison measurements, such as the conventional method of determination of acoustic response and pulse waveforms of the radiated signals in water, are also reported as a function of tapering in disk thickness. A correlation is discussed between results obtained from different measurements. © *1997 Acoustical Society of America.* [S0001-4966(97)02401-6]

PACS numbers: 43.38.Fx, 43.38.Ar, 43.35.Yb [SLE]

INTRODUCTION

Uniform thickness piezoelectric crystals are employed in conventional ultrasonic transducers for use as transmitters and receivers. Such transducers produce a narrow-band transmit response when operated at resonance. A slight variation of any crystal parameter such as disk thickness that may be introduced in the processing of the disk, however, leads to a significant variation in transducer response,[1–8] which in turn could help in detecting the extent of nonuniformity involved in transducer dimensions.

Tapering of the transducer could be caused inadvertantly or could be introduced with a definite purpose. In the former case, the nonuniformity in thickness is small, whereas in the latter it may assume larger proportions. Nowadays the tapered transducers are finding increasing applications in NDT and medical fields. For example, a linearly tapered piezoelectric disk can be used as a broadband transducer with nearly uniform transmit response[5–10] even without employing a lossy backing layer. Using such tapered disks, the complicated and relatively expensive backing treatment is avoided. Moreover, localized concentration of vibration on a crystal can be achieved fairly easily by tapering the disk. The displacement amplitude of such a vibrating disk has been shown to be higher at the thicker part even for a very slight taper. As a specific case, a point source transducer may be designed using a centrally thick tapered transducer. In the cw mode, a varying frequency can selectively excite resonant portions of tapered ceramic and thereby achieving frequency-controlled acoustic beam translation. A tapered phased array has been found to offer an accurate and efficient heating of deep-seated tumors with the least heating of surrounding normal tissues and thus is useful for hyperthermia treatment.[11] An accurate knowledge of the vibrational characteristics of tapered transducers would therefore be essential to obtain the right performance of the transducer and can be used as an important design parameter.

Although a geometric measurement of taper is easy to make, it cannot provide precise information about the transducer construction because of factors such as bonding, clamping, etc., in transducer construction. One has to take recourse to an acoustical method. The measurement of acoustic pressure has often been used for such studies. In the present work the object of investigation is to examine whether vibrational amplitude measurements can be utilized to provide necessary information about the tapered transducer. The point-to-point study by vibration amplitude measurements would complement the overall study made by acoustic pressure measurements and give a connective insight for better understanding of design perspectives. Moreover, whereas the acoustic measurements give only a measure of the damping of transducer performance caused by tapering, the vibration measurements provide a number of parameters that can be used to distinguish transducers even with very slight difference of thickness gradient.

Vibrational characteristics of transducers having a lack of parallelism between the major plane surfaces of piezoceramic elements can be accurately and conveniently studied by the measurement of the normal displacement amplitude patterns near resonance using phase-locked laser interferometry technique with high accuracy (3%–7%).[12,13] The main feature of this technique is that the effect of relatively low-frequency (up to 10 kHz) high amplitude ($\sim\mu$m) environmental vibrations on the low amplitude ($\sim$nm) high-frequency ($\sim$MHz) ultrasonic vibrations (to be measured) is minimized.[12] It is proposed in the present work to study the vibrational characteristics of tapered transducers along with acoustic response and pulse waveforms in the vicinity of resonance or its overtones and to examine the effect of ta-

pering in disk thickness on transducer response. Transducer conductance is also studied in the preassigned frequency range and results are compared with those obtained from vibrational measurements. An attempt is then made to correlate the variation in transducer (disk) dimensions with corresponding variations in the transducer response thus obtained. Such studies could be of significant use in the estimation of design configuration of specific ultrasonic transducers discussed above and in the detection of slight tapering that may be introduced inadvertantly. Each transducer used employs longitudinal thickness drive vibrations of thick wedge-shaped disks of lead-zirconate titanate ceramic.

I. EXPERIMENT

A. Laser interferometer system

Figure 1 presents the block diagram of the phase-locked laser interferometer system that employs an He–Ne laser beam of wavelength 632.8 nm and power 5 mW. The basic arrangement is very similar to that of the Michelson interferometer, but modified for compensation of the effect of low-frequency ($\lesssim$10 kHz) ambient vibrations in the signal beam.[12,14–16] The technique utilizes the electro-optic Pockels effect[17,18] by introducing a Pockels cell unit consisting of suitably oriented KD*P crystals in the path of circularly polarized laser light. Use of this unit and an appropriate feedback system in between the reference and signal beam circuitries makes it possible to shift the frequency of the reference beam by the same amount as the shift generated in the signal beam by low-frequency ambient vibrations.[15,17] The feedback circuit, having a characteristic time constant which is small compared to the time of one cycle of environmental vibration, is unable to respond to the high-frequency ultrasonic signals to be detected at the output of the interferometer.[12] As the signal and reference beams are orthogonally polarized, the amplitude fluctuations from interference can be detected by a second polarizing beam splitter placed at 45° to the plane of polarization of the beams. With correct gain in the feedback loop, the interferometer is balanced and very sensitive to small ultrasonic displacements (ξ) in the range 0–150 nm are found, which are computed using the following relation

$$\xi(t)=\sin^{-1}[(V_i/V_0)\cdot(\lambda/4\pi n)], \tag{1}$$

where V_i is the output signal of the interferometer. V_0 is the reference or fringe voltage, λ is the wavelength in vacuum of the laser light used (632.8 nm), and n is the refractive index of the medium concerned. For $\xi \ll \lambda/4$. Equation (1) may be expressed as

$$\xi(t)=[(V_i/V_0)\cdot(\lambda/4\pi n)], \tag{1'}$$

which has been used in the present study to compute the required displacement amplitude.

B. Measurements of vibrational amplitude, acoustic output, and electric conductance

To study the vibrational amplitude patterns, each transducer was excited with a Wavetek (model 166) function generator connected with a pulse generator (Systronics, model 1110/D) to generate the required sinusoidal tone-burst signal of 0.8-ms duration. The excitation voltage is mentioned in each corresponding figure (Fig. 3). As the transducers produced relatively weaker peak amplitudes at higher harmonics, the excitation voltage at those frequency ranges as increased to obtain prominent vibration amplitude. The point at which the excitation voltage was increased is denoted by a dashed vertical line in the figure.

For the measurement of acoustic response in water at a

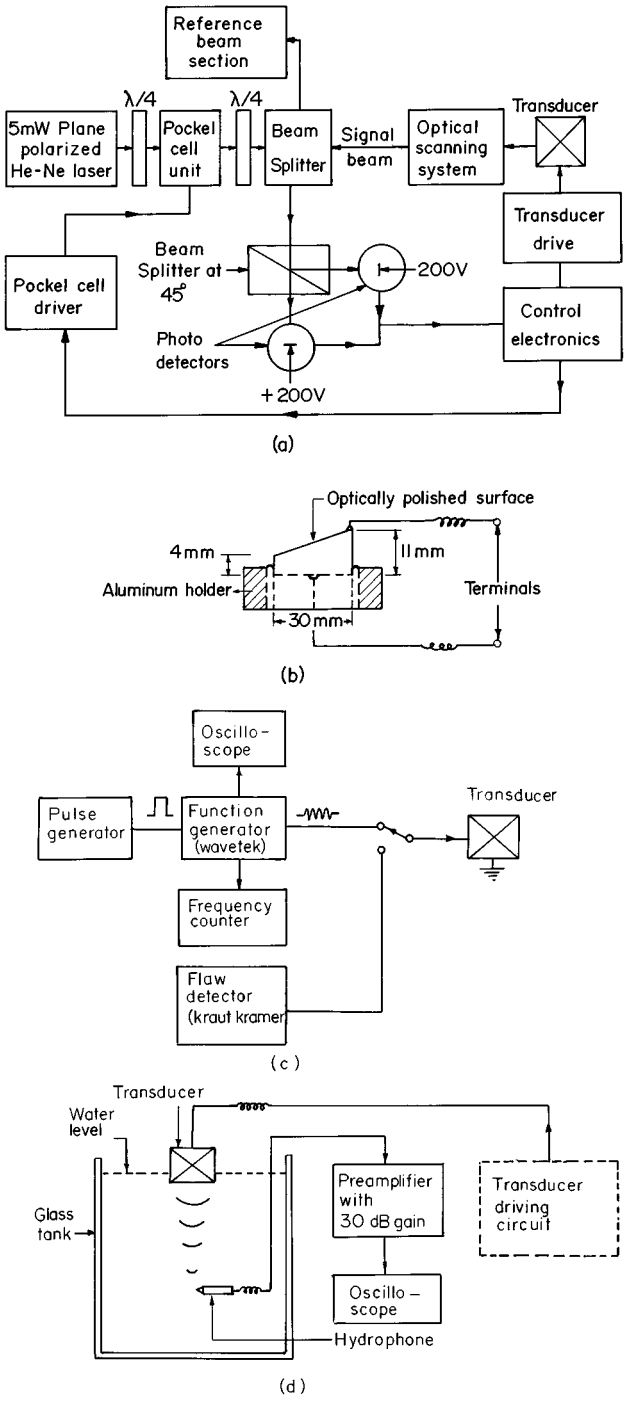


FIG. 1. (a) Schematic block diagram of optical phase-locked laser interferometer. (Quarter wave plates are indicated by λ/4.) (b) Sectional view of mounted transducer 4. (c) Schematic of excitation of the transducer. (d) Schematic of the measurement of acoustic response of the transducer using a probe hydrophone.

TABLE I. Physical parameters of transducer elements. Diameter (D) of each disk=30±0.02 mm.

Transducer serial number	Tapering in disk thickness (t) in mm	Thickness gradient ($m=\Delta t/D$)	Charge const. (d_{33}) pC/N
1	11.5 (const.)	0	320
2	11 to 12	0.03	336
3	6 to 10	0.13	310
4	4 to 11	0.23	298

far field axial point, the transducer was excited by the similar process [Fig. 1(c)], as described above, and a PVDF probe hydrophone (Medisonics, needle type) connected with a preamplifier (Medisonics Mk II, 30-dB gain) and an oscilloscope was used as a receiver [Fig. 1(d)]. The tip of the hydrophone was placed at an axial distance of more than $2a^2/\lambda$ from the transducer surface, where ''a'' and ''λ'' are, respectively, the radius of the transducer and the wavelength of radiation in water. The transducer impulse response was studied by driving it with spike signals obtained from a Krautkramer (model USIP-11) flaw detector operating at specific frequencies. For acoustical measurements the transducer was connected through a two-way key [Fig. 1(c)] to arrange for the excitation either by tone-burst signals or spike signals. The vibration amplitude measurements were taken in air and the transmitting response measurements in water because of the high absorption coefficient and the consequent absorption correction in air.

Electrical conductance of the transducers has been studied in the frequency range 100–1000 kHz covering the fundamental and higher harmonics, at a step frequency of 1 kHz. A vector impedance analyzer (Hewlett-Packard 4192A) was used for this measurement which provides the maximum excitation voltage of 1.1 V to the transducer concerned.

C. Fabrication of the transducers

Each transducer consists of an axially polarized cylindrical disk of lead-zirconate-titanate ceramic with zero or linear tapering in thickness. The ceramic disks were fabricated at the National Physical Laboratory, India, and have the elastic and piezoelectric properties equivalent to those of standard PZT-5A ceramic having the dielectric constant $k_3^T=1700$ and dielectric loss factor tan $\delta=0.025$. The sectional view of the transducer with highest tapering is shown in Fig. 1(b). One of the two major (plane) surfaces has been optically polished and then silver plated for use in the interferometer. Two electrical terminals have been taken from the two silver plated surfaces where an air-drying conducting silver paint was used to make electrical contacts. The dimensions and piezoelectric charge constants (d_{33}) of the four disks employed in the study are tabulated in Table I. The study thus comprises the evaluation of performance characteristics of ultrasonic transducers with a different degree of tapering in disk thickness.

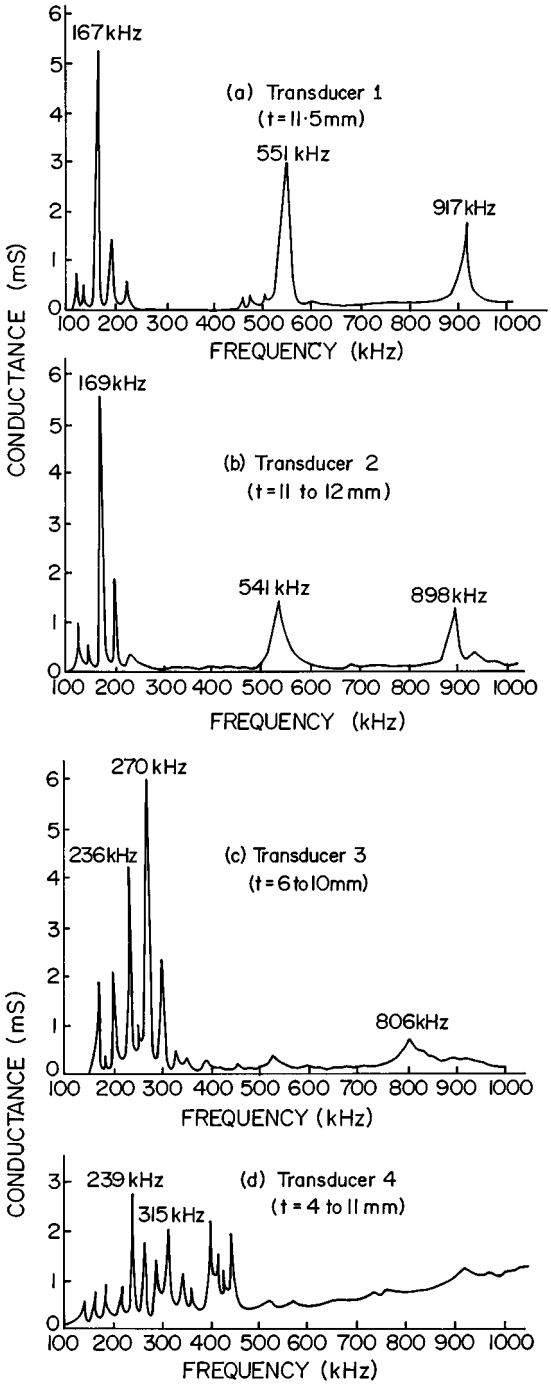


FIG. 2. Variation of conductance with frequency for different tapered transducers.

II. RESULTS AND DISCUSSSIONS

Studies were carried out on the vibrational characteristics together with the electrical as well as acoustical behavior of the four transducers in the frequency range 100–1000 kHz, which covers the fundamental resonance and higher-order overtones. Results of measurements of various parameters are reported in the following subsections.

A. Study on electrical conductance (G)

Figure 2 represents the conductance patterns of the four transducers in the preassigned frequency range. Transducer 1

TABLE II. Comparison of the electrical responses of different tapered thickness transducers.

Transducer serial number	Tapering in disk thickness in mm	No. of secondary peaks near resonance	Ratio in dB between main peak ampl. to av. of secondary peak ampl.	Frequency spreading of peaks near resonance (kHz)
1	11.5 (const.)	2	14.5	80
2	11 to 12	2	13.0	95
3	6 to 10	4	5.9	170
4	4 to 11	6	4.8	315

with no tapering in disk thickness (with a tolerance of ± 0.02 mm) shows a fairly monochromatic response [Fig. 2(a)] with a prominent resonance peak at 167 kHz which is surrounded by two smaller peaks (-14.5 dB with respect to the main one) appearing in a frequency band of 80 kHz (Table II). Transducer response at the third harmonic (at 551 kHz) appears with the peak amplitude 5 dB lower than that of the fundamental one. A fifth harmonic peak at 917 kHz (-9.9 dB) is also observed.

The conductance pattern of transducer 2 with slight tapering in disk thickness (the thickness gradient being 0.03) produces the main resonance peak at 169 kHz [Fig. 2(b)] accompanied by two close-by subsidiary peaks (-13 dB reference to the main one) of slightly higher amplitude compared to that of the corresponding ones appearing in the previous case. Moreover, the frequency spread for smaller peaks (95 kHz) is also higher compared to the previous case. Transducer response at the third harmonic is found to be much weaker (-11.4 dB) compared to that at the fundamental and also with reference to the corresponding amplitudes observed for zero tapering. The fifth harmonic peak (at 898 kHz) is still weaker (-12 dB). Electrical measurements thus present the differences in transducer characteristics for transducers with slight variation of tapering, through the amplitude ratio between main-to-subsidiary peaks near fundamental resonance and through that between fundamental-to-harmonic ones. Unlike the two cases discussed, the conductance pattern of transducer 3 with a thickness gradient of 0.13 presents a number of smaller peaks [Fig. 2(c)] in the fundamental resonance region in a frequency spread of 170 kHz (Table II) which is much broader than that observed in the two cases discussed earlier. The transducer thus shows a multiresonance characteristic at the fundamental with the main peak at 270 kHz and almost negligible response at higher harmonics. An introduction of an additional damping to the transducer by employing a suitable backing could result in a flatter vibrational pattern and thus lead to a broadband response at the fundamental region. The transducer with highest tapering in disk thickness (the gradient being 0.23), transducer 4, produces several secondary peaks of nearly the same order of amplitude which appear in fundamental resonance region in a frequency range of 315 kHz. The transducer shows no response at higher harmonics. It is further noticed that the peak value of the fundamental resonance conductance (G_m) in this case is much less (<6 dB) than that observed in each of the earlier three cases studied, although the values of piezoelectric charge constant (d_{33}) of all the disks are of the same order (Table I). The low value of G_m could be attributed to the processing condition for the present disk. The response of this transducer is observed to have the widest frequency spread for the discrete peaks near resonance, among all the four transducers concerned.

The electrical response of the transducers thus makes it obvious that a tapering in disk thickness results in the multiresonance pattern, i.e., an inhomogeneous broadening in the frequency-domain response. It also leaves a transducer less active at the harmonic region.

B. Study on the vibrational amplitude (ξ) characteristics

The displacement amplitude patterns of the four transducers in the preassigned frequency range, measured on the central points of radiating surfaces, are depicted in Fig. 3. Each transducer was excited with a sinusoidal tone-burst signal of duration 0.8 ms and a p–p driving voltage (V_D) as mentioned in each corresponding figure.

The displacement amplitude pattern of transducer 1 (no tapering) [Fig. 3(a)] produces nearly a similar sort of response as obtained from the conductance behavior [Fig. 2(a)]. Apart from the main resonance peak at 167.1 kHz (with amplitude $\xi_m = 56.8$ nm or $\xi_m/V_D = 2.03$ nm/V) and a few discrete peaks in its vicinity, there appears a continuous distribution of average amplitude of 7.5 nm in the frequency range 130–200 kHz (Table III). Such a vibrational characteristic implies that the transducer does vibrate up to an observable extent over this frequency band, which cannot be detected from the corresponding electrical measurement [Fig. 2(a)]. This could be mainly due to the low excitation voltage (1.1 V, which is the maximum limit of the experimental setup for the HP 4192A impedance analyzer) applied in the latter case. The displacement amplitude measurement is thus found to be more nearly precise and informative compared to the electrical measurements. It may be mentioned that a uniform parallel-plate transducer is expected to produce a sharp resonance with negligible spurious vibrations in its vicinity. It thus appears from the vibrational amplitude pattern of transducer 1 that there exists some nonuniformity in the disk employed for it.

Figure 4 displays the traces of displacement amplitude profiles on resonating transducer surfaces (at main resonance frequencies) as measured along a diameter connecting points on maximum and minimum thickness. The increasing thickness or the positive thickness gradient ($+ve\Delta t$) is indicated in Fig. 4(b)–(d) through arrow marks. The displacement amplitude profile for transducer 1 [Fig. 4(a)] is observed to be nonuniform, although axially symmetric with a central maximum and no side lobes. Similar forms of profiles have been reported in the literature by earlier authors like Shaw,[19] Lypacewicz and Filipczinski,[20] and Arnold and Martner[21] for short barium titanate and PZT transducers.

Transducer 2, the slightly tapered one (with gradient 0.03), produces the fundamental resonance (at 169.3 kHz) with $\xi_m/V_D = 1.61$ nm/V [Fig. 3(b)] which is accompanied by a few discrete peaks along with a continuum of average displacement amplitude 6 nm in the frequency range 120–

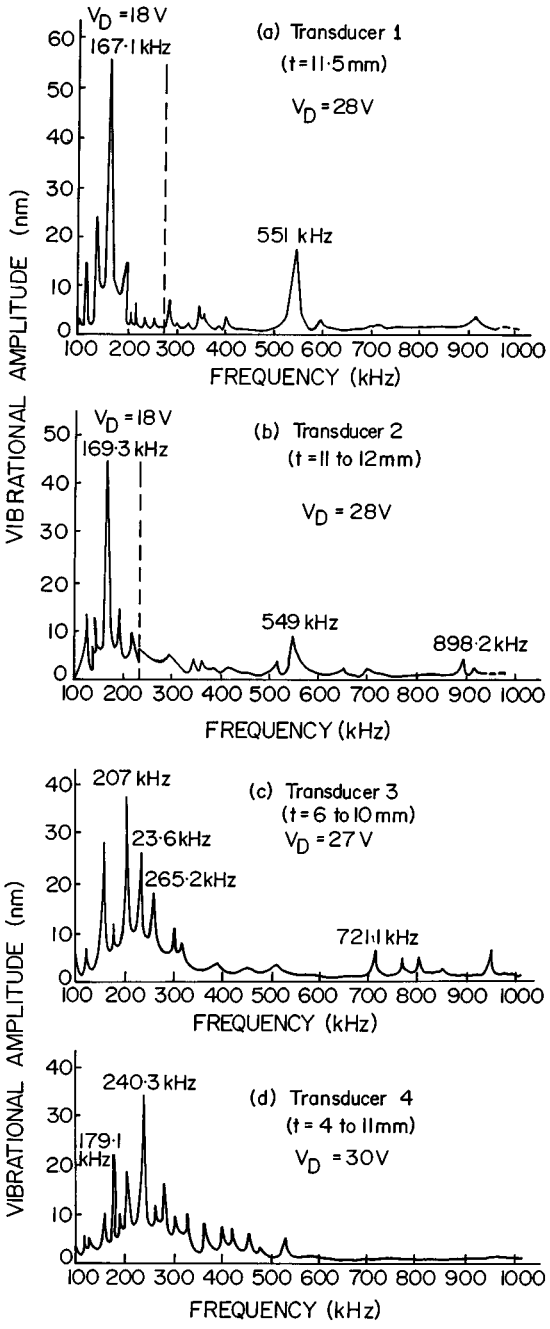


FIG. 3. Variation of vibrational amplitude with frequency for different tapered transducers.

230 kHz (Table III). A comparison between Fig. 3(a) and (b) reveals a spreading of the continuum of transducer vibration near resonance and an observable lowering of higher harmonic amplitude caused by slight tapering. In a more nearly ideal case of transducer 1, which in the present case is likely to have some nonuniformity in its configuration, the above comparison would produce a clearer distinction between the vibration patterns of the two transducers concerned. The displacement amplitude profile on the surface of transducer 2 not only represents a nonuniform distribution of ξ but an axial asymmetry is noticed with more concentration of vibration in the thicker part of the disk. An identical distribution was theoretically estimated by Loutzenheiser and Denkmann[1] for a truncated, linearly tapered quartz crystal

TABLE III. Relative performance of the four tapered transducers as observed from displacement amplitude measurements.

Transducer serial number	Ratio in dB of main peak amplitude to that of continuous distribution	Freq. spread of continuous distribution of peaks in kHz
1	17.3	70
2	16.8	110
3	16.2	220
4	16.0	310

strip in thickness vibration. It is noticed that although the two 6-dB points are equally separated in the profile from the central peak for transducer 1 [Fig. 4(a)], there is a difference of 2 mm for the corresponding values of transducer 2 [Fig. 4(b)] where the peak is shifted slightly toward the higher thickness side. The displacement amplitude characteristics, as studied by optical phase-locked laser interferometry, thus provide significant information regarding the true vibration patterns of the transducers. Use of a slightly tapered disk, such as that used in transducer 2, reveals an estimation of the obtainable precision of present experimental setup.

Figure 3(c), the displacement amplitude pattern of transducer 3, exhibits several prominent peaks (with the highest one of $\xi_m/V_D=1.4$ nm/V) of varying amplitude (15–40 nm) in the fundamental resonance region. This multiresonance pattern is again superimposed over a continuum of displacement amplitude in a wider frequency range (130–350 kHz). The transducer response does not show any prominent peak at higher harmonics. The displacement amplitude profile [Fig. 4(c)] in this case shows more nonuniformity in the distribution of normal displacement amplitude compared to the previous case.

Variation of ξ with frequency for transducer 4, formed with the highest tapered disk [Fig. 3(d)], presents several peaks in the fundamental resonance region, the pattern showing great resemblance to the corresponding electrical re-

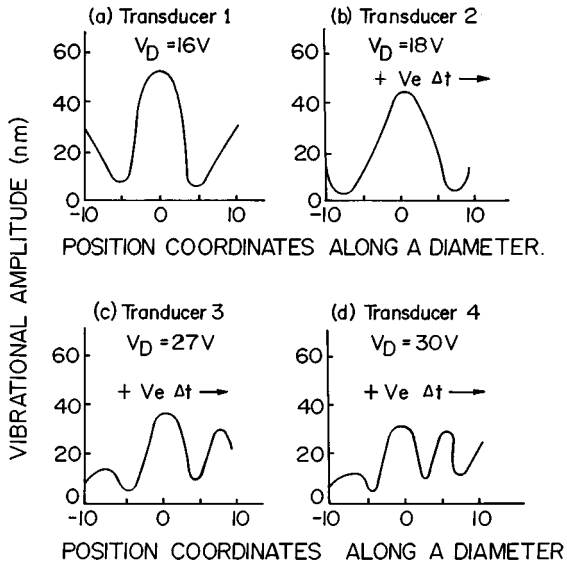


FIG. 4. Displacement amplitude profiles on resonating transducer surfaces.

sponse [Fig. 2(d)]. The highest one at 240.3 kHz appears with $\xi_m/V_D=1.15$ nm/V. An observable continuum of relatively high displacement amplitude appears over the frequency range 150–460 kHz, the spread being the highest of this kind studied here (Table III). The transducer, therefore, can be fairly operated through a wide frequency range for considerably high uniform output over which the discrete peaks are superimposed. The transducer, however, shows no response, at higher harmonics. The displacement amplitude profile [Fig. 4(d)] shows the maximum nonuniformity observed thus far in the present study. The gradual decrease in ξ_m/V_D with increasing tapering can be explained in the following way. As the number of resonance peaks or the continuum near resonance spreads with increasing tapering, the acoustic energy of the transducer at resonance gets distributed among all the peaks instead of being concentrated to a single resonance peak. This could lead to the lowering of the main peak amplitude.

The effect of tapering in disk thickness on the vibrational characteristics of ultrasonic transducers is observed in the appearance of multiresonance patterns in the fundamental resonance region and in the widening of frequency spread or the continuum of vibration amplitude around the resonance frequency. Additionally, the displacement amplitude profile on the resonating transducer surface becomes axially asymmetric, showing higher concentrations of vibration in the thicker part.

Vibrational amplitude measurements thus provide the following three parameters for the analysis of transducer performance:

(i) Frequency spread of resonance peaks or a continuum of ξ;
(ii) Asymmetry in profile with higher amplitude in thicker part;
(iii) Lowering of the higher harmonic peak amplitude of tapered transducer.

C. Study on the acoustic output of the transducers

Vibrational characteristics of a transducer are commonly studied through its temporal and frequency response when employed as a transmitter. Several experimental studies on acoustical measurements with tapered ultrasonic transducers and their theoretical interpretations are reported in the literature[1,7,8] where the impulse response (pulse waveform of acoustic output) and frequency response are mainly described. For example, in their study on tapered ceramics[8] by Barthe and Benkeser, the theoretically evaluated electric input impedance, spectral content, pulse, and continuous-wave pressure profile responses obtained by using a model were found to agree well with the measured response of a tapered piezoelectric transducer. In the present work it is proposed to examine whether there is any correspondance between the results on vibrational patterns and those obtained from conventional acoustical measurements with tapered ultrasonic transducers, to throw light on transducer design aspects from the knowledge of vibrational characteristics. To find that correspondence the effect of tapering in disk thickness has been studied by observing the impulse response, i.e., pulse waveforms of different transducers when driven with a spike voltage. Transducers with uniform thickness (transducer 1), slightly tapered (transducer 2), and highly tapered (transducer 4) disks have been used for this purpose. Photographs of the pulse waveforms obtained by using a probe hydrophone placed at a far-field axial point for each of the above three cases are shown in Fig. 5(a)–(c). It is very clear from the figure that a transmitter with parallel faced element gives rise to an output signal having the longest duration (the 6-dB diminution of peak amplitude is obtained at a duration of 24 μs). A slightly tapered transducer [Fig. 5(b)] produces a shorter time response (6-dB diminution at 10-μs duration), whereas the corresponding time is least for the highly tapered [Fig. 5(c)] transducer for which the 6-dB diminution is achieved at 4 μs. A comparison of Figs. 3, 4, and 5 thus reveals that the shorter the time response of acoustic output from the transducer, the wider the frequency spread of discrete peaks or the continuum near resonance and the more nonuniform the profile corresponding to a highly tapered disk.

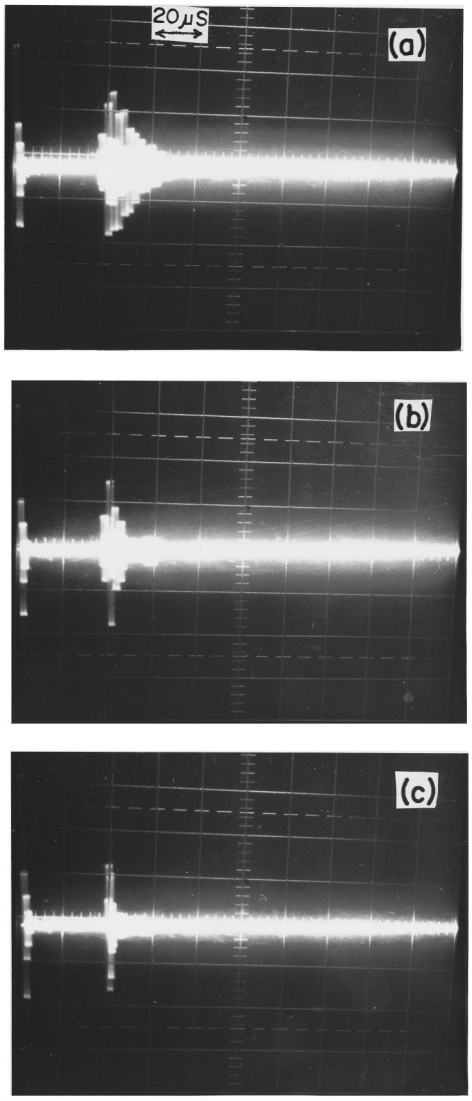


FIG. 5. Photographs showing the pulse waveforms from different transducers: (a) transducer 1 ($t=11.5$ mm); (b) transducer 2 ($t=11$–12 mm); (c) transducer 4 ($t=4$–11 mm).

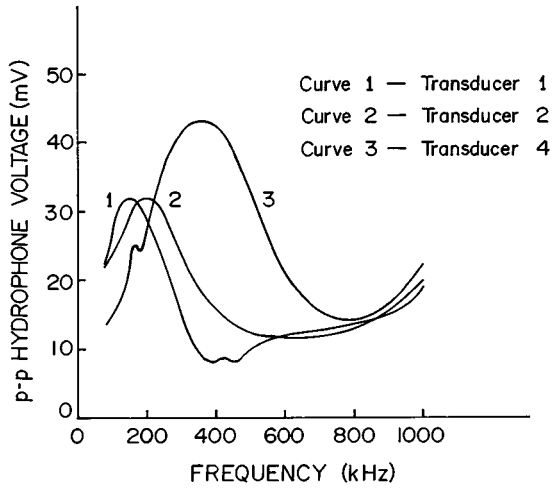


FIG. 6. Transmitting response of different transducers.

In the next step, tone-burst sinusoidal excitation with a single cycle was employed to drive the transducer operating as a transmitter, and the frequency response near resonance was measured for vibration in water. Studies were carried out at a far-field axial point of each transducer using the same probe hydrophone (Medisonics) and the results are depicted in Fig. 6. With increasing tapering the 3-dB bandwidth increases, resulting in a lowering of the Q factor of the transducer. A significantly high acoustic output is obtained with a highly tapered transducer. This could be because of the fact that tapering by gradually diminishing the thickness along a diameter results in a reduction of mass of the whole disk. Consequently, the relatively lighter transducer is capable of vibrating with a higher amplitude and thereby transmitting more acoustic energy into the surrounding medium, even when driven by the same excitation voltage as applied to the more uniformly thick-disk transducers. An increase in resonance frequency (f_r) is also observed in each case for the measurements in water. Although an increase in loading impedance is expected to decrease the value of f_r, an increase in this value caused by specific backing or loading, which is rather unexpected, has also been reported in the literature.[22]

III. CONCLUSION

The effect of tapering on the disk thickness of ultrasonic piezoelectric transducers on their characteristics is mainly manifested by a frequency spread of resonance peaks or a continuum in the vicinity of resonance, as observed from vibrational amplitude and conductance patterns. The transducer becomes less active at higher harmonics because of tapering. Tapering also causes the displacement amplitude profiles on resonating transducer surfaces to be axially asymmetric, where the vibration is more concentrated toward the thicker part. As it is observed that even a slight tapering affects the vibrational pattern up to an observable extent, the vibrational amplitude measurements would be very useful for detection of any tapering that may be introduced even by oversight.

ACKNOWLEDGMENTS

The authors are thankful to the Director, National Physical Laboratory, India, for giving permission to communicate the paper for publication. One of the authors (R.M.) is grateful to the Council of Scientific and Industrial Research, India for providing her financial assistance in the form of a research fellowship.

[1] C. B. Loutzenheiser and W. J. Denkmann, ''Thickness/twist vibrations of a truncated linearly tapered crystal strip,'' J. Acoust. Soc. Am. **41 (4, Pt.2)**, 962–968 (1967).
[2] Y. Jayet, M. Perdrix, and R. Goutte, ''Effects of damping of ultrasonic transducers due to lack of parallelism in the piezoelectric element,'' Ultrasonics **21**, 179–183 (1981).
[3] B. M. Alekseev, D. B. Dianov, and S. P. Karuzo, ''Tapered piezoelectric bar transducer with transverse polarization of the piezoceramic,'' Sov. Phys. Acoust. **23**, 1–4 (1977).
[4] R. L. Whitman, L. J. Laub, and W. J. Bates, ''Acoustic surface displacement measurements on a wedge-shaped transducer using an optical probe technique,'' IEEE Trans. Sonics Ultrason. **SU-15**, 186–189 (1968).
[5] Y. Tomikawa, H. Yamada, and M. Onoe, ''Wide band ultrasonic transducer using tapered piezoelectric ceramics for non-destructive inspection,'' Jpn. J. Appl. Phys. Suppl. 1 **23**, 113–115 (1964).
[6] T. Kobayashi, ''Wedged ultrasonic transducers having thickness extensional vibration,'' J. Acoust. Soc. Jpn. **38**, 748–754 (1982).
[7] P. G. Barthe and P. J. Benkeser, ''A staircase model of tapered thickness piezoelectric ceramics,'' J. Acoust. Soc. Am. **89**, 1434–1442 (1991).
[8] Y. Jayet, J. C. Baboux, F. Lakestani, and M. Perdrix, ''Theoretical and experimental investigation of a piezoelectric transducer with a nonparallel faced wearplate,'' IEEE Trans. Sonics Ultrason. **SU-32(b)**, 835–839 (1985).
[9] R. D. Mindlin, ''High efficiency vibrations of crystal plates,'' Q. Appl. Math. **19**, 51–61 (1961).
[10] D. I. Bolef and J. G. Miller, ''High-frequency continuous wave ultrasonics'' in *Physical Acoustics*, Vol. 8, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1971), pp. 95–201.
[11] P. J. Benkeser, L. A. Frizzell, K. B. Ocheltree, and C. A. Cain, ''A tapered phased array ultrasound transducer for hyperthermia treatment,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **UFFC-34(4)**, 446–453 (1987).
[12] D. R. Bacon, ''Primary calibration of ultrasonic hydrophones using optical interferometry,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **UFFC-35(2)**, 152–161 (1988).
[13] T. K. Saksena, J. N. Som, A. Kumar, and Y. Kumar, ''Study on low amplitude vibrations in air of an ultrasonic transducer with a phase locked laser interferometer,'' Ind. J. Pure Appl. Phys. **30**, 733–739 (1982).
[14] D. R. Bacon, ''The improvement and evaluation of a laser interferometer for the absolute measurement of ultrasonic displacements in the frequency range upto 15 MHz,'' NPL Report (U.K.), AC-109 (1986).
[15] L. E. Drain and B. C. Moss, ''An improvement and evaluation of a laser interferometer for the absolute measurement of displacements in ultrasonic vibrations up to 15 MHz,'' NPL Report (U.K.), AERE G-3525 (1988).
[16] J. H. Speake, ''An absolute method of calibrating ultrasonic transducers using laser interferometry,'' in Proc. on Evaln. & Calibn. of Ultrason. Transducers, London, 106–114 (1977).
[17] L. E. Drain and B. C. Moss, ''The frequency shifting of laser light by electro-optic technique,'' Opto-Electronics **4**, 429–439 (1972).
[18] C. F. Buhrer, L. R. Bloom, and D. H. Baird, ''Electro-optic light modulation with cubic crystals,'' Appl. Opt. **2(8)**, 839–846 (1963).
[19] E. A. G. Shaw, ''On the resonant vibrations of thick barium titanate disks,'' J. Acoust. Soc. Am. **28**, 38–50 (1956).
[20] G. Lypacewicz and L. Filipczinski, ''Measurement method and experimental study of ceramic transducer vibrations,'' Acustica **24**, 216–221 (1971).
[21] J. S. Arnold and J. G. Martner, ''Description of the resonances of short solid barium titanate cylinders,'' J. Acoust. Soc. Am. **31**, 217–226 (1959).
[22] D. F. Smith and C. V. Cagle, ''Ultrasonic testing of adhesive bonds using the Fokker bond tester,'' Mater. Eval. **24**, 362–370 (1966).

A new method for measuring *in situ* concrete elastic constants using horizontally polarized conical transducers

T.-T. Wu[a] and J.-S. Fang
Institute of Applied Mechanics, National Taiwan University, Taipei, Taiwan

(Received 15 December 1995; accepted for publication 16 August 1996)

This paper presents a result on the measurements of longitudinal and Rayleigh wave velocities in a concrete plate using horizontally polarized conical transducers. The characteristics of horizontally polarized surface responses due to a point source applied at a half-space and a plate are analyzed and discussed. The analyses show that the horizontal component of a skimming longitudinal wavefront due to a point source is much larger than its vertical component. This suggests the utilization of a horizontal component of a surface response to detect the skimming longitudinal wavefront. In addition, theoretical analysis shows that by using the horizontal component to determine the Rayleigh wave velocity, the source-to-receiver distance can be reduced significantly. Homemade horizontally polarized conical transducers were fabricated to demonstrate the feasibility of the present method on measuring the dynamic concrete elastic constants. Experimental results on the measurements of the horizontal component of a surface response due to a point impact were in good agreement with the theoretical prediction. In conclusion, it is noted that, based on this proposed method, concrete dynamic elastic constants can be determined through a single transient elastic wave experiment. [S0001-4966(97)04512-8]

PACS numbers: 43.40.At, 43.20.Bi, 43.58.Dj [CBB]

INTRODUCTION

The elastic constants of concrete are usually determined by the destructive method, which consists of using an uniaxial test to determine the Young's modulus of a standard cylindrical specimen. In the field of the nondestructive evaluation of concrete, the ultrasonic wave velocity is usually measured and utilized to predict the strength of concrete.[1,2] In addition to the ultrasonic method, various methods based on the point–source/point–receiver technique, which utilizes transient elastic wave response, have been proposed. Several successful applications[3–7] are presented and detailed.

In a recent paper, Wu *et al.*[8] proposed a method for determining the dynamic elastic constants of a concrete specimen using transient elastic waves. In the paper, the Rayleigh wave velocity was determined based on the cross correlation method and the longitudinal wave velocity, which was determined by measuring the longitudinal wavefront arrival. However, the major limitation of measuring the longitudinal wave velocity of a concrete specimen by the aforementioned method is that it requires the presence of a perpendicular corner. Thus, the longitudinal wave needs to travel around a corner, giving a bigger displacement jump at the point of the longitudinal wavefront arrival when a National Bureau of Standards (NBS) conical transducer[9] (vertically polarized and normal to the specimen surface) is used. Furthermore, in the Rayleigh wave velocity measurement, the source-to-receiver distance is relatively long. Therefore, these limitations hinder the broader application of this method.

In this paper, the limitations of the previous proposed method[8] on measuring the longitudinal wave and the Rayleigh wave velocities of concrete are removed. A new method based on the measurements of horizontally polarized surface responses, parallel to the surface, is proposed. In this new method, the longitudinal wavefront can be identified from the surface response directly. In addition, the source-to-receiver distance in determining the Rayleigh wave velocity is reduced. The advantages of using the horizontal components of surface responses to determine longitudinal and Rayleigh wave velocities will be presented in this paper. This new method allows concrete dynamic elastic constants to be determined through a single transient elastic wave experiment.

In the following sections, analysis and discussion of the horizontal components of surface responses of a homogeneous elastic half-space due to a steel ball impact are detailed. A new method based on the use of two homemade horizontally polarized conical transducers for detecting longitudinal and Rayleigh wave velocities is proposed. The proposed method is then applied to measure the elastic constants of a concrete plate. Satisfactory results with our new method were obtained.

I. ANALYSIS OF SURFACE RESPONSES OF A HALF-SPACE

Most current applications of point–source/point–receiver methods are based on analyzing vertically polarized displacement signals by use of a NBS conical transducer. However, as a horizontally polarized wave receiving device is available, some of the key characteristics of a horizontally polarized transient elastic wave signal can be adopted. In this section, the wave responses of a half-space surface induced by a point source with an unit-step time function, as well as by a low velocity point impact, will be given. The simulation of the Rayleigh wave velocity measurement using horizontal

[a]Corresponding author.

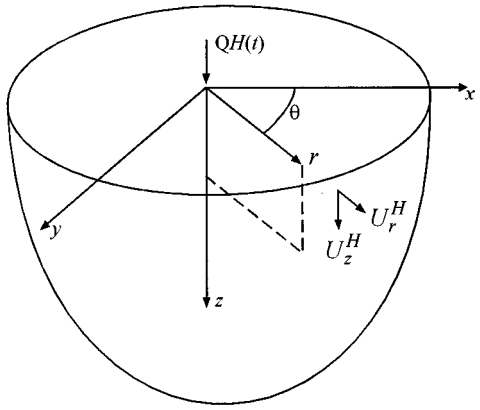


FIG. 1. Coordinates of a half-space with a point load applied at the origin.

components of surface responses is also demonstrated in this section. A computer program written by Hsu[10] is employed to calculate the transient elastic wave responses in a concrete plate.

A. Surface responses due to an impact source

In a homogeneous isotropic half-space with density ρ and Lamè constants λ, μ, the corresponding longitudinal wave speed is $C_L = \sqrt{(\lambda+2\mu)/\rho}$ and the transverse wave speed is $C_T = \sqrt{\mu/\rho}$. In the following, the vertical and radial components of surface responses of a half-space $z \geqslant 0$ (Fig. 1) are denoted as u_z and u_r, respectively. A normal point force of magnitude $QH(t)$ is applied at the origin [$H(t)$, is the heaviside time function]. The displacements are nondimensionalized according to

$$u_z^{H*} = \frac{\pi\mu r}{Q} u_z^H, \quad u_r^{H*} = \frac{\pi\mu r}{Q} u_r^H, \tag{1}$$

while the nondimensional time is

$$t^* = \frac{C_T t}{r}, \tag{2}$$

and r is the radial coordinate.

As pointed out by Chang,[11] for the case of a low-energy impact, the force-time relation appears to fall somewhere between the sine and sine-squared approximations. In this paper, a half $\sin^{3/2} t$ function is used to model the force-time function of a steel ball impacting elastically on an elastic material. The duration T_c (the contact time of the steel ball impact) of the impact is nondimensionalized according to $T_c^* = C_T T_c / r$. The contact time[12] of an elastic impact is a function of several factors: the radius, Poisson's ratio, Young's modulus of the steel ball, the drop height, and the elastic constants of the target specimen. For the case of a point source with an arbitrary force-time function acting on a half-space, the wave response can be obtained by convoluting the force-time function with the Green function of the half-space, or by superimposing the wave response due to the heaviside source. Herein, the superposition method similar to the method described in Ref. 8 is adopted.

Figure 2 shows the vertical (solid line) and horizontal (broken line) components of the surface wave signal generated by the impact of a steel ball at the origin of the half-space. The nondimensional duration T_c^* used in this simulation is 0.3. The arrival of the longitudinal wave (P), transverse wave (S), and Rayleigh wave (R) are marked in Fig. 2. We note that[8] the arrival of the Rayleigh wave does not coincide exactly with the first dip of the vertical component (solid line) of the wave response. The arrival of the longitudinal wavefront is relatively small when compared with the rest of the signal. In addition, the transverse wave arrival is continuous at the vertical displacement signal. In contrast to the vertical component, the horizontal component (broken line) of the longitudinal wave gives a much clear displacement jump at the wavefront arrival. Furthermore, the Rayleigh wave arrival induces a sharp corner and can be identified easily.

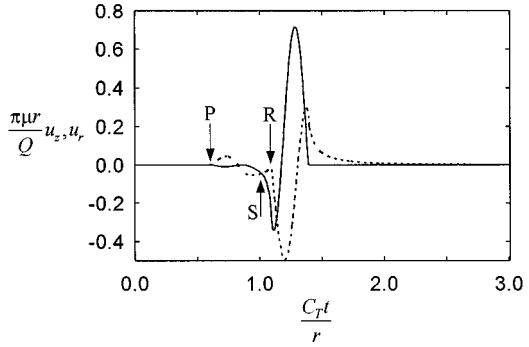


FIG. 2. The surface vertical displacement due to the impact of a steel ball at the origin of a half-space.

Figure 3 is the replot of Fig. 2 with an enlarged vertical scale to reveal the amplitude of the longitudinal wavefront. From this replot it is clear that the amplitude of the horizontal component is four to five times larger than that of the vertical component.

B. Simulation of Rayleigh wave velocity measurement

In Ref. 8 an analysis using two surface responses (vertical components) to determine Rayleigh wave velocity was given. In this section, a parallel analysis for the case of using horizontal components of surface responses to determine Rayleigh wave velocity is given.

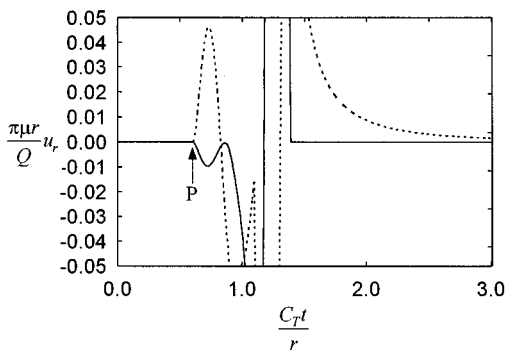


FIG. 3. The surface radial displacement due to the impact of a steel ball at the origin of a half-space.

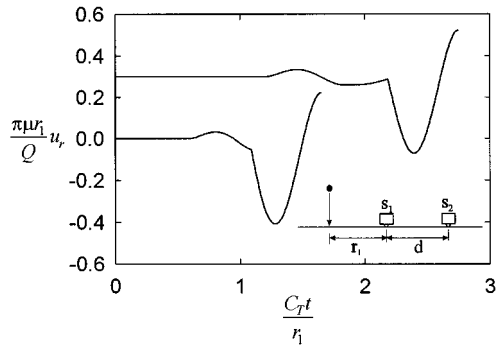


FIG. 4. The relative positions of an impact source and two receivers (S_1 and S_2) and the corresponding surface responses.

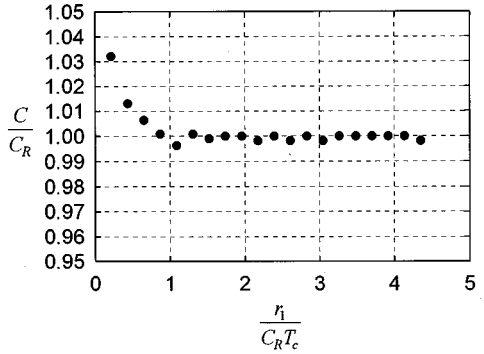


FIG. 5. The influence of r_1 on the calculated Rayleigh wave velocity using the cross correlation method (half-space).

The relative positions of an impact source and two receivers (S_1 and S_2) are shown in Fig. 4. The distance between the source and the first receiver S_1 is r_1, and the distance between the first and second receiver S_2 is d. In order to utilize the cross correlation method to determine the Rayleigh wave transit time between the first receiver and the second receiver, it is required that the wave responses received at these two receivers must be predominated by the Rayleigh waves. Therefore, the distance r_1 must be large enough to ensure that the Rayleigh wave carries most of the energy.

Letting $h_1(t)$ and $h_2(t)$ be the signals recorded at S_1 and S_2, respectively, the cross correlation function of these two signals is defined as

$$z(t)=\int_{-\infty}^{\infty} h_1(\tau)h_2(t+\tau)d\tau. \tag{3}$$

If the signals $h_1(t)$ and $h_2(t)$ are similar with the time delay Δt, the maximum of $z(t)$ occurs at $t=\Delta t$. In other words, the time delay between $h_1(t)$ and $h_2(t)$ can be determined by finding the corresponding time at which $z(t)$ is maximum.

The signals shown in Fig. 4 are the numerical simulations of the horizontal components of the surface responses recorded at receiver S_1 (lower trace) and S_2 (upper trace). The half $\sin^{3/2} t$ force-time function is utilized in the simulation. The nondimensional impact duration is $T_c^* = 0.6$, while $r_1=d$. The relative amplitude of the surface response at receiver S_2 in Fig. 4 has been multiplied by a factor of 1.5. On employing the cross correlation method for the two signals, the time difference of these two signals can be obtained and thus the Rayleigh wave velocity can be determined. The signal length used in the cross correlation calculation is up to the first peak after the Rayleigh wave arrival.

Figure 5 shows the results of r_1 which varies from $0.1d$ to $2d$ at an interval of $0.1d$. The horizontal axis is nondimensionalized with the exact Rayleigh wave velocity C_R and the impact duration T_c, i.e., $r_1^* = r_1/C_R T_c$ and the vertical axis is the ratio of the calculated and the exact Rayleigh wave velocity. The results in Fig. 5 show that the calculated Rayleigh wave velocities are almost equal (within an error of 1%) to the exact Rayleigh wave velocity as the nondimensional distance between the source and the first receiver is greater than 0.5. We note that the results shown in Fig. 5 is different from those of when vertical components of the surface responses[8] are used. In the latter case, the calculated Rayleigh wave velocity approaches the exact value gradually as the distance of the source to the first receiver increases. To within an 1% of error, r_1^* should be greater than 3.

Based on this study, we conclude that the use of horizontal components of surface responses to determine Rayleigh wave velocity can effectively reduce the distance between the source and the first receiver significantly. For example, the case where $T_c=20$ μs and Rayleigh wave velocity $C_R=2500$ m/s, the minimum distance between the source to the first receiver is only about 2.5 cm when the horizontal component measurement method is used. With the same testing geometry and elastic parameters, the minimum source-to-receiver distance is about 15 cm when the vertical component measurement of surface responses is utilized. We note that to measure the local elastic properties of a concrete structure, the source to the first receiver distance should be maintained as small as possible.

II. ANALYSIS OF SURFACE WAVE RESPONSE IN A PLATE

The difference between a concrete half-space and a concrete plate is the appearance of one extra rebounding surface in the latter case. Since a reflected longitudinal wave may arrive earlier than the Rayleigh surface wave (when the source-to-receiver distance is larger than the thickness of the plate), a variation on the calculated Rayleigh velocity is expected.

The configuration of the source and receiver arrangement in a plate is shown in Fig. 6. The thickness of the plate is denoted as H, while the solid curve represents the radial component of the surface wave signal generated by the impact of a steel ball at the origin of the plate. The nondimensional duration T_c^* used in this simulation is 0.5, i.e., $T_c^* = C_T T_c/H = 0.5$. The first receiver is located on the same surface with a distance H to the source. The broken curve in Fig. 6 represents the radial component of the surface response at a half-space with the same source and receiver arrangement. In the case of the plate response (solid line), the arrival of the longitudinal wave (P), Rayleigh wave (R), and the reflected longitudinal wave (2P) are marked. On the ar-

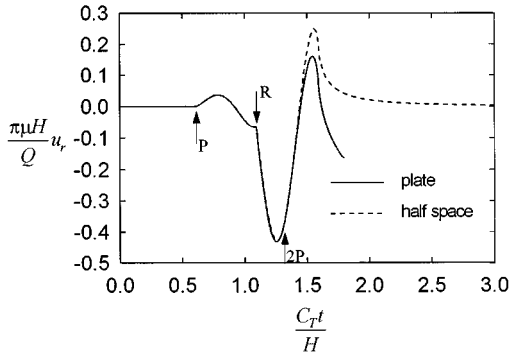


FIG. 6. Comparison of surface radial displacements of a plate (solid curve) and a half-space (broken curve) due to the impact of a steel ball with the source-to-receiver distance H.

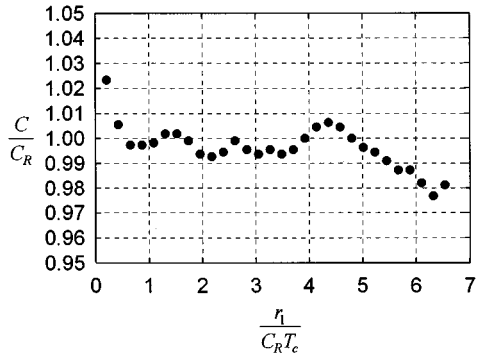


FIG. 8. The influence of r_1 on the calculated Rayleigh wave velocity using the cross correlation method (plate).

rival of the reflected longitudinal wave from the bottom surface of the plate, the wave signal deviates from the half-space response. Figure 7 shows similar results for the case when the source-to-receiver distance is changed to $2H$. In this case, the reflected longitudinal wave (2P) arrives earlier than the Rayleigh wave (R), and thus a slight signal deviation between these two signals occurs.

Figure 8 shows the numerical results of the calculated Rayleigh wave velocities using the horizontal component of the plate surface responses. In our calculations, r_1 varies from $0.1H$ to $2H$ at an interval of $0.1H$. From the results, we find that the calculated Rayleigh wave velocities are scattered around the true value of the Rayleigh wave velocity (within 1% of error) for $0.5<r_1/C_R T_c<5.5$. In this region, the influence of the reflected longitudinal wave on the determination of the Rayleigh wave velocity using the cross correlation method is surprisingly small and may be explained by the following. In the determination of the Rayleigh wave velocity using the cross correlation method, the signal length utilized is only up to the first peak after the arrival of the Rayleigh wave. In this case, the signal difference between a half-space and a plate response (horizontal component) is small. This is due to the fact that most of the shapes of the responses are dominated by the big downward dip caused by the Rayleigh wave.

From the above discussions, we conclude that the present method of determining Rayleigh wave velocity can be applied not only in a half-space but also in a plate specimen with reasonable accuracy.

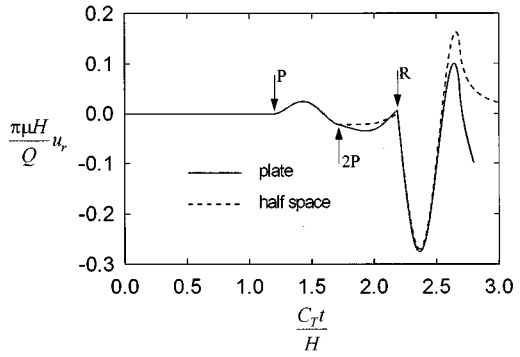


FIG. 7. Comparison of the surface radial displacements of a plate (solid curve) and a half-space (broken curve) due to the impact of a steel ball impact with the source-to-receiver distance $2H$.

III. HORIZONTALLY POLARIZED CONICAL TRANSDUCER

The NBS conical transducer which was designed for the purpose of detecting transient elastic waves or acoustic emission wave signals has found many successful applications. It is a highly sensitive, broadband (10 kHz–2 MHz) device which measures vertical components of wave responses on the surface of a solid. In a previous paper, Proctor[13] developed a companion transducer which measures the tangential (horizontally polarized) component of the same configuration.

In our study, two homemade horizontally polarized conical transducers were fabricated. The design of the horizontally polarized conical transducer is similar to the NBS conical transducer, except for the substitution of the longitudinal piezoelectric conical element in the NBS conical transducer with a shear conical element. The active element is a cone-shaped piece of lead-zirconate-titanate (PZT-5A, shear). The tip of the conical element is about 1.5 mm in diameter. The backing brass cylinder is large in size when compared to the element both in length and radial dimensions. The conical element was mounted on a brass cylinder with conductive silver epoxy. The tip of the conical transducer is coated with gold and is used as the ground port of the sensing circuit. In the measurement of a nonconducting material such as concrete, aluminum foil is placed between the conical transducer and the testing sample. Since the conical element is a shear type piezoelectric element, the polarized direction of the conical transducer is marked on the brass backing for reference.

To test the performance of the homemade horizontally polarized conical transducer, a transient elastic wave experiment was conducted on a steel block (20 cm in diameter and 15 cm in length). The steel block is large enough to be considered a homogeneous half-space. The longitudinal and transverse wave velocities of the steel block are $C_L=5922$ m/s and $C_T=3234$ m/s, respectively. The solid line in Fig. 9 denotes the measured radial component of the surface response due to a 2-mm steel ball impact on the steel half-space, while the broken line is the simulated surface response. The time origin of the impact was not determined

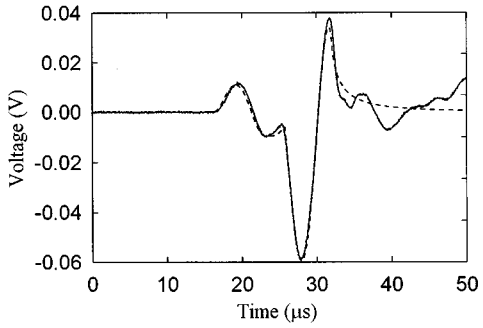


FIG. 9. The measured (solid line) and calculated (broken line) radial components of the surface responses due to the impact of a 2-mm steel ball at the steel half-space.

and hence the time origin in Fig. 9 is a relative value. The source-to-receiver distance of the experiment is 7 cm. In addition, the first peak of both the simulated and the measured signals were adjusted to have the same magnitude. The result shows that the homemade horizontally polarized conical transducer gives a faithful measurement of the radial component of the surface response for waves arriving before the Rayleigh wave packet. To demonstrate the advantages of using a horizontally polarized conical transducer in detecting longitudinal wavefront arrival, the vertical components of surface responses generated by a steel ball impact were recorded simultaneously with a commercial NBS conical transducer. The NBS conical transducer was placed at the opposite end of the homemade horizontally polarized conical transducer with the same source-to-receiver distance, i.e., 7 cm. The solid line in Fig. 10 is the recorded vertical component of the surface response, while the broken line represents the radial component.

From the results, it is obvious that measuring radial components of surface responses is a much better choice for detecting longitudinal wave arrival.

IV. MEASUREMENTS OF LONGITUDINAL AND RAYLEIGH WAVE VELOCITIES

Instead of finding the time origin of an impact source, we utilized two transducers to measure the time delay between the wavefront arrivals received at the two different positions. The source and the receivers were located on the same straight line. Since the distance between the receivers is known, the wave velocity can be obtained without knowing the time origin of the impact source. In this study, two homemade horizontally polarized conical transducers are utilized to measure the radial component of the surface displacement signals. The received voltage signals from the conical transducers were amplified by a preamplifier and recorded by a 100-MHz digital oscilloscope (LeCroy 9314L).

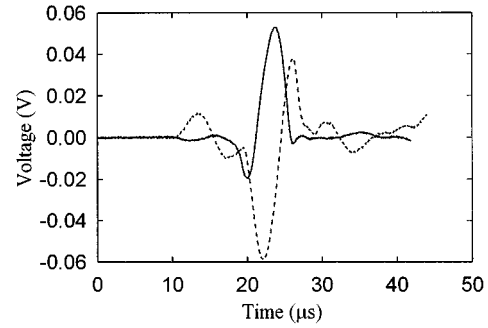


FIG. 10. The measured vertical component (solid line) and radial component (broken line) of surface response at a half-space.

Before applying this proposed method to a concrete specimen, the method was first applied to an aluminum specimen (Al 6061-T651) with dimensions 20 cm×20 cm×15 cm. A steel ball with a 1-mm diameter was the impact source. The source to the first receiver distance r_1 ranged from 4 to 10 cm and the distance between the two receivers was 9 cm. The measured longitudinal and Rayleigh wave velocities were $C_L=6430\pm69$ m/s and $C_R=2912\pm5$ m/s, respectively. For comparison, the longitudinal, transverse, and Rayleigh wave velocities of the aluminum specimen were determined using an ultrasonic method and the results obtained are $C_L=6398$ m/s, $C_T=3128$ m/s, and $C_R=2921$ m/s, respectively. Our results show that the current proposed method gives an accurate and reliable measurement of longitudinal and Rayleigh wave velocities.

In the following section, results on applying the proposed method to determine the longitudinal and Rayleigh wave velocities of a concrete specimen are given. The dimensions of the concrete specimen used in this study were 60 cm wide, 55 cm long, and 13 cm thick. The density of the concrete specimen was $\rho=2427$ kg/m^3. A steel ball with a 4.75-mm diameter was the impact source.

A. Determination of the longitudinal wavefront

In our longitudinal wavefront measurement the vertical sensitivity of the oscilloscope was set at 10 mv/div to increase the resolution of the measured arrival time. Experimental results on the concrete specimen show that the arrival of the skimming longitudinal wave gives a big displacement jump at the horizontal component of the surface response. To enhance the resolution of the delay time measurement, the recorded signal was enlarged and processed by a linear phase finite impulse response filter which was embedded in the digital oscilloscope.

To measure the longitudinal wave velocity of the concrete specimen, the impact source to the first receiver distance was kept at the value of $r=15$ cm. The distance between the first receiver and the second receiver was then varied from 5 cm to 17 cm at an interval of 1 cm. The source and the receivers were located along the same line.

Figure 11 shows the results for the longitudinal wavefront measurements conducted at the bottom of the concrete specimen. Results show a linear relationship between the arrival time delay and the separation of the two receivers. For the top of the concrete specimen, the results were similar. The least-square fittings of the measured longitudinal wave velocities on both the bottom and the top surfaces of the concrete plate are

Bottom surface: $C_L=4602\pm76$ m/s;

Top surface: $C_L=4240\pm101$ m/s.

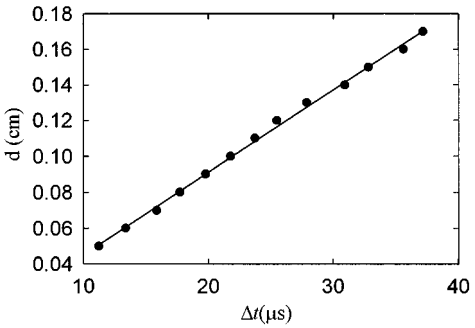


FIG. 11. The results for the longitudinal wavefront measurements conducted at the bottom of the concrete specimen.

The difference of the longitudinal wave velocities between the top and the bottom surfaces is solely due to the gravity effect in the construction of a concrete plate, i.e., the coarse aggregate tends to concentrate more at the bottom surface of the plate.

B. Determination of the Rayleigh wave velocity

As discussed in the previous section, the influence of the reflected longitudinal wave on the Rayleigh wave velocity determination using the cross correlation method is small. In order to verify that this effect is negligible in concrete measurement, transient elastic wave experiments with r_1 ranging from 5 to 22 cm were conducted. The distance between the first and the second receivers was 13 cm, which was the same as the thickness of the concrete plate. The same experiments were performed on both the bottom and the top surfaces of the concrete plate. The measured results are shown as solid circles (bottom surface) and triangles (top surface) in Fig. 12. The variations of the measured Rayleigh wave velocities with respect to the distance r_1 showed a similar trend as that predicted from numerical simulations (Fig. 8). The average values of the Rayleigh wave velocities for the bottom and top surfaces of the concrete plate are

Bottom surface: $C_R = 2548 \pm 16$ m/s;

Top surface: $C_R = 2302 \pm 24$ m/s.

The determined Rayleigh wave velocities of the bottom and top surfaces show a consistent trend to that of longitudinal wave velocities, i.e., the velocity of the bottom surface is higher than that at the top surface.

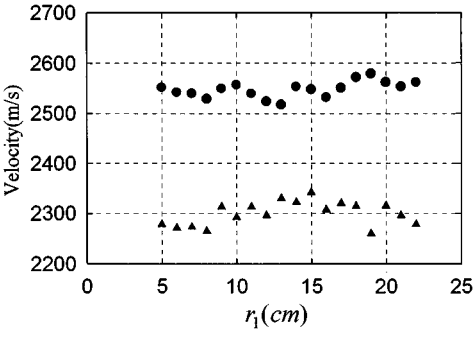


FIG. 12. The Rayleigh wave velocities measured at the bottom surface (solid circles) and top surface (triangles) of the concrete specimen.

TABLE I.

	λ (Gpa)	μ (Gpa)	E (Gpa)	ν
Bottom surface	13.56	18.92	45.74	0.209
Top surface	12.99	15.32	37.67	0.229

C. Discussions

The results of Secs. IV A and IV B show that the measurement error of the longitudinal wave velocity is larger than that of the Rayleigh wave velocity. This difference may be due to the different methods adopted in the determination of the transit time of the longitudinal wave and Rayleigh wave. In the Rayleigh wave velocity measurement, a segment (to the first peak of the Rayleigh wave) of the signal was utilized. However, in the longitudinal wave velocity measurement, only the signal around the longitudinal wavefront arrival was utilized. In the latter case, any small noise around the longitudinal wavefront will further induce error in the arrival time determination. The distance between the two receivers also plays an important part in the measurement error of both the longitudinal wave and Rayleigh wave velocities. In practice, the receivers can be fixed with a designed receiver–holder to hold the accuracy of the distance.

V. DETERMINATION OF DYNAMIC ELASTIC CONSTANTS OF CONCRETE

With the Rayleigh and the longitudinal wave velocities known, the transverse wave velocity of the concrete specimen can be obtained by solving the well-known Rayleigh wave equation

$$\left(2-\frac{C_R^2}{C_T^2}\right)^2 = 4\left(1-\frac{C_R^2}{C_L^2}\right)^{1/2}\left(1-\frac{C_R^2}{C_T^2}\right)^{1/2}. \tag{4}$$

The results obtained are as follows:

Bottom surface: $C_T = 2792$ m/s;

Top surface: $C_T = 2513$ m/s.

Once the longitudinal and the transverse wave velocities of an isotropic material are known, the Lamè constants λ, μ, or the Young's modulus and Poisson's ratio can be obtained in a straightforward way as those shown in Table I.

VI. CONCLUDING REMARKS

This paper presents a new method for measuring the longitudinal and the Rayleigh wave velocities based on the measurements of horizontally polarized surface responses. The characteristics of horizontally polarized surface responses due to a point source applied either on a half-space or a plate are analyzed and discussed. Based on the analysis, the utilization of the horizontal component to detect the skimming longitudinal wavefront is strongly recommended. Further, theoretical analysis shows that the use of horizontal components of surface responses to determine Rayleigh wave velocity can reduce the source-to-receiver distance sig-

nificantly. Experimental verification of the proposed method applied to a concrete plate was conducted and the results show that the method is indeed promising. In conclusion, we note that based on this proposed method, concrete dynamic elastic constants can be determined through a single transient elastic wave experiment.

ACKNOWLEDGMENTS

The authors thank the financial supports from the National Science Council (NSC84-2621-P-002-012B) of the Republic of China. The authors also thank Dr. Nelson Hsu of NIST for offering us the Plate Program.

[1] ''Standard Test Method for Pulse Velocity Through Concrete,'' ASTMC 597-71 (1979).
[2] A. E. Ben-Zeitun, ''Use of pulse velocity to predict compressive strength of concrete,'' Int. J. Cement Composites Light Weight Concrete **8**(1), 51–59 (1986).
[3] N. J. Carino, M. Sansalone, and N. N. Hsu, ''A point source–point receiver pulse-echo technique for flaw detection in concrete,'' ACI J. Proc. **83**, 199–208 (1986).
[4] N. J. Carino and M. Sansalone, ''Detection of voids in grouted ducts using the impact-echo method,'' ACI Mater. J. **89** (3), 296–303 (1992).
[5] K. Y. Kim, W. Sachse, and A. G. Every, ''On the determination of sound speeds in cubic crystals and isotropic media using a broad band ultrasonic point-source/point-receiver method,'' J. Acoust. Soc. Am. **93,** 1393–1406 (1993).
[6] J.-F. Chai, and T.-T. Wu, ''Determinations of anisotropic elastic constants using laser generated surface waves,'' J. Acoust. Soc. Am. **95,** 3232–3241 (1994).
[7] T.-T. Wu, J.-S. Fang, and P.-L. Liu, ''Detection of the depth of a surface-breaking crack using transient elastic waves,'' J. Acoust. Soc. Am. **97,** 1678–1686 (1995).
[8] T.-T. Wu, J.-S. Fang, G.-Y. Liu, and M.-K. Kuo ''Determination of elastic constants of a concrete specimen using transient elastic waves,'' J. Acoust. Soc. Am. **98,** 2142–2148 (1995).
[9] T. M. Proctor, ''An improved piezoelectric acoustic emission transducer,'' J. Acoust. Soc. Am. **71,** 1163–1168 (1982).
[10] N. N. Hsu, ''Dynamic Green's functions of an infinite plate: A computer program,'' National Bureau of Standards, NBSIR, 85–3234 (1985).
[11] C. Chang, ''Impact force measurement by signal deconvolution,'' MS thesis, Cornell University, 1983.
[12] W. Goldsmith, *Impact: The Theory and Physical Behavior of Colliding Solids* (Edward Arnold, London, 1965).
[13] T. M. Proctor, ''A high fidelity piezoelectric tangential displacement transducer for acoustic emission,'' J. Acoust. Emiss. **7** (1), 41–47 (1988).

The response of an elastically supported infinite Timoshenko beam to a moving vibrating mass

Seroj Mackertich
Pennsylvania State University—Harrisburg, Middletown, Pennsylvania 17057-4898

(Received 10 May 1996; accepted for publication 16 September 1996)

An analysis is made of the problem of vibrations of an infinite beam on elastic foundation excited by a moving and vibrating mass. The solution is presented within the framework of a beam theory which includes the effects of shear deformation and rotatory inertia. An example is provided where the displacement is calculated. © *1997 Acoustical Society of America.* [S0001-4966(97)05001-7]

PACS numbers: 43.40.Cw [CBB]

INTRODUCTION

The beam on elastic foundation is a common technical problem in engineering and many solutions have been proposed in recent years.[1–6]

As a problem of this kind, Duffy[6] treated the dynamic response of an infinite railroad track to a moving vibrating mass. He studied the effect of the moving and vibrating load as a function of both the mass and driving frequency of the load as well as the physical properties of the track, and obtained the deflection curves for several values of load velocity. The problem was studied on the basis of the classical Bernoulli–Euler beam theory. Chonan[7] studied the problem of moving harmonic load on an elastically supported Timoshenko beam. He compared his results with those by the Euler beam theory obtained in Ref. 8. In this paper, we find the response of an elastically supported infinite Timoshenko beam to a moving vibrating mass.

I. FORMULATION OF THE PROBLEM

Consider an infinite Timoshenko beam on an elastic foundation, initially undeformed and at rest that is excited by a constant velocity moving, vibrating mass as shown in Fig. 1. The equation of motion governing transverse vibration of the beam including shear deformation and rotatory inertia on elastic foundation can be written as[9]

$$EI\,\frac{\partial^2\theta}{\partial x^2}+k'AG\left(\frac{\partial y}{\partial x}-\theta\right)=\zeta I\,\frac{\partial^2\theta}{\partial t^2}, \tag{1}$$

$$k'AG\left(\frac{\partial^2 y}{\partial x^2}-\frac{\partial\theta}{\partial x}\right)+P(x,t)-ky=\zeta A\,\frac{\partial^2 y}{\partial t^2}, \tag{2}$$

where y is transverse displacement of the beam, P is the time-and-space-dependent applied force per unit length, E is Young's modulus, I is the moment of inertia of the cross-sectional area, ζ is density, k' is the shear coefficient, A is the cross-sectional area, G is the shear modulus, and k is subgrade reaction modulus. The general force $P(x,t)$ may be written as

$$P(x,t)=\left[F\,\cos\,\omega t-M\left(\frac{\partial^2 y}{\partial t^2}+2\nu\,\frac{\partial^2 y}{\partial x\,\partial t}+\nu^2\frac{\partial^2 y}{\partial x^2}\right)\right]\delta(x-\nu t),$$

where ν is the velocity of moving mass, M is the moving mass, F is the amplitude of the force whose frequency of vibration is $\omega/2\pi$, $\delta(x)$ is the Dirac-delta function, and parentheses inside the brackets represent the transverse acceleration of the mass.

Equations (1) and (2) can be made dimensionless by measuring lengths in terms of the radius of gyration of the cross section and measuring time in terms of the period of vibration of the undeformed beam on the elastic foundation.

We introduce

$$\omega_0=\sqrt{\frac{k}{\zeta A}}\quad r_0=\sqrt{\frac{I}{A}}\quad \nu_0=r_0\omega_0,$$

$$\alpha=\frac{\nu}{\nu_0}\quad \Delta=\frac{M}{\zeta A r_0}\quad W=\frac{y}{r_0},$$

$$a_2=\frac{\sqrt{E/\zeta}}{\nu_0}\quad a_1=\frac{\sqrt{K'G/\zeta}}{\nu_0},$$

$$z=x-\nu t\quad \tau=\omega_0 t\quad R=\frac{z}{r_0},$$

$$F_0=\frac{F}{kr_0}\quad \gamma=\frac{\omega}{\omega_0}. \tag{3}$$

Substitution of (3) into Eqs. (1) and (2) gives the dimensionless equations of motion:

$$(a_2^2-\alpha^2)\,\frac{\partial^2\theta}{\partial R^2}+a_1^2\,\frac{\partial W}{\partial R}-a_1^2\theta+2\alpha\,\frac{\partial^2\theta}{\partial R\,\partial\tau}=\frac{\partial^2\theta}{\partial\tau^2}, \tag{4}$$

$$(a_1^2-\alpha^2)\,\frac{\partial^2 W}{\partial R^2}-[\Delta\,\delta(R)+1]\,\frac{\partial^2 W}{\partial\tau^2}-a_1^2\,\frac{\partial\theta}{\partial R}$$

$$-W+2\alpha\,\frac{\partial^2 W}{\partial\tau\,\partial r}=-F_0\,\cos\,\gamma\tau\delta(R). \tag{5}$$

To complete the formulation, the initial conditions and boundary conditions are chosen as

$$W(R,0)=\frac{\partial W}{\partial\tau}\,(R,0)=0,$$

and we require that the solution must vanish as $R\to\pm\infty$. We solve Eq. (4) and Eq. (5) using double Fourier transforms. If we assume that $W(R,\tau)$ possesses a Fourier transform $W(\Gamma,\tau)$ defined by

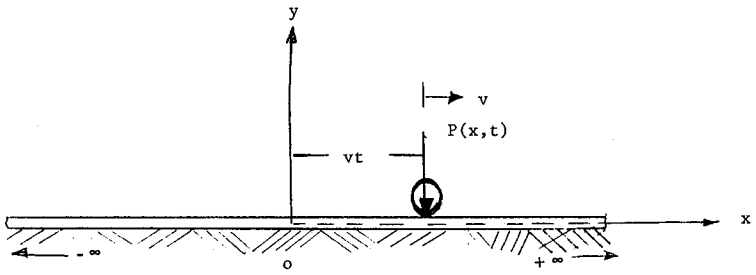


FIG. 1. Elastically supported infinite beam.

$$\overline{W}(\Gamma,\tau)=\int_{-\infty}^{\infty} W(R,\tau)e^{-i\Gamma R}\,dR, \tag{6}$$

and that $\overline{W}(\Gamma,\tau)$ possesses a Fourier transform $\overline{W^*}(\Gamma,\Omega)$ defined by

$$\overline{W^*}(\Gamma,\Omega)=\int_{-\infty}^{\infty} \overline{W}(\Gamma,\tau)e^{-i\Omega\tau}\,d\tau. \tag{7}$$

Equations (4) and (5) can be transformed into

$$-(a_2^2-\alpha^2)\Gamma^2\overline{\theta^*}+i\Gamma a_1^2\overline{W^*}-a_1^2\overline{\theta^*}-2\Gamma\alpha\Omega\,\overline{\theta^*}=-\Omega^2\overline{\theta^*}, \tag{8}$$

$$-(a_1^2-\alpha^2)\Gamma^2\overline{W^*}+\Delta\Omega^2W^*(o,\Omega)+\Omega^2\overline{W^*}-i\Gamma a_1^2\overline{\theta^*}-\overline{W^*}-2\Gamma\alpha\Omega\,\overline{W^*}=-\pi F_0[\,\delta(\Omega-\gamma)+\delta(\Omega+\gamma)], \tag{9}$$

and eliminating $\overline{\theta^*}$ from Eq. (8) and Eq. (9) one obtains

$$\overline{W^*}(\Gamma,\Omega)=\{\pi F_0A_1[\,\delta(\Omega-\gamma)+\delta(\Omega+\gamma)]+\Delta A_1\Omega^2W^*(0,\Omega)\}/A_2,$$

where

$$A_1=(a_2^2-\alpha^2)\Gamma^2+a_1^2+2\Gamma\alpha\Omega-\Omega^2 \tag{10}$$

and

$$A_2=[(a_1^2-\alpha^2)A_1-a_1^4]\Gamma^2+A_1(1-\Omega^2)+2\Gamma\alpha\Omega A_1.$$

Applying inverse Fourier transform to Eq. (10), one obtains

$$W^*(R,\Omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\{\pi F_0A_1[\,\delta(\Omega-\gamma)+\delta(\Omega+\gamma)]+\Delta A_1\Omega^2W^*(o,\Omega)\}e^{i\Gamma R}\,\frac{d\Gamma}{A_2}. \tag{11}$$

To evaluate the integral in Eq. (11), one can extend the variable Γ into the complex plane and use the method of Cauchy contour integration for their evaluation. The integrand of Eq. (11) has four simple poles; two of them lie in the upper-half-plane (Γ_1 and Γ_2) and two of them in the lower half-plane (Γ_3 and Γ_4), the exact values of the poles must be found numerically. When the integrand of Eq. (11) has poles of the second order, difficulties arise in the evaluation of the integral it does not exist and we encounter a resonance effect in the sense that the amplitude of deflection increase beyond bound. Evaluating the integral in Eq. (11) using Cauchy contour integration method, one obtains

$$\overline{W}(R,\Omega)=i\ \mathrm{sgn}(R)\{\pi F_0[\,\delta(\Omega-\gamma)+\delta(\Omega+\gamma)]\times[A_{1n}A_{3j}e^{i\Gamma_nR}+A_{ij}A_{3n}e^{i\Gamma_jR}]\}/[A_{3n}A_{3j}-i\ \mathrm{sgn}(R)\Delta\Omega^2(A_{1n}A_{3j}+A_{1j}A_{3n})], \tag{12}$$

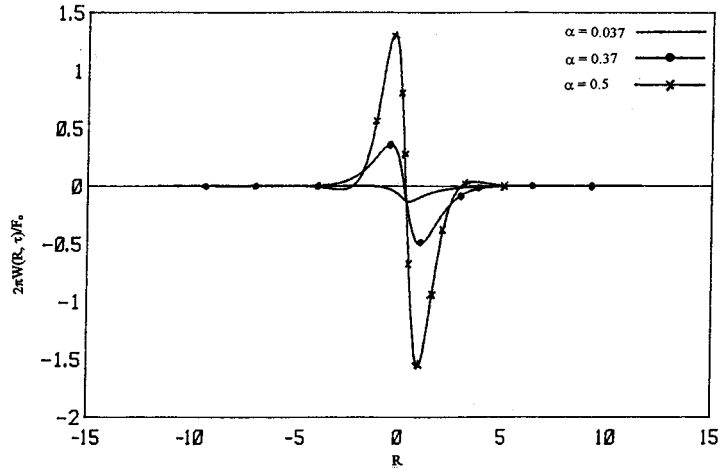


FIG. 2. Deflection of the beam with time as a function of distance from the location of the moving vibrating load. Parameters are $\Omega=0.5$, $\Delta=1$, $\tau=3$.

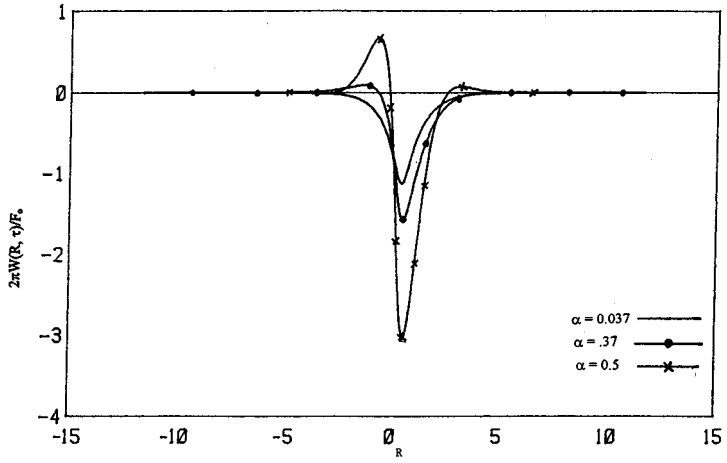


FIG. 3. Deflection of the beam with time as a function of distance from the location of the moving vibrating load. Parameters are $\Omega=0.5$, $\Delta=1$, $\tau=27$.

where

$$A_3=\frac{dA_2}{d\Gamma},$$

$$\Gamma_n\&\Gamma_j=\Gamma_1\&\Gamma_2 \quad \text{for } R\geqslant 0,$$

$$\Gamma_n\&\Gamma_j=\Gamma_3\&\Gamma_4 \quad \text{for } R<0,$$

$$A_{1n}=A_1(\Gamma=\Gamma_n), \quad A_{1j}=A_1(\Gamma=\Gamma_j),$$

$$A_{3n}=A_3(\Gamma=\Gamma_n), \quad A_{3j}=A_3(\Gamma=\Gamma_j)$$

$$\operatorname{sgn}(R)=1 \quad \text{when } R\geqslant 0,$$

and

$$\operatorname{sgn}(R)=-1 \quad \text{when } R<0.$$

Applying inverse Fourier transform on time to Eq. (12), one obtains

$$W(R,\tau)=\frac{1}{2\pi}\int_{-\infty}^{\infty}[i\ \operatorname{sgn}(R)\pi F_0\delta(\Omega-\gamma)][A_{1n}A_{3j}e^{i\Gamma_n R}+A_{1j}A_{3n}e^{i\Gamma_j R}](1/\beta)e^{i\Omega\tau}d\Omega$$
$$+\frac{1}{2\pi}\int_{-\infty}^{\infty}[i\ \operatorname{sgn}(R)\pi F_0\delta(\Omega+\gamma)]\times[A_{1n}A_{3j}e^{i\Gamma_n R}+A_{1j}A_{3n}e^{i\Gamma_j R}](1/\beta)e^{i\Omega\tau}\,d\Omega. \tag{13}$$

Evaluating the integral in Eq. (13), one obtains

$$W(R,\tau)=(i\ \operatorname{sgn}(R)F_0/2\pi)\{[\widetilde{A}_{1n}\widetilde{A}_{3j}e^{i(\Gamma_j R+\gamma\tau)}+\widetilde{A}_{1j}\widetilde{A}_{3n}e^{i(\widetilde{\Gamma}_j R+\gamma\tau)}](1/\widetilde{\beta})\overline{A_{1n}}\,\overline{A_{3j}}e^{i(\overline{\Gamma}_n R-\gamma\tau)}+\overline{A_{1j}}\,\overline{A_{3n}}e^{i(\overline{\Gamma}_j R-\gamma\tau)}](1/\overline{\beta})\}, \tag{14}$$

where

$$\widetilde{A}_{1n}=A_{1n}(\Omega=\gamma), \quad \overline{A_{1n}}=A_{1n}(\Omega=-\gamma),$$

$$\widetilde{A}_{1j}=A_{1j}(\Omega=\gamma), \quad \overline{A_{1j}}=A_{1j}(\Omega=-\gamma),$$

$$\widetilde{A}_{3n}=A_{3n}(\Omega=\gamma), \quad \overline{A_{3n}}=A_{3n}(\Omega=-\gamma),$$

$$\widetilde{\beta}=\beta(\Omega=\gamma), \quad \overline{\beta}=\beta(\Omega=-\gamma),$$

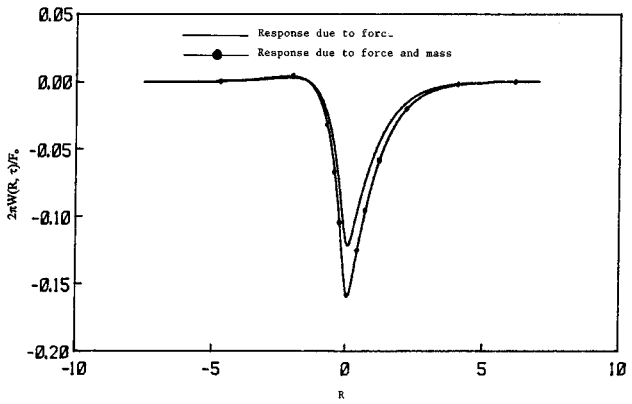


FIG. 4. Deflection of the beam with time as a function of distance from the location of the moving vibrating load. Parameters are $\Omega=0.5$, $\tau=3$, $\alpha=0.037$.

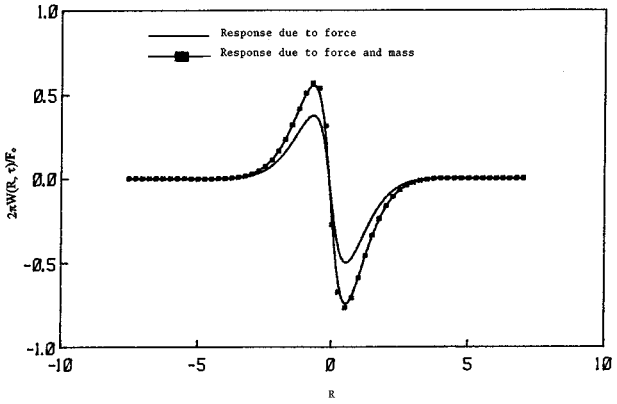


FIG. 5. Deflection of the beam with time as a function of distance from the location of the moving vibrating load. Parameters are $\Omega=0.5$, $\tau=3$, $\alpha=0.37$.

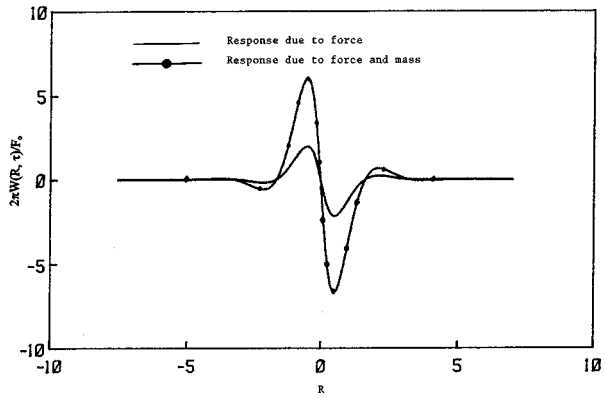


FIG. 6. Deflection of the beam with time as a function of distance from the location of the moving vibrating loads. Parameters are $\Omega=0.5$, $\tau=3$, $\alpha=0.5$.

and

$$\beta=A_{3n}A_{3j}-i\ \mathrm{sgn}(R)\Delta\Omega^2(A_{1n}A_{3j}+A_{1j}A_{3n}).$$

Equation (14) is the deflection of infinite beam on elastic foundation excited by moving and vibrating mass.

II. RESULTS AND CONCLUSIONS

In this study using double Fourier transform and Cauchy contour integration we obtained solution for infinite long beam on elastic foundation excited by moving, vibrating force and mass equation (14).

Numerical results are computed for material properties listed below:

$E=206.92$ GPa, $G=82.37$ GPa,
$k'=1.2$, $k=689.41$ MPa,
$I=40.28\times10^{-6}$ m^4, $A=1.55\times10^{2}$ m^2,
$\zeta=7860$ kg/m^3, $\Delta=1$.

Figures 2 and 3 show the dimensionless deflection of beam for various mass velocities.

Figures 4, 5, and 6 represent response of the Timoshenko beam to moving and vibrating force (Chonan), and moving and vibrating force and mass.

Comparing response of infinite long beam excited by moving and vibrating force, and moving and vibrating force and mass Figs. 4, 5, and 6 the mass contribution to the deflection of the beam is significant especially when the velocity of the load increases.

[1] M. Hetenyi, *Beams on Elastic Foundation* (Univ. of Michigan, Ann Arbor, MI, 1946).
[2] S. H. Crandall, ''The Timoshenko Beam on an Elastic Foundation,'' Proc. 3rd Midwestern Conference Solid Mech., pp. 146–159 (1957).
[3] T. M. Wang and J. E. Stephens, ''Natural Frequencies of Timoshenko Beams on Pasternak Foundation,'' J. Sound Vib. **51**, 149–155 (1977).
[4] B. Y. Ting and E. F. Mockry, ''Beam on Elastic Foundation Finite Element,'' J. Struct. Div. ASCE **110**(10), 2324–2339 (1984).
[5] Y. C. Lai and B. Y. Ting, ''Dynamic Response of Beams on Elastic Foundation,'' J. Struct. Div. ASCE, **118**(3) 853–858 (1992).
[6] D. G. Duffy, ''The Response of an Infinite Railroad Track to a Moving, Vibrating Mass,'' J. Appl. Mech. Trans. ASME **57**, 66–73 (1990).
[7] S. Chonan, ''Moving Harmonic Load on an Elastically Supported Timoshenko Beam,'' Z. Angew. Math. Mech. **58**, 9–15 (1978).
[8] P. M. Mathews, ''Vibrations of a Beam on Elastic Foundation,'' Z. Angew. Math. Mech. **38**, 105–115 (1958).
[9] R. W. Clough and J. Penzien, *Dynamics of Structures* (McGraw-Hill, New York, 1993).

Active control of axial-flow fan noise

Gerald C. Lauchle, John R. MacGillivray, and David C. Swanson
Penn State University, Graduate Program in Acoustics and Applied Research Laboratory, P.O. Box 30, State College, Pennsylvania 16804

(Received 14 February 1996; accepted for publication 19 July 1996)

Discrete-frequency axial-flow fan noise reduction using active noise control is described. The unique aspect of the current research is the use of the fan itself as the antinoise source in the active noise control scheme. This is achieved by driving the entire fan unit axially with an electrodynamic shaker which mechanically couples the solid surfaces of the fan to the acoustic medium. The fan unit is thus transformed into a crude loudspeaker. A near-field microphone serves as an error sensor, where transfer function measurements between the electrical input to the shaker and the electrical output of the microphone are found to be reasonably free of phase distortions and linear. A feedforward algorithm utilizing the output of a tachometer as a reference signal is used. The experimental apparatus is composed of a baffled fan unit in a free field. A small cylindrical flow obstruction is placed on the inlet side of the fan to enhance noise emissions at the blade-pass frequency and harmonics. The experiment successfully demonstrates the concept of active control of tonal fan noise using a shaken fan as the cancellation source. For the fan operating in a planar baffle, the fundamental blade-passage frequency sound-pressure level at the location of the error sensor is reduced by 20 dB, while the second and third harmonic levels are reduced by 15 and 8 dB, respectively. Placing a cabinet enclosure over the baffled fan did not affect these results significantly, and free-field sound power measurements indicate similar level reductions with the active control in operation. [S0001-4966(97)00112-4]

PACS numbers: 43.50.Ed, 43.50.Ki, 43.28.Ra [GAD]

INTRODUCTION

Under the long-wavelength (relative to rotor diameter) constraints of compactness, a subsonic fan, either baffled or unbaffled, radiates acoustic energy because of both steady and unsteady aerodynamic forces generated by the rotating blades. For typical fans operating in air, the blades are usually assumed rigid. The steady forces on fan blades generate the thrust of the fan, but because they are rotating, they also create a steady-loading noise component. The mechanism for steady-loading noise (Gutin noise) is the time-dependent distance between the individual steady forces on the blades and the observer. This component radiates sound at the blade-passing frequency (BPF) given by the number of fan blades multiplied by the shaft speed. The predominant direction of radiation is 90° to the fan axis, where the time variation in distance between any given blade and a fixed observation point is maximum. This type of noise begins to dominate other fan noise mechanisms only under the conditions of very high blade tip speeds (very high subsonic and supersonic conditions) and clean inflow/outflow conditions.

Unsteady blade forces result when the blades pass through spatially nonuniform, time invariant flow fields. Such situations occur when the fan is operated close to obstructions that can disrupt an otherwise uniform inflow. As the blades pass through these regions, the magnitude and direction of the local velocity incident to the blade sections varies with circumferential position. This gives rise to a local blade section lift and drag force that varies periodically with time. Dipole sound is produced by these fluctuating forces at harmonics of the BPF. Peak sound pressure occurs along the axis of the fan for the lower harmonics. In addition to the spatially nonuniform, time invariant flows, random flow variations due to turbulence and unsteady upstream conditions may also be present. These stochastic flow variations cause random blade forces which lead to a broadband component of sound radiation. The level of sound at the harmonics of the BPF, however, are usually many dBs above the broadband components. The BPF tones are therefore important to suppress (initially) in a fan noise reduction program.

There has been considerable interest and fundamental research on the use of active noise control (ANC) to reduce the level of discrete-frequency noise radiated from fans, blowers, and turbomachines. One of the first demonstrations of this technology was by Ffowcs Williams[1] on a British Gas Corp. gas turbine. Koopmann *et al.*[2] and Neise and Koopmann[3,4] were able to actively control the tonal emissions from a centrifugal blower operating in a duct, while Mendat *et al.*[5] achieved active attenuation of the random noise components as well. Felli *et al.*[6] demonstrated active blower noise control in a duct while using the reciprocal characteristics of a loudspeaker to permit replacement of the more conventional microphone error sensor by a loudspeaker. Ducted propeller or blower ANC is relatively straightforward to accomplish for the plane-wave propagating pressure components due to their one-dimensional nature. To actively cancel higher-order duct modes requires an array of synchronously phased antinoise sources in the duct.[3,4] Sutliff and Nagel[7] have also made progress at doing this for a ducted propeller through use of a feedforward ANC algorithm that uses a rotor blade position sensor as a reference.

Studies of active noise control for fans situated in the

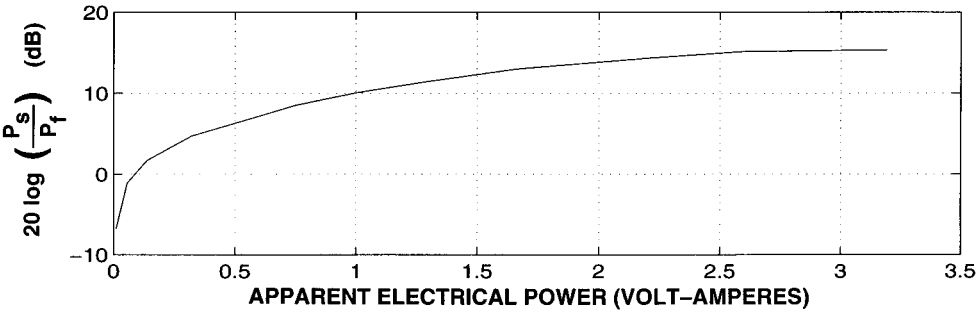


FIG. 1. On-axis far-field sound-pressure level from the shaken fan unit relative to the BPF tonal sound-pressure level radiated by the same fan in free-delivery operation. The independent variable is the apparent electrical power supplied to the mechanical shaker.

open field or in baffles are not numerous. Quinlan[8] was successful in reducing the blade-passage tones of an axial-flow cooling fan mounted in a planar rigid baffle. Located next to the fan in the baffle was the antinoise loudspeaker. The baffle apparently tends to bring the fan directivity into axial symmetry and makes it more uniform. With the antinoise source and baffled (compact) fan noise source having similar directivities, global far-field noise reduction of 10 dB was achieved for the first two harmonics of the blade-passage frequency components.

The goal of the research presented in this paper is to investigate the use of the fan itself as an antinoise source in the active control of the tonal emissions from an acoustically compact, baffled, axial-flow fan. Chiu *et al.*[9] used coherence function measurements between a small, fan-mounted force sensor and a far-field microphone to show that the radiation at the first several harmonics of the BPF was due entirely to the total integrated (over the plane of the rotor) unsteady rotor force. This result suggests that the appropriate antinoise actuator in ANC schemes for compact fans should be a mechanical shaker, or similar device that can generate controlled unsteady forces on the primary source. Furthermore, a shaken fan secondary source, if acoustically efficient, would be collocated with the primary aerodynamic fan noise sources which would be a significant advantage in global noise control where directivity issues are of concern. In the experiments described below, we use a small, commercially available electronic cooling fan as a the primary fan noise source. It is mounted directly to a electrodynamic shaker and the entire assembly is mounted in a rigid planar baffle. A feedforward ANC algorithm is used to control the shaker so that global cancellation of the far-field tonal fan noise emissions is achieved. A tachometer provides the reference signal for the algorithm, which consists of a pulse for each blade passage, and a microphone is used to supply the error signal. A feasibility study is also described that provides information on the efficiency of a shaken fan as an acoustic radiator.

I. FEASIBILITY OF USING SHAKEN FAN AS ANTINOISE SOURCE

A basic issue that needs to be addressed in the determination of whether or not a particular fan unit can be successfully implemented in the proposed ANC scheme is whether the shaken fan unit produces substantial acoustic radiation for a reasonable power input to the shaker. Another issue to be addressed is the fan noise directivity. It is desirable to have the directivity patterns of the primary and secondary sources identical. These issues are addressed experimentally for a given fan unit.

A. Shaken fan feasibility tests

Experiments involving shaker-induced radiation from the fan unit were conducted in the flow-through anechoic chamber[10] located at the Applied Research Laboratory of Penn State University. A Nidec, 82-mm-diam plastic fan was fitted with an aluminum disk on the back of its frame, which allowed for connection to a Wilcoxon, Type F3 electromagnetic shaker via a stinger which was fabricated from stainless-steel rod. The stinger was aligned along the axis of fan rotation. The shaker/fan assembly was mounted to a rigid stand such that the fan axis was vertical. Tests were conducted with the stand isolated above the anechoic wedges; the fan was unbaffled. The effect of the fan flow field on the radiation from the shaken fan unit was investigated by simply operating the fan simultaneously with the shaker.

A B&K Type 4136 microphone (1/4-in.) was suspended 1 m above the fan blades on the inlet side; its signal was analyzed on an HP 35 665A spectrum analyzer. In order to calculate the apparent electrical power[11] consumed by the shaker, a voltmeter was placed in parallel and an ammeter in series with the electrical connections to the shaker. For certain tests the fan was swiveled 180° to reverse the direction of airflow.

The fan was shaken at 264 Hz which is its free-delivery fundamental BPF. Figure 1 shows the measured on-axis acoustic pressure (p_s) at 1 m as a function of the apparent electrical power supplied to the shaker. Here, the dB level is referenced to the on-axis acoustic pressure (p_f, also at 1 m) which is produced by the normal free-delivery operation of the fan at the BPF (while the shaker was off). It is seen that the shaken fan radiation matches the pressure amplitude generated by the aerodynamic fan noise mechanisms when approximately 0.1 W of apparent power is supplied to the shaker. Figure 1 also demonstrates that with a sufficient supply of electrical power, the acoustic pressure response from the shaken Nidec fan can produce much greater acoustic pressures than the fan in operation. These results suggest that the shaken fan can act as a crude loudspeaker and is hence capable of serving as the antinoise source in ANC applications.

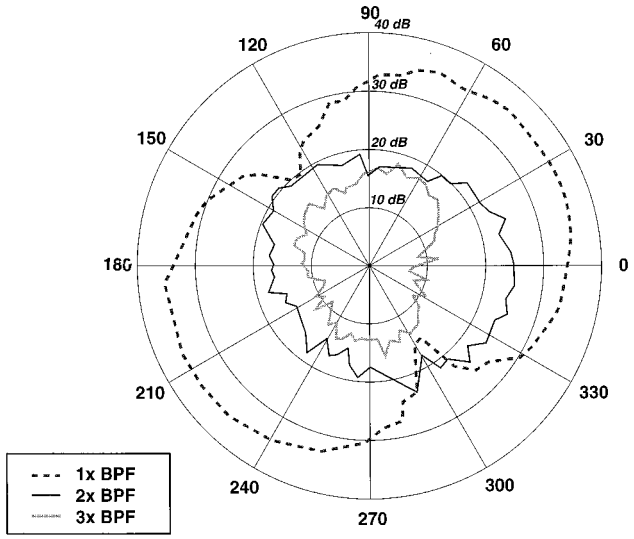


FIG. 2. Far-field directivity patterns for the unbaffled fan radiation at the first three harmonics of the BPF. The 90° axis corresponds to the fan axis on the inlet side.

Additional data of the type shown in Fig. 1 have been obtained for off-axis positions of the microphone, the fan, and shaker in operation at the same time, and again with the flow direction reversed.[12,13] In all of these situations the shaken fan was found to create sufficient acoustic energy to warrant its use as an antinoise source. However, it was observed that when the flow from the fan was reversed and directed toward the microphone (which was 1-m away and uninfluenced by the flow), the shaken fan acoustic pressure amplitude was about 2 dB less than the case when the flow was away from the microphone. This is explained as a change in radiation impedance due to flow. An analysis by Muehleisen[14] has predicted this level of change for the typical mean flow velocity of this fan.

B. Fan directivity characteristics

Unbaffled axial flow fan units are reported to produce skewed directivity patterns.[8] A nonsymmetric fan directivity pattern would suggest that complicated ways of shaking the fan would be necessary in order to most effectively cancel the fan noise. Radiation patterns for the type of fan units considered here are measured under unbaffled and baffled conditions in order to obtain knowledge about the source type and to determine the effect of the baffle on the radiation pattern. All directivity patterns were measured at 1 m from the fan in the free field. A movable microphone boom provided measurements in increments of 5°. The unbaffled directivity patterns for the first three harmonics of the BPF of the Nidec fan are shown in Fig. 2. In order to obtain results which were repeatable within ± 1.5 dB, each data point required 300 spectral averages over a frequency bandwidth of 1.6 kHz. The sampling rate was 4096 Hz. It is clear from the results shown on this figure that the unbaffled fan has complicated directivity characteristics. The pattern for the first harmonic could be interpreted as dipolelike, but it is skewed off-axis by some 60°. The patterns for the second and third harmonics cannot be interpreted in terms of simple dipole radiation patterns. These results are very consistent with those reported by Quinlan[8] for a different but similar fan. The reason for the skewness is not known precisely, but it may be a result of asymmetries in the fan construction that result in peak aerodynamic forces being directed off-axis.

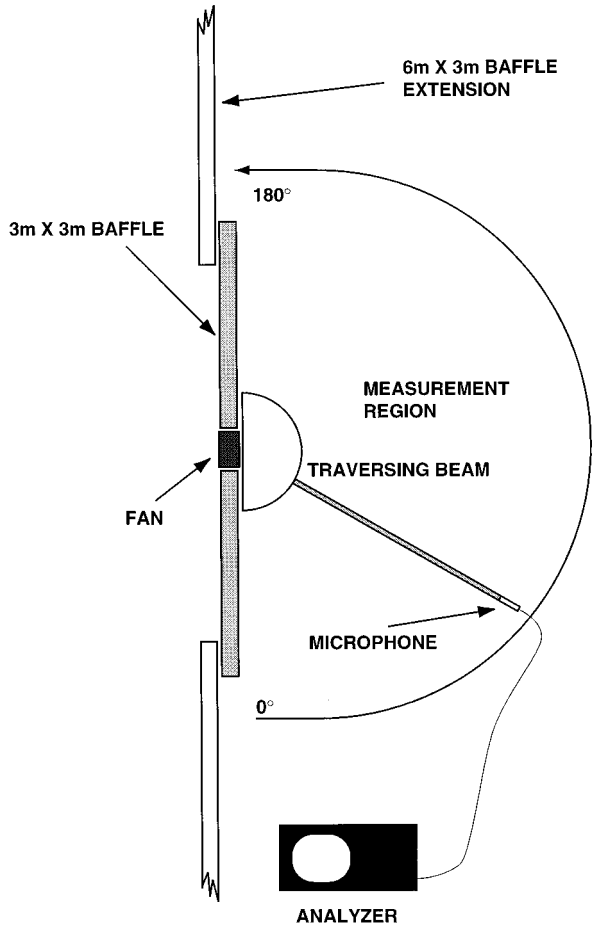


FIG. 3. Sketch of the experimental setup for measuring baffled fan directivity patterns.

When the acoustically compact fan unit is shaken as described above, one would expect dipole directivity of the sound pressure with an on-axis peak. Placement of the fan in a baffle, such as that depicted schematically in Fig. 3, would transform the dipole source directivities into monopole directivities. Figure 4 shows the baffled fan directivity patterns measured on the inlet side of the fan. The patterns are clearly more uniform than in the unbaffled case, which is favorable from the ANC viewpoint. The directivity patterns on the outlet side of the fan were measured also[13] and found to be nearly identical to those shown in Fig. 4. These experimental results confirm the findings of Quinlan.[8] They indicate that baffled fan ANC should be simpler to achieve than unbaffled fan ANC when the secondary source is of the dipole type (loudspeaker or shaken fan unit) and located in the planar baffle containing the fan.

II. EXPERIMENTAL ACTIVE FAN NOISE CONTROL

An experiment is constructed to demonstrate reduction in tonal noise from an axial-flow fan in which the fan is

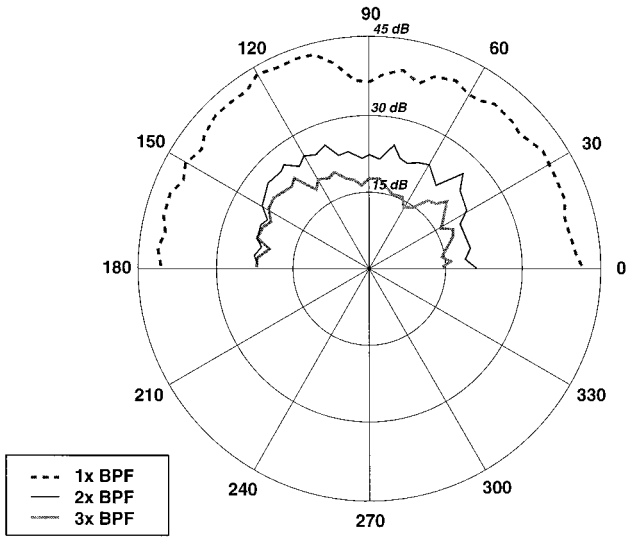


FIG. 4. Far-field directivity patterns for the baffled fan radiation from the inlet side at the first three harmonics of the BPF.

shaken to produce antisound. A Filtered-X control algorithm[15,16] is implemented and responds only to the tonal components in the fan spectrum for this experiment. The Filtered-X algorithm is a feedforward algorithm that uses an independent reference signal related to the tones of the primary source. Because the tones of interest are the harmonics of the BPF, a simple optical-type tachometer is used for the reference.

A. Experimental setup

Elements of the Filtered-X active noise control experimental setup are shown schematically in Fig. 5. In this system, the primary noise is the noise generated by the fan in operation and the secondary noise is the acoustic signal produced by the shaken fan. The error signal is the electrical output of a microphone placed on the inlet side of the fan unit, and an independent reference signal is provided by the optical tachometer. The optical sensor is placed above the fan and is sensitive to the passing of reflective strips located on the leading edge of each blade. The signal from the sensor is coherent at the BPF in addition to the higher harmonics because the voltage pulse is rectangular in shape; the Fourier transform of these periodic pulses produces a harmonic train. The optical sensor is insensitive to the motion of the axially shaken fan because the reflective strips, which the optical sensor monitors, pass orthogonally to the axis of the fan. An independent reference signal for the controller is therefore produced by the optical sensor.

The height of the error microphone above the fan unit was typically a fan diameter or more and was often moved to verify noise cancellation observations. The seven-bladed fan was mounted in a plywood baffle such that its inlet side was flush to the surface of the baffle. The gap between the fan housing and the baffle was 5 mm such that the housing was not in physical contact with the baffle. The fan was supported completely by the Wilcoxon shaker which was rigidly mounted to a platform located underneath the baffle. A small cylindrical rod was placed across the center of the fan at an axial distance of approximately $0.1R$ from the fan hub, where R is the fan blade tip radius. This distance is close enough to cause an increase in the tonal fan noise components due to the wake of the obstruction.[17] Enhancement of the BPF tones by operating the fan in a time-invariant, nonuniform inflow field seems appropriate because most practical installations result in such an inflow. The experiment was

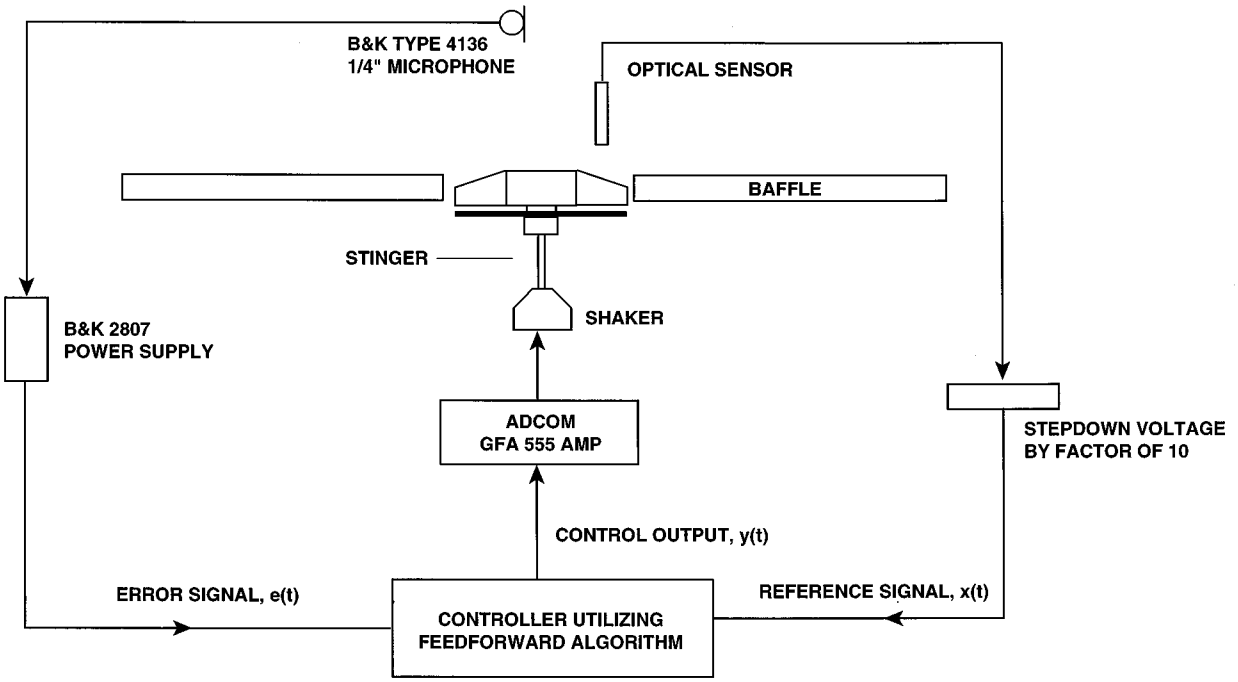


FIG. 5. Sketch of the experimental setup used to demonstrate active fan noise control.

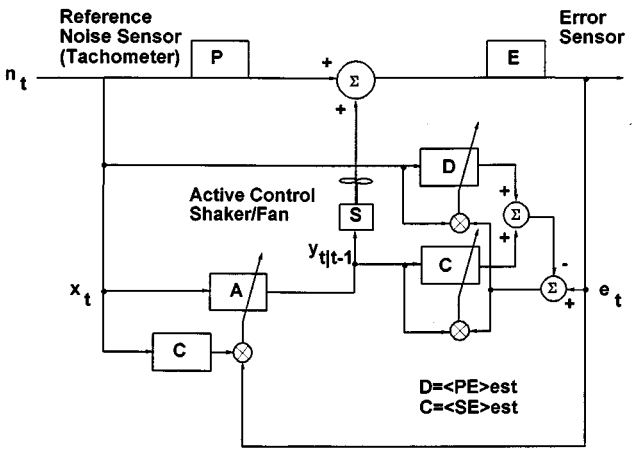


FIG. 6. Filtered-X adaptive noise cancellation block diagram used for axial fan noise reduction.

conducted in the anechoic chamber described in Sec. I A, with acoustic spectral analysis being performed using the same HP analyzer. Sound power measurements followed standard procedure.[18]

B. Control algorithm

Figure 6 presents a block diagram representing the signal processing and control for the active fan noise controller using a Filtered-X algorithm.[15] The reference signal x_t is generated by the optical sensor described in Sec. II A. The Fourier transform of the reference signal pulse train reveals exactly the same harmonic frequencies as the tonal acoustic noise generated by the fan. The block P in Fig. 6 represents the transfer function relating the magnitude and phase of the reference signal pulse train to that for the acoustic noise at the fan location. The block S represents the transfer function for the shaker/fan assembly relating the electrical input $y_{t/t-1}$ to the acoustic response at the fan location. The block E represents the transfer function of the error sensor where the input is the acoustic noise in the near field of the fan, and the output is the electrical response of the microphone to this field. Clearly, if the adaptive control filter A adapts to a transfer function approximating $-$P/S, the vibrations of the shaker/fan assembly and the unsteady forces of the blades are superimposing in a way which suppresses the radiated acoustic tonal noise. This is because the reference signal only contains harmonics coherent with the tonal acoustic noise.

We chose a passive system identification strategy[16] for the adaptive controller because we do not wish to risk increasing the broadband noise of the actively controlled fan. For the reference signal to be properly correlated to the error signal in the adaptive least-mean-square (LMS) algorithm for the controller A, we must filter x_t by the transfer function represented by the product SE, where the output is the electrical signal from the error microphone and the input is the electrical control signal to the shaker/fan assembly. The C adaptive block in Fig. 6 models the SE transfer function and the D adaptive block models the PE transfer function at the frequencies present in the reference signal x_t and the control output signal $y_{t/t-1}$. The notation $y_{t/t-1}$ depicts the fact that the digital control signal at the shaker at time t was generated by the adaptive filter last updated at time $t-1$. This delay is important physically as it indicates an unavoidable linear phase component in the SE transfer function. If the filter model C has the proper phase response at the reference signal frequencies, the Filtered-X should converge without difficulty. It is necessary to update the three LMS filters D, C, and A in real time due to the changing transfer function responses in a real fan noise application.

To insure C has the correct phase, a second system identification filter D is used to model the forward PE. The sum of the outputs of the forward model D and error model C give a prediction of the error signal e_t. The difference between the true error signal from the microphone and the prediction of the error signal is used to update the two LMS filters used in the passive system identification. Its operation is self-correcting as long as a reference signal x_t and a control signal $y_{t/t-1}$ are present (i.e., both signals are nonzero). If the control output is low, then the error signal is dominated by the forward plant and D is a good match to PE, allowing C to model SE with the residual. Conversely, if the control output is exceedingly high, the error signal is dominated by the SE loop allowing C to closely model the error plant and D to model PE with the residual. If the error signal becomes quite small (the goal of the ANC system), all of the adaptive filters slow down and converge on the desired result. Passive on-line system identification is important to axial fan ANC because flow rates and the corresponding plant time delays are always changing and the addition of broadband noise is unacceptable. It is noted that SE is the transfer function defined as the response of the error sensor when a white noise input signal to the shaker amplifier is applied in the absence of a primary excitation signal. Figure 7 shows the magnitude and phase of this function determined experimentally. The linear phase response is expected for a simple delay path associated with acoustic propagation from the fan to the error sensor.

A numerical simulation of the operations depicted in Figs. 5 and 6 has been carried out.[13] The results indicate that the algorithm of Fig. 6 operates only on causal, periodic signals. Discrete-frequency noise was thus canceled completely in the simulations. The simulation also included broadband random components of primary noise, but these were not canceled by the Filtered-X algorithm because they simply were not part of the tachometer reference signal. If a microphone reference sensor was used, rather than the tachometer, the broadband noise detected at the fan would very likely have low coherence with the far-field acoustic broadband noise due to localized turbulent flow noise at the reference position which would not be present at the far-field position. A method for suppressing local turbulent pressure fluctuations on a microphone in a flow field has recently been demonstrated,[19] but was not used in the subject investigation.

For the active fan noise control experiments, the Filtered-X algorithm was programmed in C on a WE-DSP32C floating-point digital signal processing board which was installed in an IBM PC equipped with an Intel 486DX processor operating at 33 MHz with a Windows 3.1 real-time

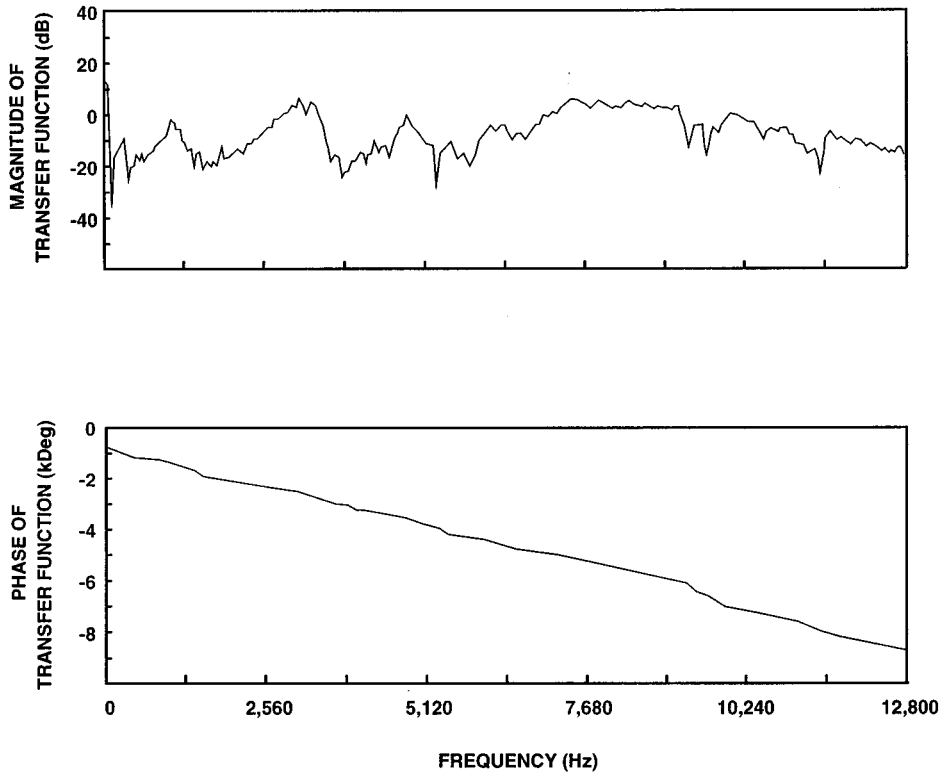


FIG. 7. Frequency response characteristics of the measured error path defined by the ratio of the Fourier transform of the input to the shaker-to-the Fourier transform of the microphone output.

user interface, also written in C. The sample rate of the controller was selected to be 2 kHz and 24 dB/octave low-pass filters were set at 900 Hz and placed at the pre-A/D and post-A/D stages of the digital signal processing board. Twenty taps (coefficients) with a step size of 0.0005 were selected for the adaptive control filter. Ten taps with a step size of 0.04 were selected for the error plant SE and ten taps with a step size of .05 were selected for the forward plant PE. On-line passive identification of the error plant (being the path representing the input to the shaker amplifier-to-the output of the error sensor) was executed in real time with the adaptive control. Using the largest step size for the D filter (in Fig. 6) modeling PE allows it to converge fastest, followed by C which models SE, and then finally the A adaptive filter, which converges to a transfer function approximating -P/S at the frequencies of the tachometer reference signal.

C. Results

Typical sound-pressure level spectra for the error sensor microphone with and without the controller on are presented in Fig. 8. The error microphone was situated approximately 19 cm above one edge of the fan frame on the inlet side. The amount of cancellation achieved at the third and higher harmonics of the BPF was found to be sensitive to the actual location of the error sensor. This is expected because the baffled fan directivity patterns become less uniform at higher harmonics of the BPF. Comparing the two spectra in Fig. 8 reveals a 20-dB reduction of the fundamental BPF tone, while the second and third harmonic levels are reduced by 15 and 8 dB, respectively. A frequency at approximately 980 Hz is also reduced. This component is related to the shaft rotation speed of 40.57 Hz and is perhaps due to a mechanical resonance of the fan unit. The fourth harmonic of the BPF at 1.136 kHz shows an 8-dB increase with the controller on. This is suspected to be due to aliasing because the low-pass filters used in the digital signal processing board were set at 900 Hz, the sampling Nyquist frequency was 1 kHz, and the 24 dB/octave filter roll-off may not be sufficient to prevent a residual component from entering the presented spectra.

The sound power radiated by the fan with and without the controller on was measured using a standard 12-point measurement procedure[18] on the inlet side of the fan over the baffle. A hemispherical surface, 0.5 m in diameter was used. Figure 9 shows the reduction in sound power level (which is indicative of global noise reduction) with the controller on as a function of frequency. The sound power at the fundamental and second harmonic BPF tones is reduced by 13 dB and 8 dB, respectively.

The sound power level reduction is some 6 or 7 dB less than the reduction in sound-pressure level measured at the error microphone location. This observation can be explained from the directivity patterns presented in Fig. 10. These patterns were measured over the baffled fan on the inlet side at the BPF, with and without the controller on. The effect of the small cylinder placed in front of the fan is evident in the ''control off'' pattern when compared to the unobstructed case of Fig. 4. The obstruction causes an approximate 10-dB increase in sound-pressure level over most observation positions. Figure 10 reveals a null in the direc-

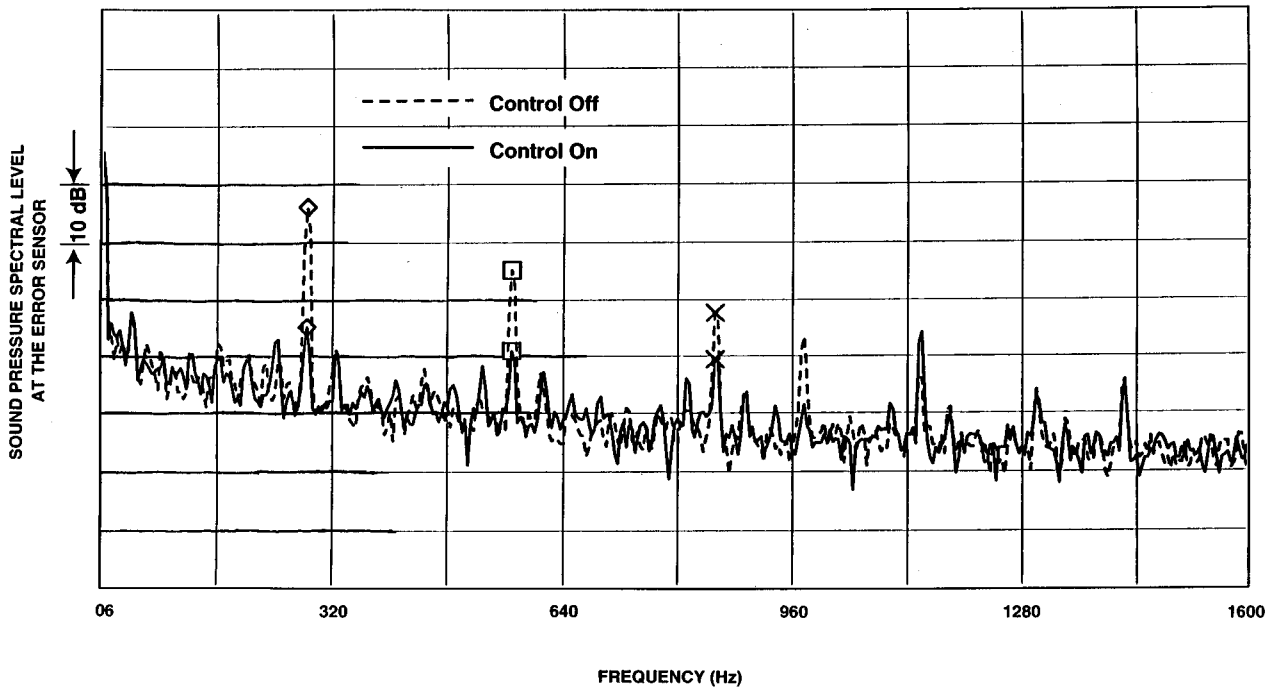


FIG. 8. Spectra of the fan sound-pressure level sensed at the error sensor position when the controller is on and off. A small cylindrical flow obstruction was placed near the fan inlet during these experiments, and the sensor is approximately 19 cm away from the edge of the fan frame, normal to the baffle.

tivity pattern along the fan axis when the control is in operation. This null clearly shows that the axial radiation is almost completely canceled by the applied axial force. Sound radiation reduction at 0° and 180° is of the order 10 dB. Obviously, these directions are less influenced by the secondary source because of its axial dipole characteristics. The error sensor was placed near the axis of the fan, so it was in the null region of the ''control on'' directivity pattern. The sound-pressure reductions are very large in this region relative to all other angular positions. This is the apparent cause of the sound power reductions being less than the sound-pressure reductions.

Additional noise reduction measurements were performed with the error microphone located on the opposite

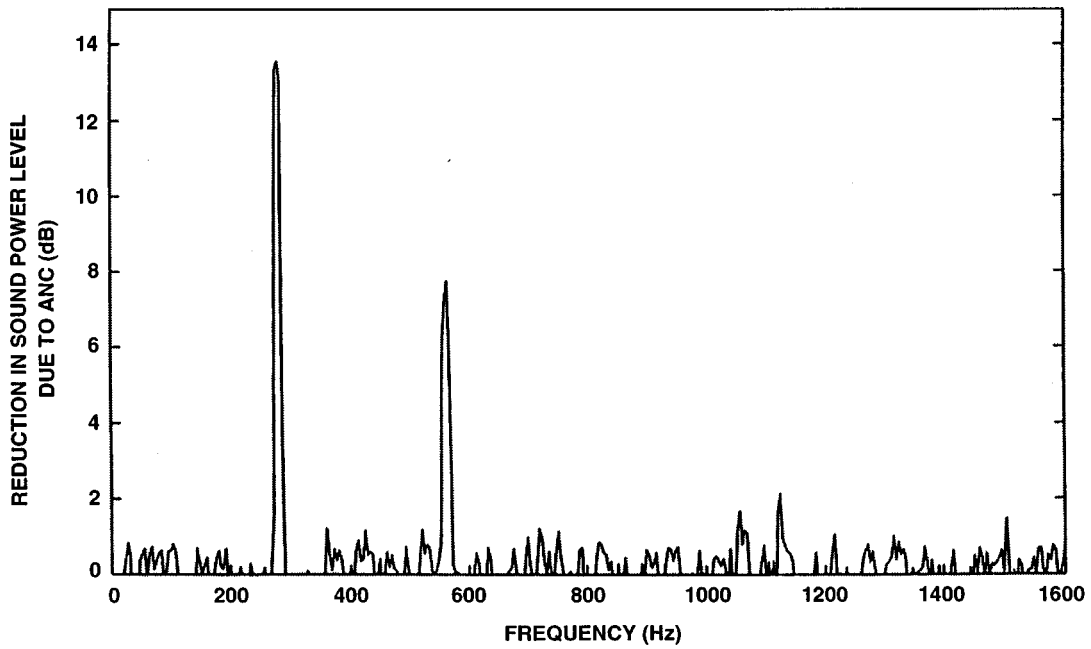


FIG. 9. Reduction in baffled axial-flow fan sound power level in dB as a function of frequency achieved using the active noise control procedures described in this paper.

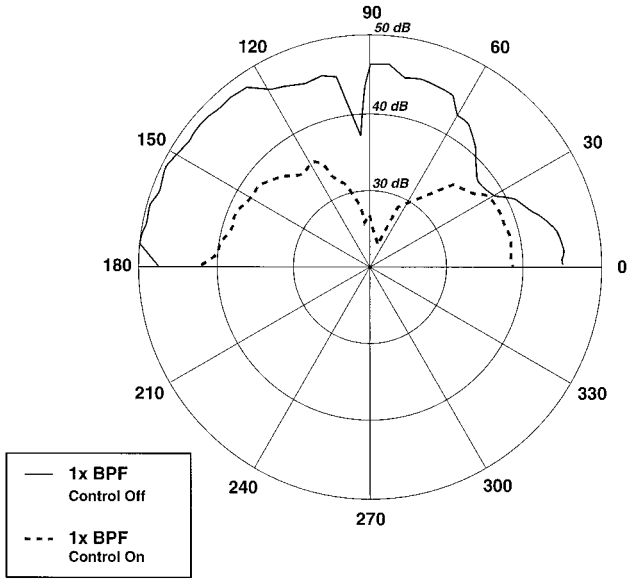


FIG. 10. Sound-pressure level directivity patterns measured for the baffled fan with the small flow obstruction in place, with and without the ANC in operation.

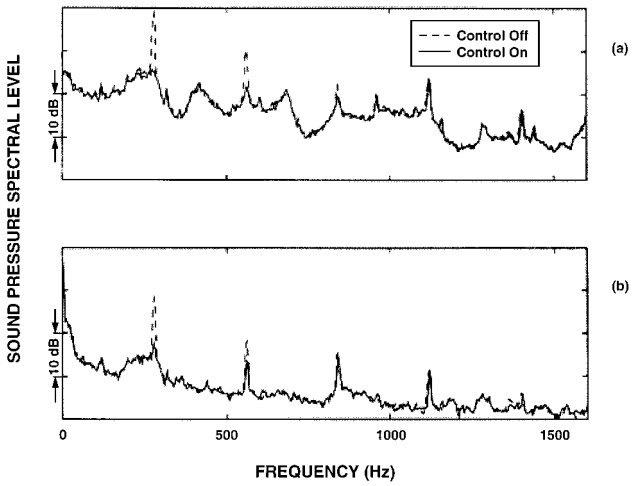


FIG. 11. Spectra of the sound pressure measured with and without the ANC in operation for the baffled fan operating inside a desktop computer cabinet: (a) the error sensor spectra measured inside the cabinet, and (b) the spectra measured at a remote location outside the cabinet.

side of the baffle (outlet side) along with a remote microphone located anywhere from 0.4 to 1.0 m away from the fan.[13] Again, the sound-pressure level at first two harmonics of the BPF was found to be reduced by 6 or more dB at the remote (and error) microphone positions. Sound power was not measured on this side of the baffle because of the proximity of the apparatus to the hard reflecting floor of the hemianechoic chamber.

As one last experiment to explore the potential applicability of the subject methodology, an empty desktop computer cabinet was placed on the planar baffle over the fan. In this arrangement, the fan pulled air into the cabinet. The air exited through the opposite side of the baffle. The error microphone was placed inside the cabinet approximately 10 cm away from the fan, and slightly off-axis. A remote microphone was positioned outside the cabinet 1 m away from the fan and 0.5 m above the baffle. Figure 11 shows the results for this experiment. The sound pressure level at the BPF is seen to be reduced by 21 dB at the external (remote) position, and by 26 dB at the internal (error microphone) position. Tone level reductions for the second harmonic are 10 and 17 dB, respectively. Sound power was not measured for the cabinet configuration, but is the subject of future applied research. The results of Fig. 11 suggest that the active noise control method employed in this study has the potential for success in a typical cooling fan application.

III. CONCLUSIONS

Although previously published research has shown that the tonal emissions from fans can be reduced through various active noise control strategies, the current research is the first to show that the fan itself, if shaken adaptively, can act as the antinoise source. This is a very significant finding because the need for a separate, secondary source is eliminated. Under the conditions of aeroacoustic compactness, the shaken fan (the secondary source) is collocated with the primary fan noise source. This tightly coupled configuration produces the excellent global noise reduction reported, and also leads to the possibility of analog feedback control strategies. This may possibly permit random noise components of the fan noise radiation spectrum, in addition to the tonal components, to be reduced. With either feedback or feedforward control, and under the premise that unsteady forces are the mechanism of subsonic fan sound production, it would be a straightforward extension of the subject methodology to utilize an internal unsteady force sensor[20] as the error sensor. Future efforts also include shaking the rotor only as opposed to shaking the entire fan assembly as was done here. Other possible modifications include multiple shakers either on individual fan blades or located at fan frame mounting lugs to aid in high-frequency cancellation when the directivity characteristics are nonuniform, and when the aeroacoustic compactness assumption is no longer valid.

ACKNOWLEDGMENTS

This work has been supported by the IBM Corp. Shared University Research Program, Dr. David Yeager, project monitor. The continuous support of the Applied Research Laboratory at Penn State University is also gratefully acknowledged.

[1] J. E. Ffowcs Williams, ''The silent noise of a gas turbine,'' Brit. Sci. News **175**, 9–12 (1981).
[2] G. H. Koopmann, W. Neise, and W. Chen, ''Active noise control to reduce the blade tone noise of centrifugal fans,'' J. Vib. Acoust. Stress Relia. Des. **110**, 377–383 (1988).
[3] W. Neise and G. H. Koopmann, ''Active source cancellation of the blade tone fundamental and harmonics in centrifugal fans,'' Proc. InterNoise 88, Avignon, France (Noise Control Foundation, Poughkeepsie, NY, 1988), pp. 801–804.
[4] W. Neise and G. H. Koopmann, ''Active sources in the cutoff of centrifugal suction fans to reduce the blade tones at higher-order duct mode frequencies,'' J. Vib. Acoust. **113**, 123–131 (1991).
[5] D. P. Mendat, K. H. Eghtesadi, M. P. McLoughlin, D. G. Smith, and E.

W. Ziegler, Jr., ''Active control of centrifugal fan noise,'' Proc. Fan Noise, an International INCE Symposium, Senlis, France (Noise Control Foundation, Poughkeepsie, NY, 1992), pp. 455–462.
[6] M. Felli, F. Cotana, and L. Santarpia, ''A loudspeaker sensor for active fan noise control: An experimental facility,'' Proc. Fan Noise, an International INCE Symposium, Senlis, France (Noise Control Foundation, Poughkeepsie, NY, 1992), pp. 471–474.
[7] D. L. Sutliff and R. T. Nagel, ''Active control of far-field noise from a ducted propeller,'' AIAA J. **33**, 231–236 (1995).
[8] D. A. Quinlan, ''Application of active control to axial flow fans,'' Noise Control Eng. J. **39**, 95–101 (1992).
[9] W.-S. Chiu, G. C. Lauchle, and D. E. Thompson, ''Subsonic axial flow fan noise and unsteady rotor force,'' J. Acoust. Soc. Am. **85**, 641–647 (1989).
[10] R. C. Marboe, G. C. Lauchle, and W. A. Kargus IV, ''Quiet wall jet facility for basic aero/hydroacoustics research,'' in ASME Publication NCA Vol. 10, *Hydroacoustic Facilities, Instrumentation, and Experimental Techniques*, edited by T. M. Farabee and R. E. Arndt (ASME, New York, 1991), pp. 69–73.
[11] W. H. Hayt and J. E. Kemmerly, *Engineering Circuit Analysis* (McGraw-Hill, New York, 1986), 4th ed., pp. 297–301.
[12] J. MacGillivray and G. C. Lauchle, ''Active control of axial flow fan noise,'' Proc. Noise-Con 94, Fort Lauderdale, FL (Noise Control Foundation, Poughkeepsie, NY, 1994), pp. 71–76.
[13] J. R. McGillivray, ''Active control of discrete-frequency noise from small subsonic fans,'' MS thesis, Penn State University, 1996.
[14] R. Muehleisen, ''Reflection, radiation, and coupling of higher order modes at discontinuities in finite length rigid walled rectangular ducts,'' Ph.D. thesis, Penn State University, 1996.
[15] P. A. Nelson and S. J. Elliott, *Active Control of Sound* (Academic, San Diego, 1992), pp. 172–175, 195–198.
[16] S. D. Sommerfeldt and J. Tichy, ''Active control of a two-stage vibration isolation mount,'' J. Acoust. Soc. Am. **88**, 938–944 (1990).
[17] K. B. Washburn and G. C. Lauchle, ''Inlet flow conditions and tonal sound radiation from a subsonic fan,'' Noise Control Eng. J. **31**, 101–110 (1988).
[18] D. A. Bies and C. H. Hansen, *Engineering Noise Control* (Unwin Hyman, London, 1988), pp. 135–139.
[19] R. S. McGuinn, G. C. Lauchle, and D. C. Swanson, ''A low flow-noise microphone for active noise control applications,'' AIAA J. (to be published).
[20] An internally mounted fan unsteady force sensor has been developed, successfully evaluated, and reported in Ref. 9.

Global attenuation of broadband noise fields using energy density control

Young C. Park and Scott D. Sommerfeldt
Graduate Program in Acoustics and Applied Research Laboratory, The Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

(Received 15 August 1995; revised 23 April 1996; accepted 7 August 1996)

The performance of the energy density control algorithm for controlling a broadband noise is evaluated in a one-dimensional enclosure. To avoid the noncausality problem of the control filter, which often happens in a frequency domain optimization, analyses presented in this paper are undertaken in the time domain. This approach provides the form of the causally constrained optimal controller. Numerical results are presented to predict the performance of the active noise control system, and indicate that improved global attenuation of the broadband noise can be achieved by minimizing the energy density, rather than the squared pressure. It is shown that minimizing the energy density at a single location yields global attenuation results that are comparable to minimizing the potential energy. Furthermore, unlike controlling the squared pressure, the energy density control does not demonstrate any dependence on the error sensor location for this one-dimensional field. A practical implementation of the energy-based control algorithm is presented. Results show that the energy density control algorithm can be implemented using the two sensor technique with a tolerable margin of performance degradation. [S0001-4966(97)02512-5]

PACS numbers: 43.50.Ki [GAD]

INTRODUCTION

Controlling the sound field in an enclosure is involved in a number of current problems of interest in active noise control. There are several active noise control algorithms derived by choosing different cost functions, such as potential energy, squared pressure, and energy density. Minimizing the potential energy yields excellent performance in terms of global attenuation.[1] However, in a practical situation, it is very difficult to measure the acoustic potential energy, so that a large number of sensors are often used to obtain an approximate measure.[2] On the other hand, the active noise control system designed to minimize the squared pressure has been widely used to control the noise in the enclosure due to the simplicity of the control structure and the efficiency in computation. However, it has been shown in previous studies[2–4] that attenuating the acoustic sound pressure at a single location in the enclosure often results in a relatively small region of control, referred to as a localized control effect.

Recently, in an attempt to simplify the control architecture, an alternative method for achieving a more global control of the sound field was developed.[3–5] This method is developed based on sensing and minimizing the energy density at discrete locations in the enclosure, so that it utilizes the concept of controlling a local variable observed at a discrete location to achieve global control. Previous work undertaken in enclosures indicates that one can often achieve improved global attenuation of deterministic signals by minimizing the acoustic energy density, rather than the squared pressure.[3–5] Also, the method has the advantage of overcoming the spillover problem that often leads to localized zones of silence when controlling the measured acoustic pressure in a field. Practical versions of such systems minimizing the energy density so far have demonstrated substantial and reliable control results in the case of deterministic signals. However, there is also a need to control broadband random noise at low frequencies in enclosures.

The objective of this study is to present numerical results that compare the global attenuation of broadband noise in a one-dimensional enclosure achieved by minimizing the energy density (which consists of the sum of the potential and kinetic energies per unit volume in the sound field), with the attenuation achieved by minimizing other acoustic parameters, such as squared pressure and potential energy. Analyses reported here are undertaken in the time domain in a manner that yields the form of the causally constrained optimal control filter.

A theoretical approach based on a frequency domain analysis enables one to establish the basic physical limitations of active noise control systems. However, this approach cannot necessarily be applied when controlling a broadband random noise, since it often yields optimal control solutions that are noncausal in the time domain,[6,7] even though such a frequency domain approach is entirely satisfactory for deterministic signals.

Another issue associated with the energy density control algorithm is that, in practical applications, multiple sensors are required to obtain error energy quantities since the acoustic velocity as well as the pressure signal should be measured to implement the algorithm.[5] In general, two microphones will be required to estimate one velocity component. Although this approach has demonstrated substantial control results, it is necessary to show that the two-microphone technique provides comparable performance to the control system employing the ideal velocity sensor. In considering a practical implementation, the optimal control filter imple-

menting the energy density control algorithm using a two-microphone sensor is presented in this paper, and its performance is demonstrated via numerical examples.

Frequency domain optimization methods based on the modal model are introduced in Sec. I. Section II presents noise control filters that are optimized under the causality constraint, and numerical results to demonstrate the performance of the optimal control filters are presented in Sec. III. Also, a two pressure microphone implementation of the energy density control is presented in Sec. IV. Section V outlines the conclusions from this work.

I. FREQUENCY DOMAIN OPTIMIZATION

The optimization of the controller in the frequency domain is based on the modal model of the sound field in an enclosure.[1,8] In this paper, the optimization is done for a one-dimensional enclosure. In addition, one primary source and one control source are considered for simplicity. To carry out this optimization, it is assumed that the enclosure is excited by a single frequency noise source. When the system is in steady state, the pressure field at the location x is given as the sum of a number of modal components,

$$p(x)=\sum_{m=0}^{\infty}(A_m+B_mQ_c)\Phi_m(x). \tag{1}$$

Here, m is the mode index, functions $\Phi_m(x)$ correspond to the eigenfunctions of the enclosure, Q_c is the complex control source strength, and the weights A_m and B_m are the modal weights associated with the primary field and the secondary control field, respectively.

The objective of the optimal control design is to compute the source strength, Q_c, so as to minimize a chosen performance function. The coefficients of the controller are optimized by using several different cost functions, such as potential energy, squared pressure, and energy density. These three performance functions can be expressed as[3]

$$I_{\mathrm{PE}}=\frac{L}{4\rho c^2}\sum_{m=0}^{\infty}|A_m+B_mQ_c|^2, \tag{2}$$

$$I_{\mathrm{SP}}=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}[A_m+B_mQ_c][A_n+B_nQ_c]^*\Phi_m(x)\Phi_n(x), \tag{3}$$

$$I_{\mathrm{ED}}=\frac{1}{2\rho c^2}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}[A_m+B_mQ_c]\times[A_n+B_nQ_c]^*F_{m,n}(x), \tag{4}$$

where

$$F_{m,n}(x)=\Phi_m(x)\Phi_n(x)+\frac{1}{k^2}\frac{\partial\Phi_m(x)}{\partial x}\frac{\partial\Phi_n(x)}{\partial x}. \tag{5}$$

Here, the subscripts PE, SP, and ED indicate cost functions corresponding to the potential energy, squared pressure, and energy density, respectively. Also, L is the length of the one-dimensional enclosure, ρ is the ambient fluid density, c is the acoustic phase speed, k is the acoustic wave number, and * denotes the complex conjugate.

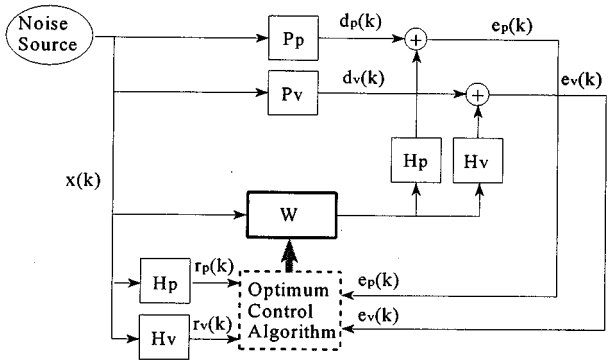


FIG. 1. A schematic diagram of the optimal noise control system.

Although the method corresponding to the potential energy is attractive for analytical work, its experimental implementation is limited by the lack of appropriate sensors to obtain a global measure of the potential energy.[2] The approach most often used in practice is the method corresponding to the cost function in Eq. (3). Since this method involves the pressure magnitude at discrete locations, it can be easily implemented. However, this approach often produces localized zones of silence instead of the desired global attenuation. The third approach, corresponding to the cost function in Eq. (4), also utilizes a local measurement, but the energy density at a discrete location yields more global information than the squared pressure control.

The results of the optimization can be expressed as[3]

$$Q_{c,\mathrm{PE}}=-\frac{\sum_{m=0}^{\infty}B_m^*A_m}{\sum_{m=0}^{\infty}B_m^*B_m}, \tag{6}$$

$$Q_{c,\mathrm{SP}}=-\frac{\sum_{m=0}^{\infty}A_m\Phi_m(x)}{\sum_{m=0}^{\infty}B_m\Phi_m(x)}, \tag{7}$$

$$Q_{c,\mathrm{ED}}=-\frac{\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}A_mB_n^*F_{m,n}(x)}{\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}B_mB_n^*F_{m,n}(x)}. \tag{8}$$

The method of minimizing the potential energy has been suggested as the optimal theoretical solution,[2] since it provides a global measure of the energy in the enclosure. The model and the optimization routines listed above are implemented in a simulation program, and results are presented in Sec. III.

II. TIME DOMAIN OPTIMIZATION

As a result of the frequency domain optimization over the whole frequency range, the optimal complex secondary source strengths are derived. However, the well-known disadvantage of this approach is that it often leads to highly noncausal impulse response functions in the time domain. Therefore, it is necessary to carry out the optimization in the time domain under the causality constraint to obtain optimal control filters that are causal in time.

A. Minimization of squared pressure

A digital model of the optimal noise control system is schematically illustrated in Fig. 1. A noise signal $x(k)$ is measured at the location of the noise source with a local

sensor. The same noise signal propagates both acoustically and structurally to the location of the error sensor, at which point it is desirable to remove the components due to the noise source. The controller drives the actuator to minimize certain parameters at the location of the error sensor. Hence, the error sensor measures the combined control actuator and the primary noise outputs as propagated to the error sensor location.

In Fig. 1, subscripts $_p$ and $_v$ indicate pressure and velocity parameters, respectively. The blocks P_p and P_v in Fig. 1 represent the transfer functions from the noise source to the acoustic pressure sensor and to the acoustic velocity sensor, respectively, and the blocks H_p and H_v denote the transfer functions from the adaptive filter output to the pressure and to the velocity sensors, respectively. Let $x(k), x(k-1), \ldots, x(k-N)$, and $w_0, w_1, \ldots, w_N$ represent reference input samples and tap coefficients of the Nth-order control filter implemented in a tapped-delay-line (TDL) structure, respectively. Also, let $h_{p,m}$, $0 \leqslant m \leqslant M$, denote the weights of the Mth-order finite impulse response (FIR) filter representing the impulse response from the filter output to the pressure error sensor. The sampled acoustic pressure signal detected by the pressure sensor is then equal to the sum of the primary pressure signal, $d_p(k)$, due to the primary noise source, and the control pressure signal due to the output of the actuator, so that

$$e_p(k) = d_p(k) + \sum_{n=0}^{N} w_n \sum_{m=0}^{M} h_{p,m} x(k-n-m). \tag{9}$$

Here, it was assumed that the filter weights, w_n, $0 \leqslant n \leqslant N$, are only slowly varying relative to the timescale of the response of the system to be controlled.[9] To simplify the equation, let $\mathbf{w}(k)$ and $\mathbf{r}(k)$ denote the $(N+1)\times 1$ weight vector and filtered reference input vector, respectively,

$$\begin{aligned} \mathbf{w} &= [w_0 \quad w_1 \quad \cdots \quad w_N]^T, \\ \mathbf{r}_p(k) &= [r_p(k) \quad r_p(k-1) \quad \cdots \quad r_p(k-N)]^T, \end{aligned} \tag{10}$$

where ^{T} denotes the matrix transpose, and the elements of the filtered reference input vector are defined as[8]

$$r_p(k-n) = \sum_{m=0}^{M} h_{p,m} x(k-n-m), \; 0 \leqslant n \leqslant N. \tag{11}$$

With the definitions in Eqs. (10), the pressure error signal can be rewritten as

$$e_p(k) = d_p(k) + \mathbf{w}^T \mathbf{r}_p(k). \tag{12}$$

An optimal weight vector can be obtained by minimizing the expectation of the square of the pressure error signal with respect to the weight vector $\mathbf{w}$. Thus, the optimal controller is designed by solving the following quadratic optimization problem:

$$\text{find} \quad \mathbf{w} \quad \text{minimizing} \quad J_{\mathrm{SP}} = E\{e_p^2(k)\}. \tag{13}$$

The cost function J_{SP} can be expanded out into a quadratic form:

$$J_{\mathrm{SP}} = E\{d_p^2(k)\} + 2\mathbf{w}^T \mathbf{P}_p(k) + \mathbf{w}^T \mathbf{R}_p(k) \mathbf{w}, \tag{14}$$

where $\mathbf{R}_p(k) = E\{\mathbf{r}_p(k)\mathbf{r}_p^T(k)\}$ and $\mathbf{P}_p(k) = E\{d_p(k)\mathbf{r}_p^T(k)\}$, respectively, denote the autocorrelation matrix of the filtered reference input and the cross correlation vector between the primary noise signal and the filtered reference input, associated with the acoustic pressure. Since the cost function J_{SP} has a unique minimum point, the gradient can be set to zero to obtain the optimal solution, given by

$$\mathbf{w}_{o,\mathrm{SP}} = -\mathbf{R}_p^{-1}(k) \mathbf{P}_p(k). \tag{15}$$

A signal processing problem related to the squared pressure control is to design an adaptive algorithm to minimize the square of the sensor output, and eventually obtain the optimal control signal by adjusting the weights of the control filter. A number of different control algorithms have been developed for implementing this active control approach. Those algorithms mostly rely on the filtered-x LMS algorithm[9–11] or the recursive LMS algorithm[12] due to the simplicity in implementing the algorithm.

B. Minimization of energy density

The acoustic energy density at the error sensor location in the field is expressed as

$$\xi = \frac{e_p^2(k)}{2\rho c^2} + \frac{\rho e_v^2(k)}{2}, \tag{16}$$

where $e_v(k)$ is the acoustic particle velocity signal at the same location where the pressure sensor is placed. These quantities specify the potential and kinetic energies per unit volume. Similar to the pressure error signal, the velocity error signal is also given as the sum of the primary velocity signal, $d_v(k)$, and the control velocity signal, i.e.,

$$e_v(k) = d_v(k) + \mathbf{w}^T \mathbf{r}_v(k), \tag{17}$$

where

$$\mathbf{r}_v(k) = [r_v(k) r_v(k-1) \cdots r_v(k-N)]^T \tag{18}$$

represents the filtered reference input vector associated with the acoustic velocity, whose elements are defined as

$$r_v(k-n) = \sum_{m=0}^{M} h_{v,m} x(k-n-m), \quad 0 \leqslant n \leqslant N. \tag{19}$$

Here, $h_{v,m}$, $0 \leqslant m \leqslant M$, denote the weights of the Mth-order FIR filter representing the impulse response from the filter output to the velocity error sensor. In this case, the designing of the optimal controller is accomplished by minimizing the expectation of the energy density function, given by Eq. (16), with respect to the weight vector $\mathbf{w}$. The optimization problem can be formulated as

$$\text{find} \quad \mathbf{w} \quad \text{minimizing} \quad J_{\mathrm{ED}} = \frac{E\{e_p^2(k)\}}{2\rho c^2} + \frac{\rho E\{e_v^2(k)\}}{2}. \tag{20}$$

Using Eqs. (12) and (17), the cost function J_{ED} can be expanded out into a quadratic form in the variable $\mathbf{w}$:

$$2\rho c^2 J_{\mathrm{ED}} = d_p^2(k) + (\rho c)^2 d_v^2(k) + 2\mathbf{w}^T \times [\mathbf{P}_p(k) + (\rho c)^2 \mathbf{P}_v(k)] + \mathbf{w}^T [\mathbf{R}_p(k) + (\rho c)^2 \mathbf{R}_v(k)]\mathbf{w}. \quad (21)$$

For an input noise signal that is a broadband white noise signal, this cost function is a positive definite quadratic function of the controller weights, so that it has a unique global minimum point.[5] The derivative of this cost function is expressed as

$$\nabla(2\rho c^2 J_{\mathrm{ED}}) = 2[\mathbf{P}_p(k) + (\rho c)^2 \mathbf{P}_v(k)] + 2[\mathbf{R}_p(k) + (\rho c)^2 \mathbf{R}_v(k)]\mathbf{w}. \quad (22)$$

We can set the gradient to zero to obtain the weight vector of the optimal controller:

$$\mathbf{w}_{o,\mathrm{ED}} = -[\mathbf{R}_p(k) + (\rho c)^2 \mathbf{R}_v(k)]^{-1} \times [\mathbf{P}_p(k) + (\rho c)^2 \mathbf{P}_v(k)], \quad (23)$$

where $\mathbf{R}_v(k) = E\{\mathbf{r}_v(k)\mathbf{r}_v^T(k)\}$ and $\mathbf{P}_v(k) = E\{d_v(k)\mathbf{r}_v^T(k)\}$ denote the autocorrelation matrix of the filtered reference input and the cross correlation vector between the primary noise signal and the filtered reference input, associated with the acoustic velocity, respectively. Since the energy density is controlled to try to achieve global attenuation of the sound field, the adaptive algorithm associated with the energy density-based control will involve two independent error signal components: pressure and velocity. An adaptive algorithm to obtain the optimal weight vector for the energy density control filter was fully developed and tested in Ref. 5. The algorithm in Ref. 5 was developed based on the filtered-x LMS algorithm, using an approach similar to the algorithm minimizing the sum of the squared pressure errors provided by multiple error sensors.

C. Minimization of potential energy

In a practical situation, it is very difficult to measure the acoustic potential energy, defined as the volume integral of the time-averaged acoustic potential energy density. However, to obtain an approximate measure it might be possible to use a number of acoustic pressure sensors evenly distributed over the entire enclosure. When an array of acoustic pressure sensors are used to approximate the potential energy in the one-dimensional enclosure, the cost function for the optimization can be expressed as

$$\hat{J}_{\mathrm{PE}} = \frac{1}{4\rho c^2} E\left\{ \sum_{i=0}^{N_e-1} e_p^2\left[\left(i + \frac{1}{2} \right) \delta x_e, k \right] \delta x_e \right\}, \quad (24)$$

where N_e represents total number of sensors, δx_e is distance between sensors which is equal to $1/N_e$, and $e_p[(i+\frac{1}{2})\delta x_e, k]$ is the output of the pressure error sensor located at $(i+\frac{1}{2})\delta x_e$.

The optimal control filter can be obtained by minimizing the cost function $\hat{J}_{\mathrm{PE}}$. This cost function is also a quadratic function of the controller weight vector. Furthermore, it has a unique global minimum value for the weight vector, so that the solution for $\mathbf{w}$ minimizing the potential energy in the enclosure estimated using a microphone array can be found by setting the gradient of the function $\hat{J}_{\mathrm{PE}}$ to zero. The optimal weight vector of the control filter is given by

$$\hat{\mathbf{w}}_{o,\mathrm{PE}} = -\left[\sum_{i=0}^{N_e-1} \mathbf{R}_{p,i}(k) \right]^{-1} \left[\sum_{i=0}^{N_e-1} \mathbf{P}_{p,i}(k) \right]. \quad (25)$$

Here, $\mathbf{R}_{p,i}(k)$ and $\mathbf{P}_{p,i}(k)$, respectively, represent the autocorrelation matrix and cross correlation vector corresponding to the ith sensor.

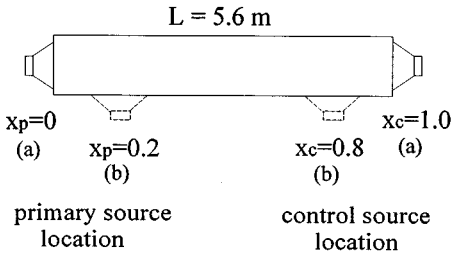


FIG. 2. System configuration considered for the computer simulation.

III. NUMERICAL RESULTS

The weight vectors of optimal controllers are derived from *ensemble averages* with the result that one filter optimum is obtained in a probabilistic sense for all realizations of the operational environment, assumed to be wide-sense stationary. However, in practical applications a nonadaptive and optimum control design involves the use of time averages, with the result that the filter depends on the number of samples used in the computation. When it is assumed that the control system is not time varying and there are sufficient input samples available, the optimal controllers are designed by minimizing time-averaged acoustic parameters, such as squared pressure, energy density, and potential energy. In this section, results from computer simulations used to test the performance of the optimal controllers are shown. In the computer simulations, the time average was used to form the cost functions being minimized.

Simulations are conducted for a one-dimensional enclosure with length $L = 5.6$ m. The simulation model considered here is presented in Fig. 2. For convenience the length of the enclosure is normalized to 1.

The model and the optimization process were implemented in simulation programs running on a PC. Based on the enclosure configuration, the frequency responses of the primary and control paths associated with the acoustic pressure and velocity were calculated using the modal model of the sound field,[1,8] with an assumed modal damping coefficient of 0.05. Impulse response estimates, i.e., p_p, p_v, h_p, and h_v, were then computed from the frequency responses. Each path was modeled as a 256-tap FIR digital filter. Prior to using the impulse response estimates in the simulation, a comparison was made between the frequency responses of the modal model and the 256-tap FIR model to establish that 256 taps was sufficient to model the primary and control paths to the accuracy required. The broadband noise signal $x(k)$ was taken to be white noise filtered through a bandpass filter with a pass band from 50 to 350 Hz, and the sampling frequency was set to 1000 Hz. Measurement noises which

are uncorrelated with the signal were added to the noise signal. The level of the measurement noise was set at -40 dB below the signal level.

Using the signal model and the impulse response estimates, 20 000 samples of the primary and the filtered reference signals were generated for both the acoustic pressure, as well as the velocity. Then, the autocorrelation matrices $\mathbf{R}_p(k)$ and $\mathbf{R}_v(k)$, and the cross correlation vectors $\mathbf{P}_p(k)$ and $\mathbf{P}_v(k)$, were estimated using the time average over the entire input samples. Finally, the weight vectors of the optimal controllers were computed using Eqs. (15), (23), and (25).

The global control of the control filter optimized in the time domain under the causality constraint was measured with the averaged power spectral density (PSD). The PSD was computed in three steps. First, frequency responses of the noise and control signal paths were calculated using the 256-tap FIR model described above. Residual signals at each of the discrete locations used were then computed using the FIR models and the noise samples. Finally, the PSD of each residual signal was computed and averaged using the equation given by

$$\overline{S(f\ )}=\frac{1}{N_o}\sum_{i=0}^{N_o-1} S\left[\left(i+\frac{1}{2}\right)\delta x_o, f\ \right], \tag{26}$$

where N_o denotes the total number of observation points, $S(x,f\ )$ is the PSD of the acoustic pressure at position x, and δx_o is the distance between adjacent observation points which is equal to $1/N_o$.

In the first configuration, denoted by (a) in Fig. 2, the primary source is positioned at one end ($x_p=0$) and the control source is placed at the other end ($x_c=1.0$). The error sensor location (x_e) is 0.7. Since the location of the error sensor is closer to the controller location than the primary source location, this configuration constitutes a causal situation for the control filter. With this configuration and FIR control filters having 64 taps, performances of the different optimization schemes were evaluated and compared with each other. Throughout the simulations, 50 microphones evenly distributed along the enclosure were used to estimate the potential energy, i.e., $N_e=50$, and the same number of microphones were used to compute the averaged PSD, i.e., $N_o=50$.

A global measure of the control that results from the frequency domain optimization is given by the potential energy in the enclosure both before and after the control is applied. In computing the optimal control strength and the potential energy, the infinite sum in Eqs. (2), (6), (7), and (8) were truncated to include the first 1500 modes. The potential energy in the enclosure as a function of frequency is shown in Fig. 3(a), and the averaged PSD achieved by using the time domain optimization is shown in Fig. 3(b), also as a function of frequency. In the figures, results obtained by minimizing the three acoustic parameters of squared pressure, energy density, and potential energy, are indicated by squared pressure control, energy density control, and potential energy control, respectively. From the results in Fig. 3(a) it can be seen that the minimization of the potential energy yields the lowest global energy, as is to be expected. However, minimizing the energy density at the single discrete location leads to results that are comparable to the results obtained by minimizing the potential energy. In fact, little difference can be found between the results indicated by energy density control and potential energy control over the entire frequency band, so that one curve is almost covered by another. On the other hand, minimizing the squared pressure actually increases the global potential energy in the enclosure at some frequencies. Furthermore, minimizing the squared pressure yields higher potential energy than minimizing the energy density as well as the potential energy at most frequencies.

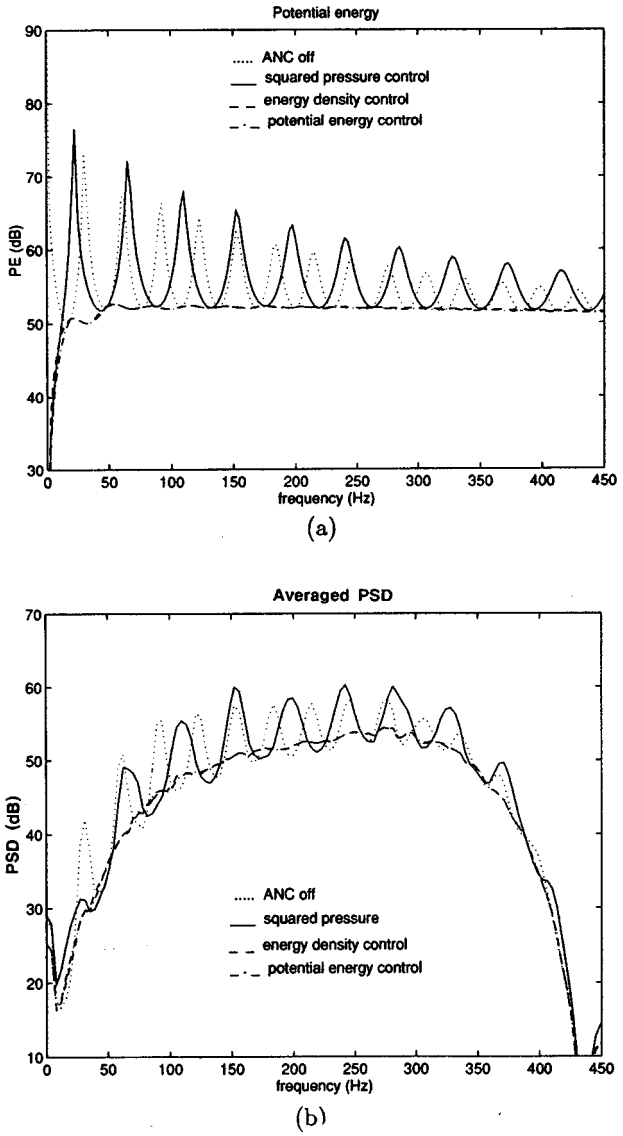


FIG. 3. Global measure of noise fields before and after the controller is applied: (a) frequency domain optimization results, (b) time domain optimization results ($x_p=0$, $x_c=1.0$, $x_e=0.7$).

Similar trends can be observed in the time domain optimization results. It should be remembered that the input signal was bandlimited from 50 to 350 Hz. Thus, the comparison between the frequency domain and time domain optimizations can be made only in that frequency range. As can be seen from Fig. 3(b), minimizing the potential energy

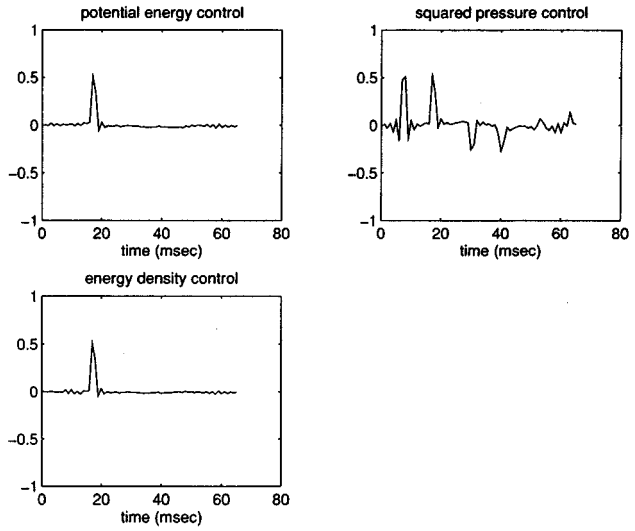


FIG. 4. Impulse responses of the optimal control filters ($x_p=0$, $x_c=1.0$, $x_e=0.7$).

yields the lowest averaged PSD results, and minimizing the energy density at a single discrete location provides results that are similar to the case of minimizing the potential energy. However, minimizing the squared pressure increases the PSD level at some frequencies, and shows poor performance at most frequencies compared to the case of minimizing the energy density as well as potential energy. Additional insight into the control effect achieved with each of the control approaches can be gained by looking at the impulse responses of the optimized controllers. Figure 4 shows the impulse responses of the FIR optimal controllers. It is clearly indicated in Fig. 4 that the energy density control provides the control filter with an impulse response which is almost identical to the one obtained by minimizing the potential energy.

In the next simulation, the error sensors were placed at the normalized position of 0.3. Since the error sensor location is closer to the primary source than the controller, this configuration would lead to a noncausal controller. Figure 5(a) shows the potential energy in the enclosure using frequency domain optimization. The averaged PSD obtained by using the time domain optimization and the corresponding impulse responses are shown in Figs. 5(b) and 6, respectively. From the results in Fig. 5(a), it can be seen that minimizing the energy density at a single discrete location leads to global potential energy results which are comparable to the case of minimizing the potential energy, while the potential energy is again significantly increased by minimizing the squared pressure at some frequencies.

However, since the configuration being tested can be considered a noncausal situation in terms of the error sensor location, the results in Fig. 5(a) cannot necessarily be used to predict the performance of the control filter being implemented in the time domain. Using the time domain optimization technique, on the other hand, one can predict the exact performance of the control filter since the optimization is undertaken in a manner that satisfies the causality constraint. Also, the control results obtained by using the time domain

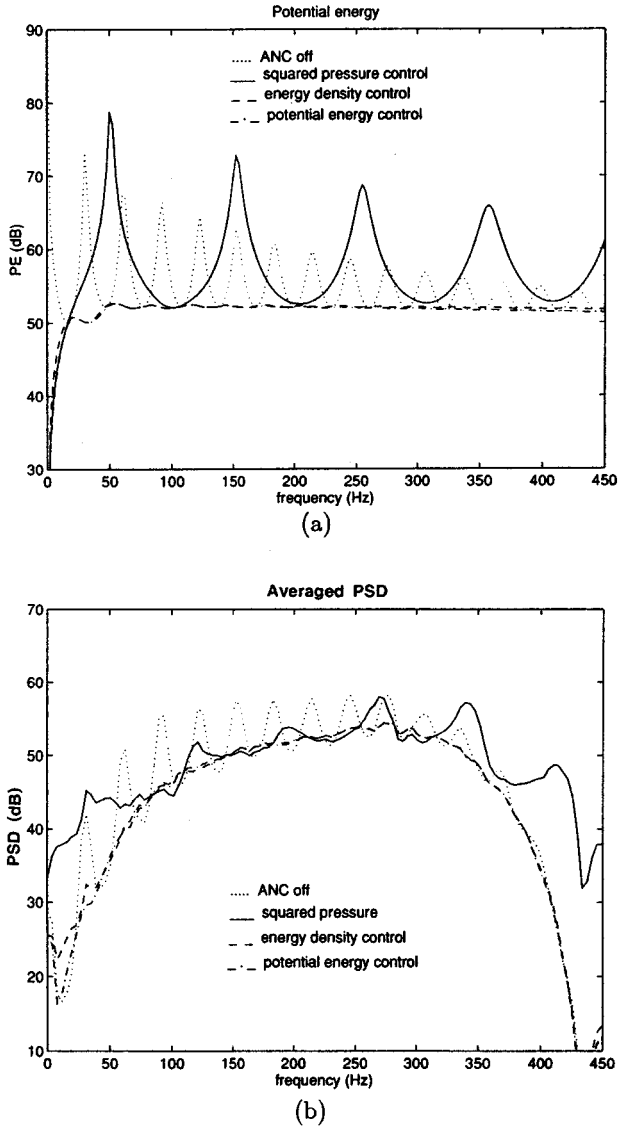


FIG. 5. Global measure of noise fields before and after the controller is applied: (a) frequency domain optimization results, (b) time domain optimization results ($x_p=0$, $x_c=1.0$, $x_e=0.3$).

optimization can differ from the results obtained by using the frequency-domain approach. As can be seen from Fig. 5, the case of minimizing the squared pressure demonstrates a difference in control results between the frequency-domain optimization and the time domain optimization. However, minimizing the squared pressure still increases the PSD levels significantly at some frequencies, the 0–100 Hz and 300–450 Hz regions in particular. However, minimizing the energy density shows the results which are similar to the case of minimizing the potential energy. Figure 6 shows the impulse responses of the optimal controllers. It is clearly shown from Fig. 6 that minimizing the energy density provides the control filter with a consistent impulse response. On the other hand, the control filter designed to minimize the squared pressure does not show the same impulse response that was obtained in the causal configuration. As a result, the algorithm yields an inconsistency in the performance as can be seen from Figs. 3(b) and 5(b). These results imply that the

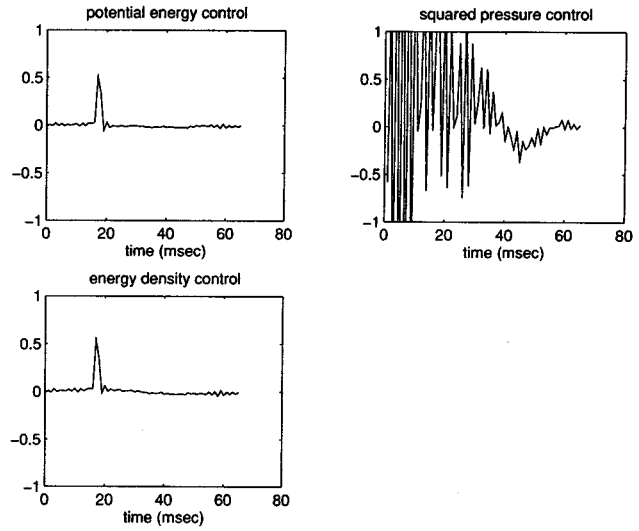


FIG. 6. Impulse responses of the optimal control filters ($x_p=0$, $x_c=1.0$, $x_e=0.3$).

performance of the controller designed to minimize the squared pressure is sensitive to the error sensor location, which imposes a limitation on practical implementations of the algorithm.

For further investigation of the inconsistency problem which may happen to the control filters in a noncausal situation, impulse responses of the optimal control filters designed to minimize the energy density and the squared pressure at several error sensor locations are illustrated in Fig. 7. The control filters were optimized for each x_e, which varies from 0.25 to 0.75. The energy density control method results in consistent control filter weights for the entire range of error sensor locations considered, while controlling the squared pressure provides highly inconsistent control filters for the error sensor locations $x_e \leqslant 0.5$, which constitute noncausal situations. The results in Fig. 7 indicate that the placement of the error sensor is not a critical issue when using energy density control, which is a significant advantage in practical applications.

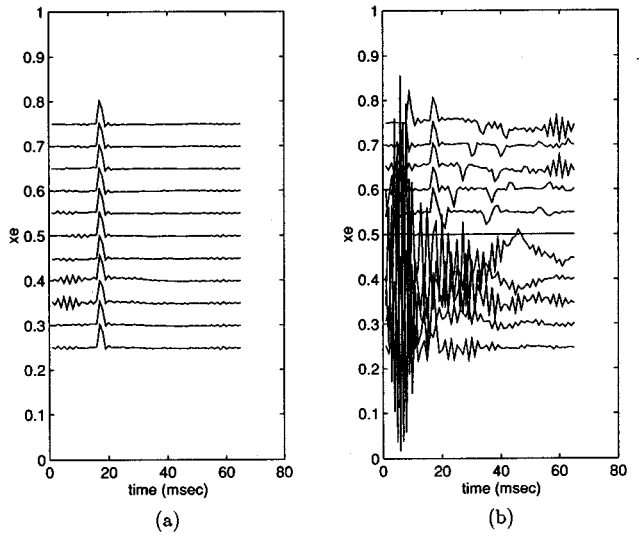


FIG. 7. Impulse responses of the control filter optimized with error sensor locations varying from 0.25 to 0.75: (a) energy density control, (b) squared pressure control ($x_p=0$, $x_c=1.0$).

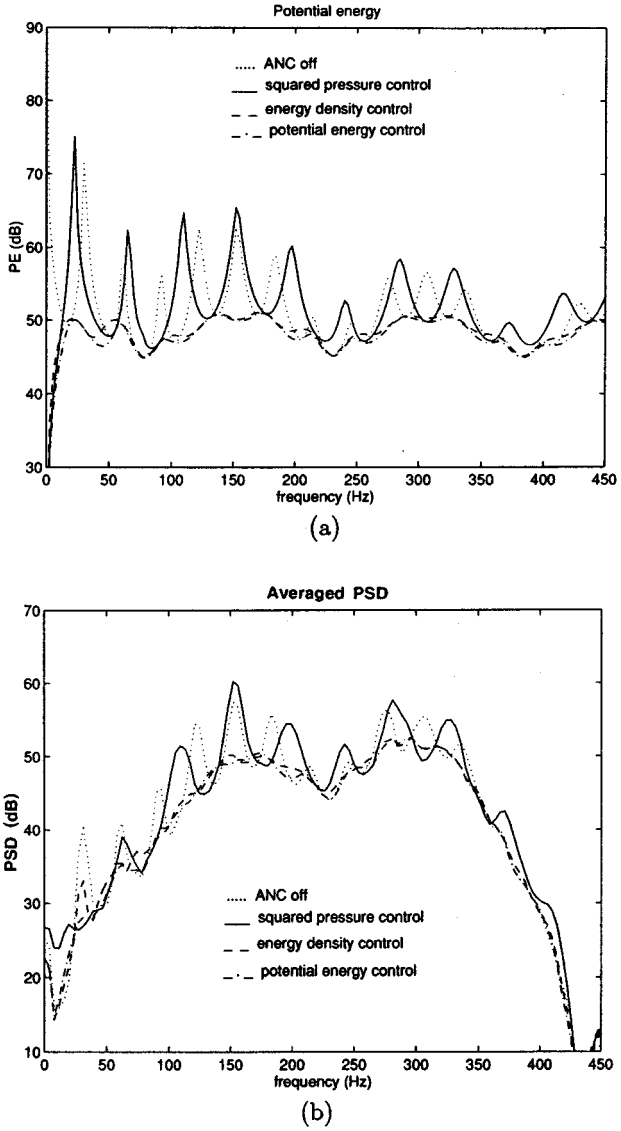


FIG. 8. Global measure of noise fields before and after the controller is applied: (a) frequency domain optimization results, (b) time domain optimization results ($x_p=0.2$, $x_c=0.8$, $x_e=0.7$).

The performance of the optimization schemes was evaluated for a different configuration of the enclosures, which is indicated by (b) in Fig. 2. In this case, the primary source is located at 0.2 and the controller is placed at 0.8. The controller position was chosen to avoid the nodal points of resonance frequencies of the enclosure, such as 30.6 Hz, 61.25 Hz, 91.8 Hz, and so on. However, it should be stated that the optimum location for the controller is not the issue investigated in this study.

Simulations were performed for two different error sensor locations: 0.7 and 0.3. Figures 8 and 9 show the results obtained with the error sensor located at 0.7. This error sensor location constitutes the configuration which provides a causal solution. Both the frequency domain optimization as well as the time domain optimization show results which are

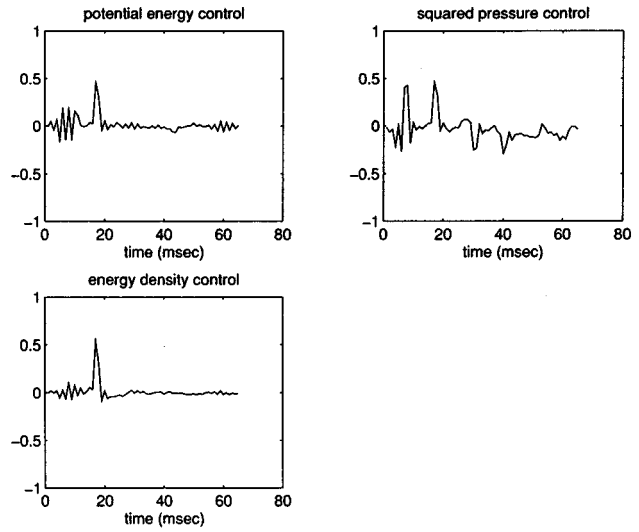


FIG. 9. Impulse responses of the optimal control filters (x_p=0.2, x_c=0.8, x_e=0.7).

similar to the previous ones obtained with the configuration x_p=0, x_c=1.0, and x_e=0.7, i.e., the energy density control shows global attenuation results which are better than the case of minimizing the squared pressure for most frequencies and comparable to the case of minimizing the potential energy.

Figures 10 and 11 show the results obtained with the error sensor located at 0.3. Minimizing the squared pressure in the time domain produces an impulse response which is different from the one obtained with the configuration x_p=0.2, x_c=0.8, and x_e=0.7. It also increases the PSD level at most frequencies. The energy density control, on the other hand, provides the control filter that is similar to the potential energy control. Figure 12 shows impulse responses of the optimal control filter designed to minimize the energy density and the squared pressure at several error sensor locations. These results again prove that, unlike the method of controlling the squared pressure, the performance of the energy density control does not depend on the error sensor location. However, the same is not true for the control source location, and, as mentioned earlier, that is not the issue of interest in this paper.

IV. IMPLEMENTATION CONSIDERATIONS FOR ENERGY DENSITY CONTROL

The implementation of the energy density control algorithm requires measurements of the acoustic velocity as well as the pressure at the error sensor location. The acoustic velocity in a one-dimensional enclosure can be obtained using a particle velocity sensor, such as a laser vibrometer or velocity microphone, or using a two-microphone technique which is typically used to measure the acoustic intensity. In this section, the energy density control algorithm is implemented in a one-dimensional enclosure, using the two-microphone technique, and its performance is evaluated.

It is generally assumed that two highly phase-matched microphones are required to obtain energy quantities when a two-microphone technique is used. However, it has been

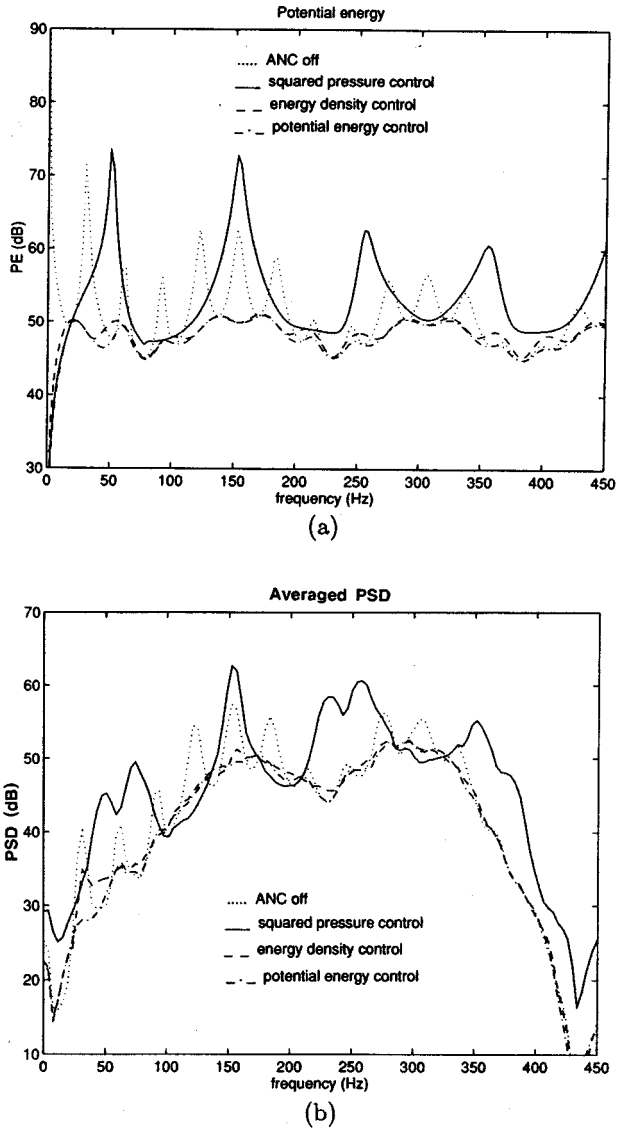


FIG. 10. Global measure of noise fields before and after the controller is applied: (a) frequency domain optimization results, (b) time domain optimization results (x_p=0.2, x_c=0.8, x_e=0.3).

shown in previous research[4,5] that low cost microphones offered by some manufacturers are sufficiently stable so that two reasonably well phase-matched microphones can be found without too much difficulty. Also, it should be mentioned that since the control system will be sensitive to all modes associated with the field in the enclosure, small measurement errors caused by the phase mismatch of the microphones are tolerable to maintain the performance desired. To conduct the analysis here, it is assumed that the magnitude and phase responses of two pressure microphones are exactly matched.

When the two sensor approach is used, the pressure and velocity in a one-dimensional enclosure are estimated using the equations, given by

$$\hat{e}_p(k)=\frac{e_{p1}(k)+e_{p2}(k)}{2}, \tag{27}$$

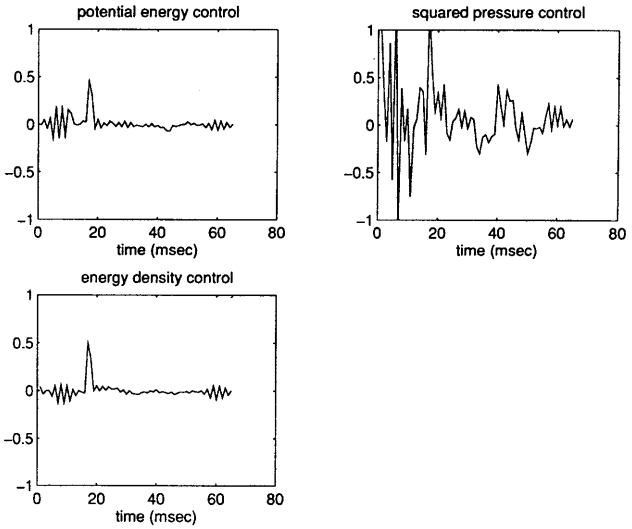


FIG. 11. Impulse responses of the optimal control filters (x_p=0.2, x_c=0.8, x_e=0.3).

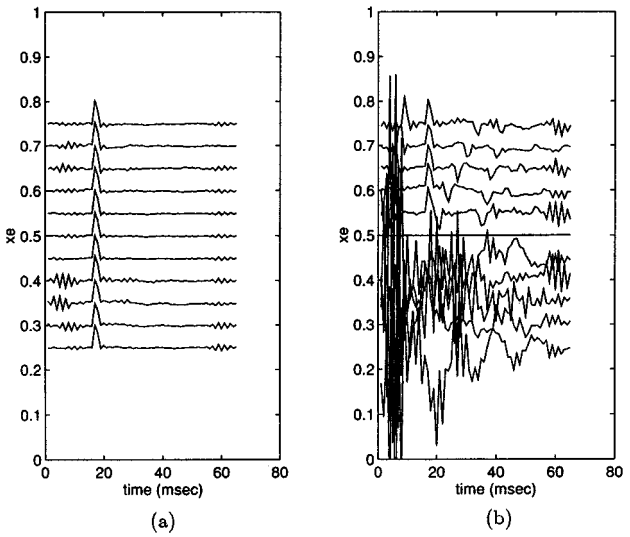


FIG. 12. Impulse responses of the control filter optimized with error sensor locations varying from 0.25 to 0.75: (a) energy density control, (b) squared pressure control (x_p=0.2, x_c=0.8).

$$\hat{e}_v(k) = -\frac{1}{\rho}\,\Xi\left\{\int_t \frac{e_{p2}(t)-e_{p1}(t)}{\Delta x}\,dt\right\}, \tag{28}$$

where $e_{p1}(k)$ and $e_{p2}(k)$ are the pressure measurements from two closely spaced microphones, Δx is the spacing between them, and $\Xi\{.\}$ denotes the continuous-to-discrete time transformation. The integration can be done using an analog integrator. However, more reliable and accurate results can be obtained by using a digital integrator. There are several possible ways of designing the digital integrators. An example of designing the digital integrator can be found in Ref. 13. Using the digital integrator, the velocity estimate can be expressed in a simple recursive form:

$$\hat{e}_v(k) = \hat{e}_v(k-1) + \alpha[e_{p2}(k)-e_{p1}(k)]\exp(-1/f_s), \tag{29}$$

where $\alpha = -1/(\rho \Delta x f_s)$ and f_s denotes the sampling frequency.

Figure 13 shows the schematic diagram of the energy density control system being implemented using the two-sensor technique. Outputs of two pressure sensors are used to estimate the acoustic pressure and velocity at the position. The estimated signals are applied to the control filter. The control filter is optimized based on the error estimates and the filtered reference inputs. To compute the filtered reference inputs, denoted by $\hat{\mathbf{r}}_p(k)$, and $\hat{\mathbf{r}}_v(k)$ in Fig. 13, the control path between the controller output and error estimates are copied to the control algorithm. In this case the control path comprises the error estimation process including the digital integrator.

Using the pressure and velocity error estimates, the weight vector of the optimal controller minimizing the energy density is expressed as

$$\hat{\mathbf{w}}_{0,\mathrm{ED}} = -[\hat{\mathbf{R}}_p(k)+(\rho c)^2\hat{\mathbf{R}}_v(k)]^{-1} \times[\hat{\mathbf{P}}_p(k)+(\rho c)^2\hat{\mathbf{P}}_v(k)], \tag{30}$$

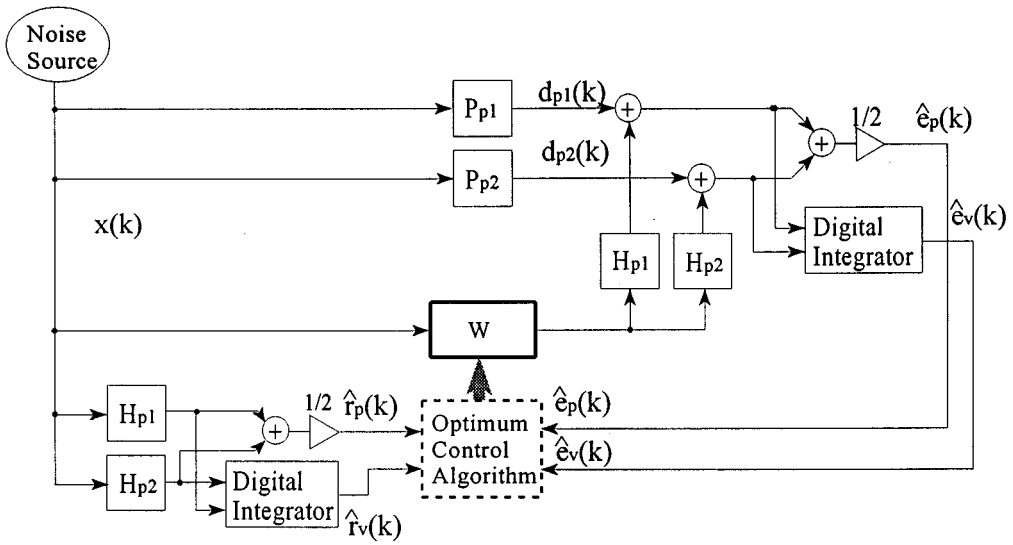


FIG. 13. Schematic diagram of the energy density control system implemented using two pressure microphones.

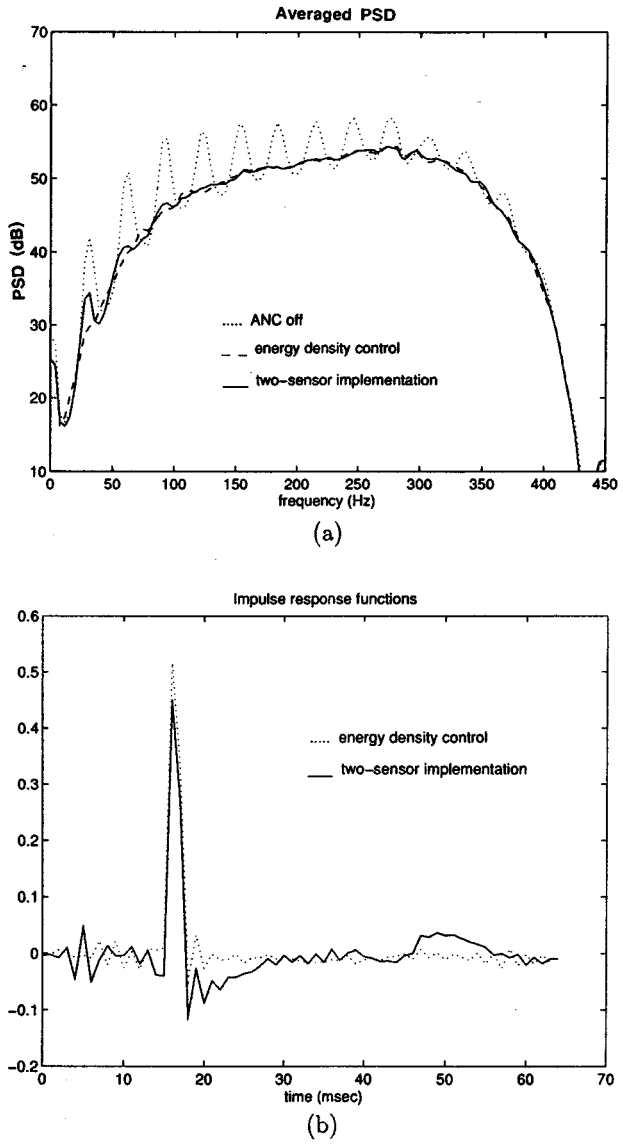


FIG. 14. The performance of the energy density control algorithm being implemented using the two-microphone technique: (a) averaged PSDs, and (b) impulse responses ($x_p=0$, $x_c=1.0$, $x_e=0.7$).

where $\hat{\mathbf{R}}_p(k)$, $\hat{\mathbf{R}}_v(k)$, $\hat{\mathbf{P}}_p(k)$, and $\hat{\mathbf{P}}_v(k)$ are the autocorrelation matrices and cross correlation matrices estimated using $\hat{\mathbf{r}}_p(k)$, $\hat{\mathbf{r}}_v(k)$.

Simulations were conducted for the case $x_p=0$, $x_c=1.0$, and $x_e=0.7$ to evaluate the performance of the two-sensor implementation. Two pressure microphones were spaced a distance of 5.0 cm (0.009 in the normalized scale). Figure 14(a) and (b) shows the averaged PSD and the impulse response. For comparison purposes, the results obtained by using the energy density control system employing the ideal velocity signal are reproduced in Fig. 14. There are several discrepancies in the results. However, overall it can be seen that the two-sensor approach yields performance which is comparable to that of the ideal energy density control. Based on these results, it can be concluded that the energy density control can be implemented using a two sensor technique without significant degradation in the performance of the original algorithm.

V. CONCLUSIONS

In this paper the performance of the energy density control method for global attenuation of broadband noise in a one-dimensional enclosure was evaluated. Numerical results were presented to compare the global attenuation achieved by minimizing the energy density with the attenuation achieved by minimizing the squared pressure or the potential energy. Optimal control filters were designed in the time domain to yield solutions that are causal in the time domain.

Numerical results have indicated that greater global control of the sound field can be achieved by minimizing the energy density, rather than the squared pressure. It has also been shown that minimizing energy density at a single location in a one-dimensional enclosure produces the global control that one would achieve by minimizing the potential energy. Another significant advantage of energy density control is that, unlike the case of minimizing the squared pressure, this control method does not demonstrate any dependence on the error sensor location. Thus, by controlling the energy density, one can overcome the limitations of the possible locations for the error sensors which exist in practical situations.

A two microphone implementation of the energy-based control algorithm was presented. Simulation results indicate that the energy density control can be implemented using the two-sensor technique with a tolerable margin of performance degradation.

[1] P. A. Nelson, A. R. D. Curtis, S. J. Elliott, and A. J. Bullmore, ''The active minimization of harmonic enclosed sound fields, Part I: Theory,'' J. Sound Vib. **117**, 1–13 (1987).

[2] A. R. D. Curtis, P. A. Nelson, and S. J. Elliott, ''Active reduction of a one-dimensional enclosed sound field: An experimental investigation of three control strategies,''J. Acoust. Soc. Am. **88,** 2265–2268 (1990).

[3] S. D. Sommerfeldt and J. W. Parkins, ''An evaluation of active noise attenuation in rectangular enclosures,'' Proc. Inter-Noise 94, Yokohama, 1351–1356 (1994).

[4] S. D. Sommerfeldt, J. W. Parkins, and Y. C. Park, ''Global active niose control in rectangular enclosures,'' Proc. Active-95, Newport Beach, CA, 477–488 (1995).

[5] S. D. Sommerfeldt and P. J. Nashif, ''An adaptive filtered-*x* algorithm for energy-based active control,''J. Acoust. Soc. Am. **96,** 300–306 (1994).

[6] A. J. Bullmore, P. A. Nelson, A. R. D. Curtis, and S. J. Elliott, ''The active minimization of harmonic enclosed sound fields, Part II: A computer simulation,'' J. Sound Vib. **117**, 15–33 (1987).

[7] S. Laugesen and S. J. Elliott, ''Multichannel active control of random noise in a small reverberant room,'' IEEE Trans. Speech Audio Process. **1**, 241–249 (1993).

[8] P. A. Nelson, and S. J. Elliott, *Active Control of Sound* (Academic, New York, 1992), Chap. 10.

[9] S. J. Elliott, I. M. Stothers, and P. A. Nelson, ''A multiple error LMS algorithm and its application to the active control of sound and vibration,'' IEEE Trans. Acoust. Speech Signal Process. **35** 1423–1434 (1987).

[10] J. C. Burgess, ''Active adaptive sound control in a duct: a computer simulation,''J. Acoust. Soc. Am. **70,** 715–726 (1981).

[11] B. Widrow and S. D. Stearns, *Adaptive Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1985), Chap. 11.

[12] L. J. Eriksson, M. C. Allie, and R. A. Greiner, ''The selection and application of an IIR adaptive filter for use in active sound attenuation,'' IEEE Trans. Acoust. Speech Signal Process. **ASSP-35**, 433–437 (1987).

[13] T. Hodges, P. A. Nelson, and S. J. Elliot, ''The design of a precision digital integrator for use in an active vibration control system,'' Mech. Syst. Sign. Process. **4**, 345–353 (1990).

Interrupted traffic noise

Piotr Kokowski
Institute of Acoustics, A. Mickiewicz University (UAM), 70-679 Poznan, Matejki 48, Poland

Rufin Makarewicz[a)]
Kyushu Institute of Design, 9-1 Shiobaru 4-chome, Minami-ku, Fukuoka 815, Japan

(Received 4 December 1995; revised 3 June 1996; accepted 26 August 1996)

A one-parameter model (ξ) for stop-and-go noise is developed. The physics of the phenomenon is based on the idea of the *linear density of radiated energy*. The methodology of ξ determination has been worked out and illustrated. The noise interaction with the surface of the ground is considered to be the only wave phenomenon that disturbs geometrical spreading. The model makes possible the calculation of the correction term in the time-average sound level, L_{AT}, as an adjustment to the constant speed model of traffic noise. © *1997 Acoustical Society of America.* [S0001-4966(97)01501-4]

PACS numbers: 43.50.Lj, 43.50.Vt [GAD]

INTRODUCTION

Noise from a restricted flow of the road traffic stream has been studied by many authors.[1–7] The most complete summary of the previous results can be found in the report by Bowlby *et al.*[3] Our approach is based on the new concept of radiation energy density. Like the others, we make use of the time-average sound level, L_{AT}, and calculate an adjustment to the noise coming from the constant speed traffic (Sec. III). The ground effect is incorporated using an expression with one parameter that can be determined by noise measurement (Sec. IV). Calculating the total radiated energy we answer the question, whether the point of stop increases noise or not (Secs. VII, VIII).

The model of noise generation by interrupted traffic flow has been derived under the following assumptions:

(a) All halting vehicles that belong to the same category are assumed to stop in the same point and operate in the same mode.

(b) Noise emission during deceleration and acceleration is described in terms of the radiation energy density as a linear function of vehicles' position. The mode of acceleration is quantified by the measurable parameter, ξ.

Preliminary measurements of the interrupted traffic noise indicate that the present method of noise prediction can be used at a distance less than 200 m from the point of stop.

I. SOUND EXPOSURE LEVEL

When traffic noise consists of similar noise events, which are produced by identical operation of vehicles belonging to the same category (e.g., automobiles), then

$$L_{AT}=L_{AE}+10\ \log\{N\cdot t_0/T\}, \tag{1}$$

where $t_0=1$ s, and N expresses the number of noise events during the time period T. The sound exposure level, L_{AE}, characterizes the single noise event,

$$L_{AE}=10\ \log\{E/p_0^2 t_0\}, \tag{2}$$

where $p_0=2\times10^{-5}$ Pa. The sound exposure, E, can be calculated by the integration of the A-weighted squared sound pressure,

$$E=\int_{-\infty}^{\infty} p_A^2(t)\,dt. \tag{3}$$

In this study we consider a one-way road, which is composed of two cruise segments, where a vehicle is moving with constant speed V_0 (Fig. 1). The deceleration and acceleration segments are joined by the point of stop, i.e., starting point, $x=0$. All halting vehicles are assumed to operate in the same way. From $x=-\infty$ to $x=-l_1$ they move with the cruise speed, V_0. Then, at $x=-l_1$ the deceleration begins. Next, at $x=0$ they start the acceleration, which finally, at the point $x=l_2$, converts again into the cruise speed. If a fraction of vehicles, $k\cdot N$ ($0\leqslant k\leqslant 1$), cruise through the site with constant speed, V_0 and the rest of vehicles, $(1-k)\cdot N$, are stopped, then the time-average sound level, L_{AT}, of the interrupted traffic is given by

$$L_{AT}=\widetilde{L}_{AT}+\Delta L_{AT}, \tag{4}$$

where

$$\widetilde{L}_{AT}=\widetilde{L}_{AE}+10\ \log\{N\cdot t_0/T\} \tag{5}$$

expresses the cruise time-average sound level for constant speed traffic, i.e., as if there is no stopped traffic at all (see Sec. III). The cruise sound exposure level is given by

$$\widetilde{L}_{AE}=10\ \log\{\widetilde{E}/p_0^2 t_0\}, \tag{6}$$

and the correction term (see Sec. VI),

$$\Delta L_{AT}=10\ \log\{k+(1-k)(E/\widetilde{E})\}, \tag{7}$$

tells us how much the time-average sound level of the constant speed traffic, $\widetilde{L}_{AT}$, is affected by the stop-and-go motion of $(1-k)\cdot N$ vehicles. Note, that for $k=1$ (green light all the time) we get $\Delta L_{AT}=0$. The sound exposures, E and $\widetilde{E}$ express the noise contribution of a single vehicle: E for the stop-and-go mode of motion and $\widetilde{E}$ for the cruising conditions. Far away from the stop point we have $E_A\rightarrow\widetilde{E}_A$ (see

[a)]On leave from the Institute of Acoustics, UAM, Poznañ, Poland.

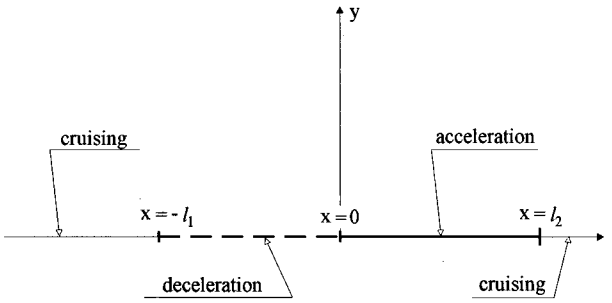


FIG. 1. A road consisting of the deceleration segment $(-l_1,0)$, acceleration segment $(0,l_2)$, and two cruise segments, $(-\infty,-l_2)$, (l_1,∞).

Sec. VII) and Eq. (7) yields $\Delta L_{AT}=0$. To predict the value of ΔL_{AT}, both sound exposures, E and $\widetilde{E}$, have to be known. However, according to definition (3), the time history of the A-weighted squared sound pressure, $p_A^2(t)$, is needed to compute E.

II. GROUND EFFECT

Above a plain surface of the ground, without buildings or any other intervening structures, the geometrical spreading of noise is affected by

(i) ground effect,
(ii) scattering due to atmospheric turbulence,
(iii) air absorption, and
(iv) refraction.

Close to the source, within 200 m, the ground effect is the most important factor. When the road is horizontal and the receivers height, H_0, is small compared to the perpendicular distance D (Fig. 2), the A-weighted squared sound pressure of the nondirectional point source is

$$p_A^2 \approx \frac{P_A \cdot \rho c}{4\pi d^2} G_A(d), \tag{8}$$

where P_A denotes the A-weighted power, ρc expresses the characteristic impedance of air, and

$$d=[(x-x_0)^2+D^2]^{1/2}, \tag{9}$$

is the horizontal distance between the source, $S(x,0)$, and receiver $O(x_0,D)$ (Fig. 3). Making use of the definition of the A-weighted sound-pressure level, we obtain the attenuation in excess of geometrical spreading [Eq. (8)],

$$\Delta L_A = 10 \log G_A(d). \tag{10}$$

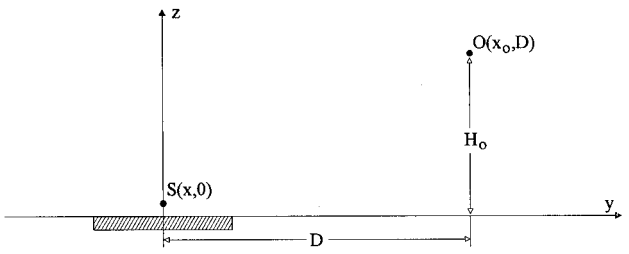


FIG. 2. A horizontal road at the perpendicular distance D from the receiver $O(x_0,D)$.

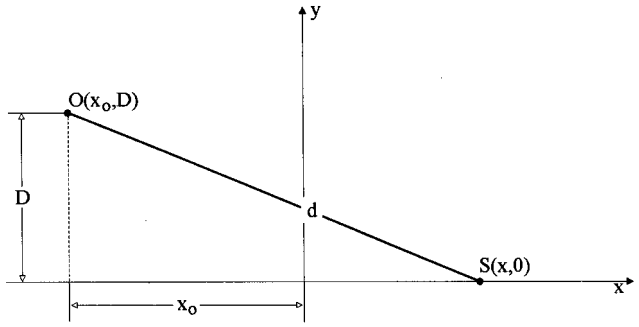


FIG. 3. Horizontal distance, d, between the vehicle S and the receiver O.

Here, ΔL_A can be seen as the measure of the ground influence on noise produced by the point source. The exact solution to that problem is known as the Weyl–Van der Pol formula.[8] Its far field approximation yields

$$\Delta L_A \rightarrow -20 \log\{d/d_0\}, \quad \text{for } d\rightarrow\infty, \tag{11}$$

where $d_0=1$ m. Unfortunately, the near-field approximation of the Weyl–Van der Pol solution $(d\rightarrow 0)$ is useless in the case under consideration, because the vehicle on a road cannot be considered as an infinitely small point source. In field measurements of ΔL_A close to the vehicle performed by Yamamoto and Yamashita[9] it was found that

$$\Delta L_A \rightarrow 3 \text{ dB}, \quad \text{for } d\rightarrow 0. \tag{12}$$

Both ''boundary conditions'' [Eqs. (11) and (12)] are met when the excess attenuation is approximated by

$$\Delta L_A = 3-10 \log\{1+d^2/d_g^2\}, \tag{13}$$

where d_g denotes the *critical distance*. Figure 4 shows ΔL_A for the point source as a function of the horizontal distance, d, for different values of the critical distance, d_g. Note, that with $d_g=1000$ m, instead of attenuation we get the noise increase: $\Delta L_A \approx 3$dB. The increase results from almost perfect reflection from the ''hard ground'' such as asphalt or concrete pavement. For $d_g=140$ m and $d_g=20$ m we obtain 3 and 6 dB decrease of excess attenuation per doubling of the distance, respectively. (Considering a road as the line source, we will obtain different characteristics of noise attenuation—see Sec. III). In this paper the critical distance, d_g, is estimated by the matching of the measured

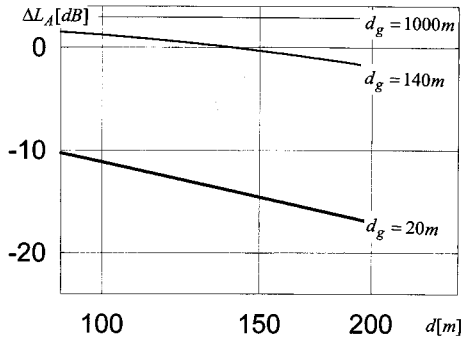


FIG. 4. Excess attenuation ΔL_A [Eq. (13)] for the critical distances $d_g=20$,140, and 1000 m, with 6, 3, and 0 dB decrease per doubling of the distance. (Noise is generated by a point source.)

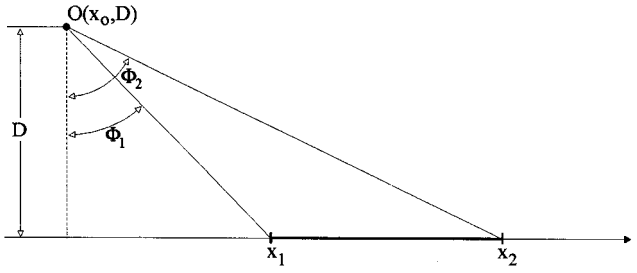


FIG. 5. Noise reaches the receiver, O, when the vehicle, S, moves between $\Phi_1(x_1)$ and $\Phi_2(x_2)$.

values with the parametrized theoretical predictions (see Sec. IV).

Using Eqs. (8) and (13) it then follows that the A-weighted squared sound pressure above the ground surface is

$$p_A^2=\frac{P_A\cdot\rho c}{4\pi d^2}\cdot\frac{2}{1+d^2/d_g^2}. \tag{14}$$

When a vehicle is cruising, the variations of $p_A^2(t)$ are caused only by the changes of the horizontal distance d [Eq. (9), Fig. 3]. In the case of stop-and-go motion, not only the distance, d, but the A-weighted power, P_A, vary with time. Thus the noise contribution coming from the road segment (x_1,x_2) can be calculated from the integral [Eqs. (3) and (14)],

$$E(x_1,x_2)=\frac{\rho c}{2\pi}\int_{x_1}^{x_2}\frac{P_A\cdot dx}{d^2\cdot V\cdot[1+(d/d_g)^2]}, \tag{15}$$

where V[m/s] is the vehicle's speed at the point x. Changing the variable of integration (Fig. 5),

$$x=x_0+D\cdot\tan\Phi, \tag{16}$$

we get [Eqs. (9) and (15)],

$$E(x_1,x_2)=\frac{\rho c}{2\pi D}\int_{\Phi_1}^{\Phi_2}\frac{P_A\cdot d\Phi}{V\cdot[1+(D/d_g\cos\Phi)^2]}. \tag{17}$$

Now we consider two types of vehicle motion: constant speed and stop-and-go.

III. CONSTANT SPEED NOISE

Constant speed motion means that a vehicle is cruising with steady speed V_0 from $x=-\infty$ to $x=+\infty$, without stopping at $x=0$ (green light). Under such conditions the A-weighted power, $P_A(V_0)$, remains constant, so we can rewrite the integral (17) in the form

$$\widetilde{E}=E_o\cdot\widetilde{F}(-\pi/2,\pi/2). \tag{18}$$

Here

$$E_o=\frac{P_A(V_0)\cdot\rho c}{4V_0D}, \tag{19}$$

expresses the cruise sound exposure in open space (far above the ground surface), and the factor

$$\widetilde{F}(-\pi/2,\pi/2)=2\left[1-\frac{D}{\sqrt{D^2+d_g^2}}\right], \tag{20}$$

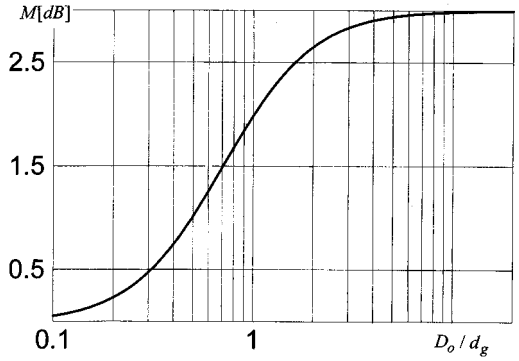


FIG. 6. Dependence of M [Eq. (26)] upon the relative distance, D_0/d_g.

determines the excess attenuation in terms of the sound exposure. Making use of the definition (2), we get the cruise sound exposure level,

$$\widetilde{L}_{AE}=L_{WA}(V_0)+10\log\left\{\frac{d_0^2}{2DV_0t_0}\left[1-\frac{D}{\sqrt{D^2+d_g^2}}\right]\right\}, \tag{21}$$

where $L_{WA}(V_0)$ is the cruise A-weighted power level. The results of many noise measurements show that

$$L_{WA}=a+m\log(V_0\cdot t_0/d_0). \tag{22}$$

For example, small vehicles and large vehicles on Japanese roads are described by $a_1=76.2, m_1=20$ and $a_2=82.6$, $m_2=20$.[10] Now we show how to use Eq. (21) for L_{WA} and d_g determination.

IV. L_{WA} AND d_g DETERMINATION

A. Method I

From Eq. (21) we obtain a theoretical prediction of the cruise sound exposure level for two distances, D_0 and $2D_0$:

$$L_{AE}(D_0)=L_{WA}+10\log\left\{\frac{d_0^2}{2D_0V_0\cdot t_0}\left[1-\frac{D_0}{\sqrt{D_0^2+d_g^2}}\right]\right\}, \tag{23}$$

$$L_{AE}(2D_0)=L_{WA}+10\log\left\{\frac{d_0^2}{4D_0V_0\cdot t_0}\left[1-\frac{2D_0}{\sqrt{4D_0^2+d_g^2}}\right]\right\}. \tag{24}$$

To estimate the critical distance, d_g, we perform simultaneous measurements of $\widetilde{L}_{AE}(D_0)$ and $\widetilde{L}_{AE}(2D_0)$, and rearrange Eqs. (23), (24) into the transcendental equation:

$$M(D_0/d_g)=\widetilde{L}_{AE}(D_0)-\widetilde{L}_{AE}(2D_0)-3\text{dB}. \tag{25}$$

The function

$$M\left(\frac{D_0}{d_g}\right)=10\log\left\{1-\frac{D_0/d_g}{\sqrt{1+(D_0/d_g)^2}}\right\}-10\log\left\{1-\frac{2D_0/d_g}{\sqrt{1+4(D_0/d_g)^2}}\right\} \tag{26}$$

is plotted in Fig. 6. Field measurements provide the right-hand side (rhs) of Eq. (25). Thus numerical value of the ratio

D_0/d_g, can be obtained from Fig. 6 by entering the abscissa with appropriate value of rhs. Finally, for a given distance D_0 between the microphone and the line of motion, one can calculate the critical distance, d_g. If the value d_g is known, L_{WA} can be found either from Eq. (23) or Eq. (24).

B. Experiment no. 1

Two simultaneous measurements of $\widetilde{L}_{AE}$, at two distances $D_0=7.5$ m and $2D_0=15$ m, have been carried out. The microphones were at heights $H_0=1.2$ m above the very soft ground and 43 automobiles (typical of European roads) have taken part in the experiment. The cruise of each vehicle was repeated five times. The vehicles were moving with steady speed $V_0\approx 15.3$ m/s. The average values of $\widetilde{L}_{AE}(D_0)$ and $\widetilde{L}_{AE}(2D_0)$ for 215 measurements are

$$\widetilde{L}_{AE}(7.5)=75.4 \text{ dB}, \quad \widetilde{L}_{AE}(15)=71.4 \text{ dB},$$

so rhs of Eq. (25) equals 1.0 dB. From Fig. 6 it is seen that it corresponds to the ratio $D_0/d_g=0.5$. Because the first microphone was at the distance $D_0=7.5$ m from the line of vehicle's motion, therefore $d_g=15$ m.

Substitution of $L_{AE}(7.5)=75.4$ dB, $D_0=7.5$ m, $V_0=15.3$ m/s, and $d_g=15$ m into Eq. (23) yields the cruise A-weighted power level: $L_{WA}=101.6$ dB.

C. Method II

Now we show how to determine L_{WA} and d_g from measurements of the cruise time-average sound level, $\widetilde{L}_{AT}$, at two distances, D_0 and $2D_0$. Assuming D expresses the distance from the center line of the road, we adopt Eqs. (1), (2), (18), (19), (20) for the ith category of vehicles,

$$\widetilde{L}_{AT}^{(i)}=\widetilde{L}_{AE}^{(i)}+10\ \log\{N_i\cdot t_0/T\}, \tag{27}$$

where

$$\widetilde{L}_{AE}^{(i)}=10\ \log\left\{\frac{P_A^{(i)}\cdot\rho c}{2DV_0t_0\cdot p_0^2}\left[1-\frac{D}{\sqrt{D^2+d_g^2}}\right]\right\}. \tag{28}$$

The logarithmic combination of $\widetilde{L}_{AT}^{(i)}$

$$\widetilde{L}_{AT}=10\ \log\left\{\sum\ 10^{\widetilde{L}_{AT}^{(i)}/10}\right\}, \tag{29}$$

gives the cruise time-average sound level of the noise coming from all vehicles, which pass by the site during the time period T. For two categories of vehicles we arrive at

$$\widetilde{L}_{AT}=10\ \log\left\{N_1\frac{P_A^{(1)}}{P_0}+N_2\frac{P_A^{(2)}}{P_0}\right\}+10\ \log\left\{\frac{d_0^2}{2DV_0T}\left[1-\frac{D_0}{\sqrt{D^2+d_g^2}}\right]\right\}, \tag{30}$$

where $P_0=10^{-12}$ W, and $d_0=1$ m. Making use of the simultaneous measurements, $\widetilde{L}_{AT}(D_0)$ and $\widetilde{L}_{AT}(2D_0)$, we obtain again Eq. (25). Its solution (see experiment no. 1) yields the critical distance d_g. The question of how to find $P_A^{(1)}$ and $P_A^{(2)}$, and finally, the A-weighted power levels, $L_{WA}^{(1)}$ and $L_{WA}^{(2)}$ arises. To solve this problem, an additional measurement is required of the cruise time-average sound level, $\widetilde{L}'_{AT}$ for flows N_1' and N_2' for the two categories of vehicle. Suppose that the results of measurements, $\widetilde{L}_{AT}(D_0)$ and $\widetilde{L}'_{AT}(D_0)$, are already known. In view of Eq. (30) we obtain the system of two linear equations with two unknowns, $P_A^{(1)}$ and $P_A^{(2)}$

$$\begin{aligned} N_1P_A^{(1)}+N_2P_A^{(2)}&=P_0\cdot 10^{[\widetilde{L}_{AT}-\delta L_{AT}]/10},\\ N_1'P_A^{(1)}+N_2'P_A^{(2)}&=P_0\cdot 10^{[\widetilde{L}'_{AT}-\delta L_{AT}]/10}, \end{aligned} \tag{31}$$

where the numerical value of

$$\delta L_{AT}=10\ \log\left\{\frac{d_0}{2D_0V_0T}\left[1-\frac{D_0}{\sqrt{D_0^2+d_g^2}}\right]\right\} \tag{32}$$

is given. For example, measuring the 1-h average sound level ($T=3600$ s) of noise generated by constant speed traffic ($V_0=15.3$ m/s with a soft ground ($d_g=15$ m), for $D_0=7.5$ m we obtain: $\delta L_{AT}=-62$ dB.

Now, assuming that the numbers of passing vehicles during the time period T are available, (N_1, N_2, N_1', N_2'), as well as the results of measurements ($\widetilde{L}_{AT}$, $\widetilde{L}'_{AT}$), Eq. (31) yield the A-weighted powers, $P_A^{(1)}$ $P_A^{(2)}$. Finally, the A-weighted power levels, $L_{WA}^{(1)}$ and $L_{WA}^{(2)}$ can be calculated. To decrease the error of L_{WA} determination, the number of passing vehicles, N, must be equal a few hundreds or more.

We will now make some comments on the ground attenuation. Field measurements indicate that on a ''hard ground'' (asphalt, concrete, etc.), the time-average level of traffic noise, $\widetilde{L}_{AT}$, decreases at a rate of 3 dB per doubling of the distance from the road center line.[11] To get such decrease within the range, $100<D<200$ m, we have to set a very large critical distance, e.g., $d_g\approx 1000$ m [Eq. (30), Fig. 6]. To obtain 4.5 dB decrease per doubling of the distance (within the same distance range) which characterizes a ''soft ground,'' we have to use $d_g=140$ m. (Such a ground is suitable for the growth of any vegetation, e.g., grass.) The critical distance $d_g=15$ m may be ascribed to a ''very soft ground,'' which yields 6 dB decrease of the time-average sound level per doubling of the distance.

We now return to Eq. (4), which plays a key role in this paper. At this point we have found a method for calculating the cruise time-average sound level, $\widetilde{L}_{AT}$ [Eqs. (21), (27), and (29)]. In the following section we will show how to calculate the correction term, ΔL_{AT} [Eq. (7)].

V. STOP AND GO NOISE

A. Cruise

To determine the correction ΔL_{AT} [Eq. (4)] caused by the interruption of continuous flow of vehicles, we start with the sound exposure, E, for the stop-and-go mode of motion. According to Fig.1, a vehicle moves with the cruise speed, V_0, along two road segments: $(-\infty,-l_1)$ and (l_2,∞). At $x=-l_1$ the deceleration begins. After the halt at $x=0$, the vehicle accelerates and achieves again the cruise velocity

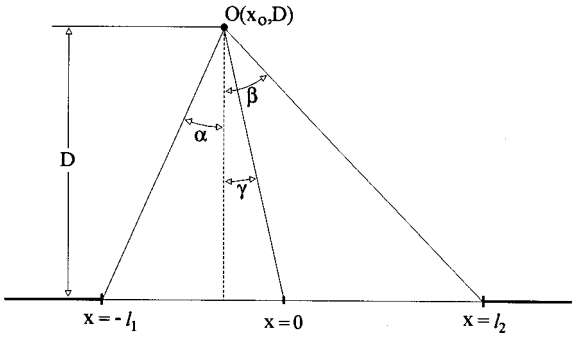


FIG. 7. Angles $(-\alpha,\gamma)$ and (γ,β) determine the location of the deceleration and acceleration segments.

V_0 at the point $x=l_2$. Thus the sound exposure of the stop-and-go motion, calculated for the receiver $O(x_0,D)$, can be broken down into five parts

$$E(x_0,D)=E(-\infty,-l_1)+E(-l_1,0)+E(0,0)+E(0,l_2)+E(l_2,\infty). \tag{33}$$

We assume the idle noise, which is produced during the halt at $x=0$, can be neglected,

$$E(0,0)\ll E(-\infty,-l_1)+E(-l_1,0)+E(0,l_2)+E(l_2,\infty). \tag{34}$$

Introducing the angles as in Fig. 7 we obtain the accounts of noise coming from both cruise segments of the road [Eq. (17)],

$$\begin{aligned} E(-\infty,-l_1)&=E_o\cdot F(-\pi/2,-\alpha),\\ E(l_2,\infty)&=E_o\cdot F(\beta,\pi/2), \end{aligned} \tag{35}$$

where E_o is the cruise sound exposure for an open space [Eq. (19)], and

$$\begin{aligned} F(-\pi/2,-\alpha)&=1-\frac{2}{\pi}\arctan\frac{l_1+x_0}{D}-\frac{D}{\sqrt{D^2+d_g^2}}\\ &\quad\times\left[1-\frac{2}{\pi}\arctan\frac{l_1+x_0}{\sqrt{D^2+d_g^2}}\right],\\ F(\beta,\pi/2)&=1-\frac{2}{\pi}\arctan\frac{l_2-x_0}{D}-\frac{D}{\sqrt{D^2+d_g^2}}\\ &\quad\times\left[1-\frac{2}{\pi}\arctan\frac{l_2-x_0}{\sqrt{D^2+d_g^2}}\right]. \end{aligned} \tag{36}$$

We are now prepared to discuss the process of noise generation during the deceleration and acceleration.

B. Deceleration

The model of the constant speed noise is relatively simple, because the A-weighted power, P_A, and the cruise velocity, V_0, remain constant. Therefore the ratio [Eqs. (18), (19)],

$$S_0=\frac{P_A(V_0)}{V_0} \tag{37}$$

remains unchanged. Note, that the unit is in joules (unit of energy) per meter (unit of length). However, the length cannot be attributed to the point source, which has no geometrical dimensions at all. Thus S_0 describes the road, i.e., the straight line which radiates sound energy, while a point source moves along with the constant velocity V_0. The unit [joule/meter] justifies the name of S_0: *density of radiated energy*. Now we can say that the cruise mode emission of noise is characterized by the uniform density of radiated energy, which is independent of the vehicle's position, x. During deceleration, both the A-weighted power and the velocity change along the road segment $(-l_1,0)$. Therefore we have to introduce S as a function of the vehicle's location, x,

$$S_1(x)=\frac{P_A(x)}{V(x)}. \tag{38}$$

The distribution of the radiated energy need not be uniform. What do we know about the function $S_1(x)$? The beginning of deceleration, $x=-l_1$, belongs to the cruise segment $(-\infty,-l_1)$, therefore [Eqs. (37), (38)],

$$S_1(-l_1)=S_0. \tag{39}$$

Later, both speed and the A-weighted power tends to zero: $V(x)\to 0$ and $P_A(x)\to 0$. This effect is caused by two mechanisms: The decrease of propulsion noise due to reduced engine speed and a decrease in tire noise due to a reduction in the vehicle's speed. Although, the A-weighted power is not exactly equal to zero (idle) at the point of stop, $x=0$, we assume,

$$S_1(0)\approx 0. \tag{40}$$

Because we do not know anything about the function $P_A(x)$, we cannot directly infer the explicit form of $S_1(x)$ between the points $x=-l_1$ and $x=0$. Therefore, a simple approximation suffices: let $S_1(x)$ decrease linearly from S_0 to 0 [Eqs. (39), (40)],

$$S_1(x)=-S_0\cdot x/l_1,\quad -l_1<x<0. \tag{41}$$

In due course, the consequences of this assumption will appear (Figs. 12–18), so we will be able to assess its correctness.

By applying Eqs. (16), (17), (41) the sound exposure of deceleration can be computed from,

$$E(-l_1,0)=E_o\cdot F(-\alpha,\gamma), \tag{42}$$

where E_o is defined by Eq. (19) and the integral (Fig. 7)

$$F(-\alpha,\gamma)=-\frac{2}{\pi l_1}\int_{-\alpha}^{\gamma}\frac{x_0+D\cdot\tan\Phi}{1+(D/d_g\cos\Phi)^2}d\Phi, \tag{43}$$

represents the result of deceleration and the ground effect. After some calculation one can show,

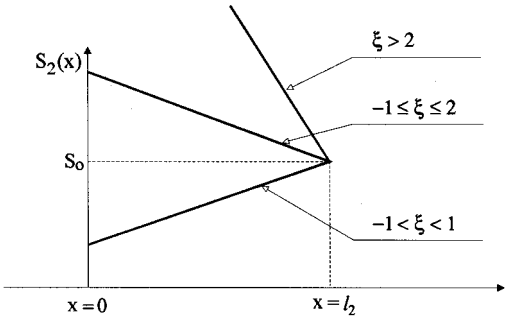


FIG. 8. The density of radiated energy, $S_2(x)$ [Eq. (47)], for careful $(-1\leqslant\xi\leqslant 1)$, normal $(1<\xi<2)$, and aggressive $(\xi\geqslant 2)$ acceleration.

$$F(-\alpha,\gamma)=\frac{2x_0}{\pi l_1}\left\{\arctan\frac{x_0}{D}-\arctan\frac{x_0+l_1}{D}+\frac{D}{\sqrt{D^2+d_g^2}}\times\left[\arctan\frac{l_1+x_0}{\sqrt{D^2+d_g^2}}-\arctan\frac{x_0}{\sqrt{D^2+d_g^2}}\right]\right\}-\frac{D}{\pi l_1}\ln\left\{\frac{(l_1-x_0)^2+D^2+d_g^2}{x_0^2+D^2+d_g^2}\cdot\frac{x_0^2+D^2}{(l_1+x_0)^2+D^2}\right\}. \tag{44}$$

C. Acceleration

During acceleration, which occurs along the road segment $(0,l_2)$, the density of radiated energy need not be uniform. Therefore we write,

$$S_2(x)=\frac{P_A(x)}{V(x)}. \tag{45}$$

At $x=l_2$ (Fig. 1), the acceleration mode reverts back to cruise mode, hence [Eqs. (37), (45)],

$$S_2(l_2)=S_0.$$

Usually, at the starting point, $x=0$, a rapid increase in engine speed makes the A-weighted power, $P_A(0)$, relatively large.[12] Due to the relatively small velocity $V(0)$, we can expect,

$$S_2(0)>S_0. \tag{46}$$

Similar to the case of deceleration, assume a linear change of S_2,

$$S_2(x)=S_0\cdot[1+\xi\cdot(1-x/l_2)],\quad 0<x<l_2. \tag{47}$$

The length l_2 is a function of ξ. Its explicit form is not important for further analysis of the problem under consideration.

Condition (46) is met when the parameter of acceleration ξ is positive. However, we cannot exclude the ''careful'' mode of acceleration. During a slow increase in engine speed we could expect $S_2(0)<S_0$. This agrees with Eq. (47) when the parameter of acceleration, ξ, is negative. Figure 8 shows the plot of $S_2(x)$ for careful $(-1\leqslant\xi\leqslant 1)$, normal $(1<\xi<2)$, and aggressive $(\xi\geqslant 2)$ acceleration. In most favorable conditions, $\xi=-1$, the density of radiated energy, $S_2(x)$, grows linearly from zero (at $x=0$) to S_0(at $x=l_2$).

Like the cruise A-weighted power level, $L_{WA}(V_0)$, and the critical distance, d_g, the parameter of acceleration, ξ, will be adjusted empirically.

Next we find a formula for the sound exposure of the acceleration part of the stop-and-go motion. By applying Eqs. (16), (17), (47) we arrive at

$$E_A(0,l_2)=E_o\cdot F(\gamma,\beta), \tag{48}$$

where E_0 [Eq. (19)] quantifies the cruise noise in open space and (Fig. 7),

$$F(\gamma,\beta)=\frac{2}{\pi}\int_\gamma^\beta\frac{1+\xi\cdot(l_2-x_0-D\cdot\tan\Phi)/l_2}{1+(D/d_g\cos\Phi)^2}d\Phi \tag{49}$$

takes account of the vehicle's acceleration and ground effect. After some calculation we get

$$F(\gamma,\beta)=F_1(\gamma,\beta)+\xi\cdot F_2(\gamma,\beta), \tag{50}$$

where

$$F_1=\frac{2}{\pi}\left\{\arctan\frac{l_2-x_0}{D}+\arctan\frac{x_0}{D}-\frac{D}{\sqrt{D^2+d_g^2}}\times\left[\arctan\frac{l_2-x_0}{\sqrt{D^2+d_g^2}}+\arctan\frac{x_0}{\sqrt{D^2+d_g^2}}\right]\right\}, \tag{51}$$

$$F_2=\left(1-\frac{x_0}{l_2}\right)F_1+\frac{D}{\pi l_2}\times\ln\left\{\frac{(l_2-x_0)^2+D^2+d_g^2}{x_0^2+D^2+d_g^2}\cdot\frac{x_0^2+D^2}{(l_2-x_0)^2+D^2}\right\}. \tag{52}$$

Note that on the extremely soft ground $(d_g\to 0)$ we get $F_1\to 0$ and $F_2\to 0$: no noise reaches the receiver. If the ground is very hard $(d_g\to\infty)$, then Eqs. (51), (52) simplify to the form,

$$F_1=\frac{2}{\pi}\left\{\arctan\frac{l_2-x_0}{D}+\arctan\frac{x_0}{D}\right\}, \tag{53}$$

$$F_2=\frac{2}{\pi}\left(1-\frac{x_0}{l_2}\right)\left\{\arctan\frac{l_2-x_0}{D}+\arctan\frac{x_0}{D}\right\}+\frac{D}{\pi l_2}\ln\left\{\frac{x_0^2+D^2}{(l_2-x_0)^2+D^2}\right\}. \tag{54}$$

VI. DETERMINATION OF ξ

A. Method I

We assume that $L_{AE}^*(0,l_2)$ is a measure of the accelerating noise for the receiver exactly opposite to the starting point, $x=x_0=0$ (Fig. 9). By virtue of Eqs. (2), (19), (48), and (50)–(52) we obtain the formula for L_{AE}^*,

$$L_{AE}^*(0,l_2)=L_{WA}(V_0)+10\ \log\left\{\frac{d_0^2}{4DV_0\cdot t_0}[F_1+\xi\cdot F_2]\right\}, \tag{55}$$

where

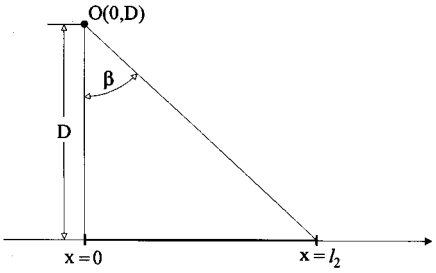


FIG. 9. Scenario of ξ measurements: the receiver $O(x_0=0)$ opposite the point of start $(x=0)$.

$$F_1(0,D,l_2)=\frac{2}{\pi}\left\{\arctan\frac{l_2}{D}-\frac{D}{\sqrt{D^2+d_g^2}}\arctan\frac{l_2}{\sqrt{D^2+d_g^2}}\right\}, \tag{56}$$

$$F_2(0,D,l_2)=F_1(0,D,l_2)+\frac{D}{\pi l_2}\ln\left\{\frac{l_2^2+D^2+d_g^2}{D^2+d_g^2}\cdot\frac{D^2}{l_2^2+D^2}\right\}. \tag{57}$$

Assuming that L_{AE}^* is the measurement output, we are able to determine the parameter of acceleration,

$$\xi=\frac{4DV_0\cdot t_0}{d_0^2\cdot F_2(0,D,l_2)}\cdot 10^{[L_{AE}^*-L_{WA}]/10}-\frac{F_1(0,D,l_2)}{F_2(0,D,l_2)}, \tag{58}$$

where $d_0=1$ m, $t_0=1$ s.

B. Experiment no. 2

At the same site as in experiment no. 1, the same 43 automobiles were used for ξ determination. Each of them started from the point $x=0$, opposite to the microphone, which was $D=7.5$ m away from the line of motion, and $H_0=1.2$ m above a very soft ground ($d_g=15$ m). The acceleration distance was $l_2=90$ m, i.e., at the position $x=90$ the vehicle achieved the cruise velocity $V_0=15.3$ m/s. Measurements of L_{AE}^* were repeated five times. The average of 215 measurements is: $L_{AE}^*=75.9$ dB.

Putting the values of $D=7.5$ m, $l_2=90$ m, and $d_g=15$ m into Eqs. (56), (57) we get, $F_1=0.55$ and $F_2=0.51$. Because the A-weighted power level (with $V_0=15.3$ m/s) equals $L_{WA}=101.6$ dB (see experiment no. 1), thus Eq. (58) yields the positive acceleration parameter: $\xi=1.3$.

C. Method II

Method I requires special conditions: The sound exposure level, L_{AE}^* has to be measured while the single vehicle accelerates. To be more realistic, imagine that we measure the time-average sound level, L_{AT}^*, of noise which comes from many vehicles of the same category. When $k\cdot N$ vehicles drive freely through the site, and $(1-k)\cdot N$ vehicles are stopped at the red light (during the time of measurement T), then [Eqs. (4)–(7)],

$$L_{AT}^*=10\ \log\left\{\frac{k\widetilde{E}}{p_0^2t_0}+\frac{(1-k)\cdot E}{p_0^2t_0}\right\}+10\ \log\left\{\frac{N\cdot t_0}{T}\right\}, \tag{59}$$

where $\widetilde{E}$ [Eq. (18)] expresses the cruise sound exposure, and [Eqs. (19), (33)–(35), (42), (48), (50)],

$$E=E_O\cdot[F(-\pi/2,-\alpha)+F(-\alpha,\gamma)+F_1(\gamma,\beta)+\xi F_2(\gamma,\beta)+F(\beta,\pi/2)], \tag{60}$$

is the sound exposure for the stop-and-go motion of a single vehicle. A brief calculation yields the parameter of acceleration

$$\xi=\frac{A-B}{F_2(\gamma,\beta)}, \tag{61}$$

where

$$A=\frac{4DV_0T}{(1-k)Nd_0^2}\,10^{[L_{AT}-L_{WA}(V_0)]/10},$$
$$B=\frac{k}{1-k}.\widetilde{F}(-\pi/2,\pi/2)+F(-\pi/2,-\alpha)+F_1(\gamma,\beta)+F(\beta,\pi/2). \tag{62}$$

The above expression can be used for $Nk\geqslant 1$, i.e., when at least one vehicle undergoes the stop-and-go motion during the time period T.

D. Example

Assume that along both cruise segments vehicles are moving with the average speed $V_0=15.3$ m/s. In such a case, the A-weighted power level equals $L_{WA}=101.6$ dB (see experiment no. 1). Now suppose L_{AT}^* is measured at the perpendicular distance from the road center, $D=50$ m, at the point $x_0=l_2/2$ m away from the starting point, $x=0$. Assuming $l_1=l_2=90$ m, and the grassy ground with $d_g=140$ m, we get [Eqs. (20), (36), (44), (51), (52)]:

$$\widetilde{F}(-\pi/2,\pi/2)=133,\quad F(-\pi/2,-\alpha)=0.047,$$
$$F(\beta,\pi/2)=0.26,\quad F(-\alpha,\gamma)=0.072,$$
$$F_1(\gamma,\beta)=0.81,\quad F_2(\gamma,\beta)=0.41.$$

Substitution of the above values into Eqs. (61), (62) yields,

$$\xi=4.7\cdot\frac{T}{(1-k)\cdot N\cdot t_0}\cdot 10^{[L_{AT}^*-70]/10}-3.2\cdot\frac{k}{1-k}-2.9, \tag{63}$$

where $t_0=1$ s. L_{AT}^* is the measured time-average sound level for $T[s]$, so the ratio $(N\cdot t_0/T)$ expresses traffic flow during the measurement, and $0\leqslant k<1$ denotes the fraction of vehicles which moves freely without stopping.

VII. DOES THE POINT OF STOP INCREASE NOISE?

A. Total radiated energy

The change in total energy output for a flow restriction may be an important parameter when comparing overall effects of different traffic control schemes. Figure 10 depicts an example of $S(x)$ variations along the road, while a vehicle decelerates from the cruise velocity, $V(-l_1)=V_0$, to the zero velocity, $V(0)=0$, and then accelerates back to the

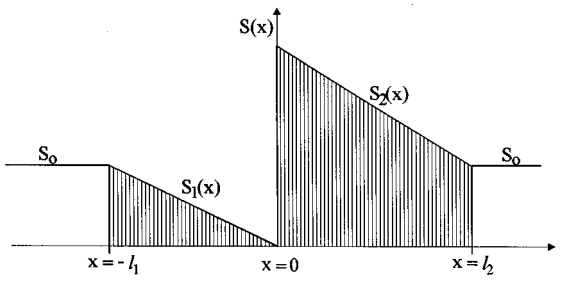


FIG. 10. The density of radiated energy for the stop-and-go mode of motion $S=S(x)$ [Eqs. (41), (47)].

cruise velocity, $V(l_2)=V_0$. It is rather a crude simplification, so the final conclusions of this and the next sections have to be seen as the first approximation of reality.

Under the definition Eq. (37), the density of radiated energy, $S(x)$, expresses the amount of energy per unit length of the road. Therefore the integral,

$$e=\int_{-l_1}^{l_2} S(x)dx \tag{64}$$

gives the total energy radiated by the road segment $(-l_1,l_2)$. During a red light, the acceleration follows the deceleration, thus the total energy equals [Eqs. (41), (47), (64)],

$$e=S_0\cdot\left[\frac{l_1}{2}+l_2+\xi\frac{l_2}{2}\right]. \tag{65}$$

The numerical value of e is represented by the shaded area in Fig. 10. When the vehicle cruises freely through the site (green light), then the total energy emitted from the same segment $(-l_1,l_2)$ is represented by the area of the rectangle (Fig. 11),

$$\tilde{e}=e_0\cdot[l_1+l_2]. \tag{66}$$

If the area of e (Fig. 10) exceeds the area of $\tilde{e}$ (Fig. 11), we could say that the vehicle's stop increases the noise, $e>\tilde{e}$. In view of Eqs. (65), (66) we obtain the following condition of the noise increase for a single passage:

$$\xi\frac{l_2}{l_1}>1. \tag{67}$$

In general, the distances l_1 and l_2 depend upon the cruise velocity V_0, i.e., the initial velocity at the onset of deceleration and the final velocity following acceleration. If the value V_0 is relatively low, then we can write $l_1\approx l_2$. Thus the decrease or increase of noise depends mainly upon the parameter of acceleration, ξ. For the normal $(1<\xi<2)$ and aggressive acceleration $(\xi\geqslant 2)$, the noise of stop-and-go motion dominates the cruise noise, so condition (67) is met. On the contrary, careful acceleration $(-1\leqslant\xi\leqslant 1)$ brings about the decrease of noise, $e<\tilde{e}$,

$$\xi\frac{l_2}{l_1}<1. \tag{68}$$

The inequalities (67), (68) are not the only criteria of the noise increase. More illustrative is the idea of the zone of influence, which we introduce in the next section.

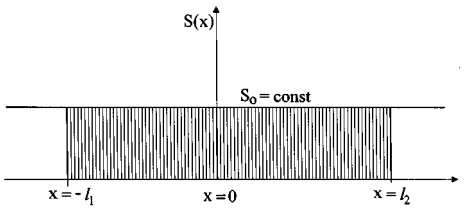


FIG. 11. The density of radiated energy for the cruise mode of motion: $S=S_0$.

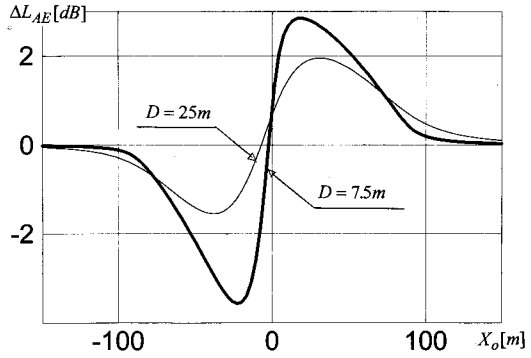


FIG. 12. Difference of the sound exposure levels, ΔL_{AE} [Eq. (72) with $d_g=140$ m, $l_1=l_2=90$ m, $\xi=1.5$] at the distances $D=7.5$ m, and $D=25$ m from the road.

B. Zone of influence

1. Sound exposure level, L_{AE}

Expression (7) contains the ratio of the sound exposures for the stop-and-go motion, E, and the constant speed motion, $\tilde{E}$. Introducing the corresponding sound exposure levels, we can write [Eq. (2)],

$$E/\tilde{E}=10^{\Delta L_{AE}/10}, \tag{69}$$

where the difference

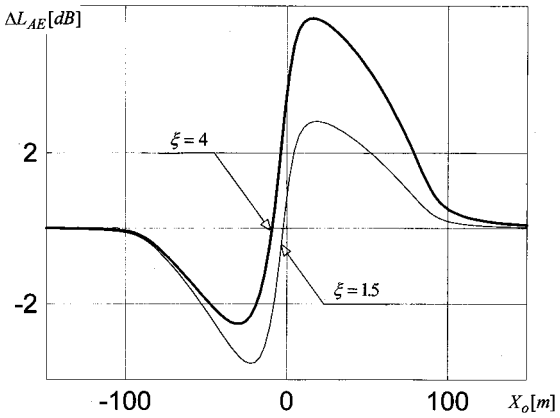


FIG. 13. Difference of the sound exposure levels, ΔL_{AE} [Eq. (72) with $d_g=140$ m, $l_1=l_2=90$ m, $D=7.5$ m] for normal, $\xi=1.5$, and aggressive, $\xi=4.0$, acceleration.

$$\Delta L_{AE}(x_0,D)=L_{AE}(x_0,D)-\widetilde{L}_{AE}(D) \tag{70}$$

will be examined later. The expression for the cruise sound exposure level, $\widetilde{L}_{AE}$, has been already derived [Eq. (21)]. The sound exposure level of noise which comes from the stop-and-go motion can be calculated by the following formula [Eqs. (2), (19), (50)–(52), (60)],

$$L_{AE}=L_{WA}(V_0)+10\ \log\left\{\frac{d_0^2}{4DV_0\cdot t_0}[F(\pi/2,-\alpha)+F(-\alpha,\gamma)+F(\gamma,\beta)+F(\beta,\pi/2)]\right\}. \tag{71}$$

Combination of Eqs. (21), (70), (71) yields the difference between the sound exposure levels,

$$\Delta L_{AE}(x_0,D)=10\ \log\left\{\frac{F(-\pi/2,-\alpha)+F(-\alpha,\gamma)+F(\gamma,\beta)+F(\beta,\pi/2)}{\widetilde{F}(-\pi/2,\pi/2)}\right\}, \tag{72}$$

where the functions $F(\cdots)$ are defined by Eqs. (20), (36), (44), (50)–(52). The difference ΔL_{AE}, as the function of the longitudinal distance x_0 from the point of start ($x=0$), is plotted in Fig. 12. It will be noted that ΔL_{AE} is lessened when the perpendicular distance D is increased. Let's examine the bahavior of $\Delta L_{AE}(x_0)$ when the receiver is close to a road, $D=7.5$ m. In the proximity of the deceleration segment $(-l_1,0)$ the cruise noise, $\widetilde{L}_{AE}$, is higher than the deceleration noise: $\Delta L_{AE}<0$. Ten meters before the point of stop ($x_0=-10$) the difference ΔL_{AE} [Eq. (72)] achieves -3.5 dB. As expected, close to the acceleration segment $(0,l_2)$, the cruise noise, $\widetilde{L}_{AE}$, is less than the acceleration noise: $\Delta L_{AE}>0$. Its maximum appears 10 m after the starting point ($x_0=+10$).

The changes of ΔL_{AE} along the road are strongly correlated with the variations of the density of radiated energy, $S(x)$ (compare Figs. 10 and 12) and dependent upon the parameter of acceleration, ξ (Fig. 13).

Does the quality of the ground surface significantly affect the value of ΔL_{AE}? Figure 14 shows the answer: no. The variations of ΔL_{AE} for soft ground ($d_g=20$ m) and hard ground ($d_g=1000$ m) are almost identical. We could say, that the influence of the stop-and-go motion on the cruise traffic noise, ΔL_{AE}, is independent of the ground covering.

Let's agree to disregard the influence of the stop-and-go motion on the total noise, when the absolute value of ΔL_{AE} is less than 2 dB. Thus the points on the (x,y) plane, which meet condition

$$|\Delta L_{AE}(x_0,D)|=2\ \text{dB} \tag{73}$$

define the borders of the influence zone. Particularly, the solution of the equation,

$$\Delta L_{AE}(x_0,D)=-2\ \text{dB} \tag{74}$$

determines the deceleration zone, and the equation

$$\Delta L_{AE}(x_0,D)=2\ \text{dB} \tag{75}$$

determines the acceleration zone. In the former case, there is an improvement of noise condition: $\Delta L_{AE}<-2$ dB. In the latter case, there is deterioration of noise condition: $\Delta L_{AE}>2$ dB.

Figure 15 shows the location of both zones on a grassy ground ($d_g=140$ m) for normal acceleration ($\xi=1.5$), while the distances of deceleration and acceleration are equal to each other, $l_1=l_2=90$ m. Note that the area of the acceleration zone is only ''a little'' larger than the area of the deceleration zone. It is in accordance with the criterion given by Eq. (67): for $l_1=l_2$ one could say that $\xi=1.5$ is ''a little'' larger than 1.

The situation changes completely when the acceleration is more aggressive: $\xi=4$. Figure 16 shows that the area of

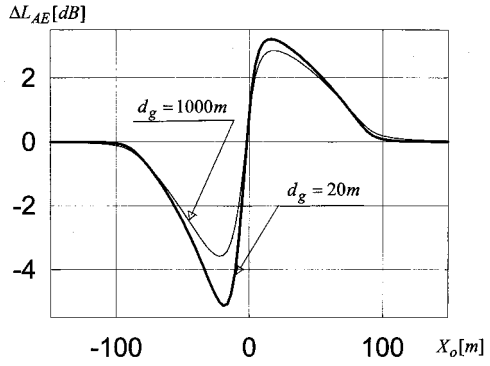


FIG. 14. Difference of the sound exposure levels, ΔL_{AE} [Eq. (72) with $d_g=140$ m, $l_1=l_2=90$ m, $D=7.5$ m, $\xi=1.5$) for very soft, $d_g=20$ m, and hard ground, $d_g=1000$ m.

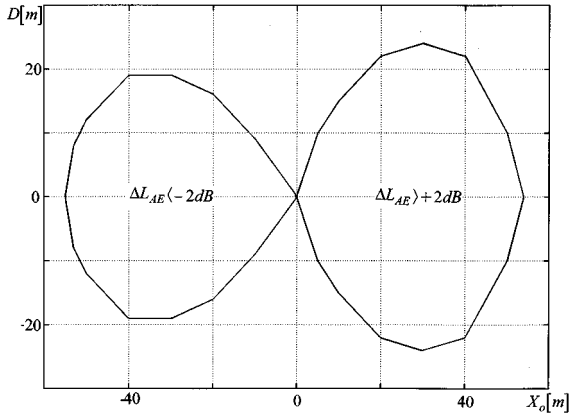


FIG. 15. The deceleration zone, $\Delta L_{AE}<-2$ dB, and the acceleration zone, $\Delta L_{AE}>2$ dB, for normal acceleration, $\xi=1.5$ [Eq. (72) with $d_g=140$ m, $l_1=l_2=90$ m].

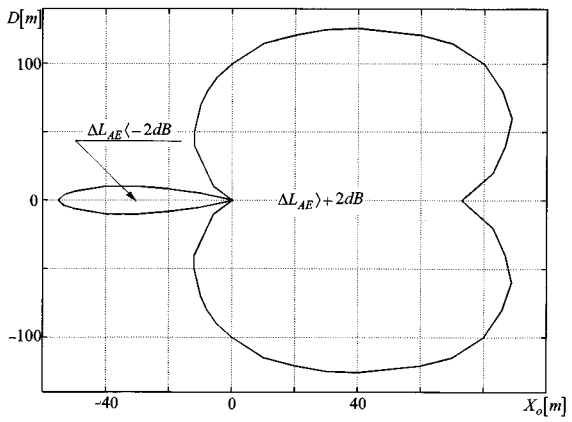


FIG. 16. The deceleration zone, $\Delta L_{AE}<-2$ dB, and the acceleration zone, $\Delta L_{AE}>2$ dB, for aggressive acceleration, $\xi=4.0$ [Eq. (72) with $d_g=140$ m, $l_1=l_2=90$ m].

the acceleration zone is about 15 times larger than the area of the deceleration zone. Without any doubts, for $\xi=4$ the stop-and-go motion increases noise.

So far the idea of the influence zones has been founded on the assumption that the noise is measured by the sound exposure level, L_{AE}. It is correct when the traffic flow is so low that a single passages occur. Otherwise, when two, three, or more vehicles approach the receiver at the same time, we have to use the time-average sound level, L_{AT}, instead of L_{AE}.

2. Time-average sound level, L_{AT}

Similar to the difference of the sound exposure levels, ΔL_{AE} [Eq. (70)], we introduce the difference between the time-average sound level for the interrupted traffic, L_{AT}, and the time-average sound level for the constant speed traffic, $\widetilde{L}_{AT}$,

$$\Delta L_{AT}=L_{AT}(x_0,D)-\widetilde{L}_{AT}(D). \tag{76}$$

ΔL_{AT}, as the function of the ratio $E/\widetilde{E}$ and the fraction of cruising vehicles k, is given by Eq. (7). Making use of Eq. (69) we obtain,

$$\Delta L_{AT}=10\ \log\{k+(1-k)\cdot 10^{\Delta L_{AE}/10}\}. \tag{77}$$

If during the time period T all vehicles halt ($k=0$) at the point of stop ($x=0$), then the difference of the time-average sound levels is the same as the difference of the sound exposure levels [Eq. (77)],

$$\Delta L_{AT}(x_0,D)\equiv\Delta L_{AE}(x_0,D). \tag{78}$$

This is true when the traffic flow is so low that no congestion takes place.

Under such condition, all conclusions regarding the effects of perpendicular distance, D (Fig. 12), mode of acceleration, ξ (Fig. 13), and ground covering, d_g (Fig. 14), are valid now, when we use the time-average sound level, L_{AT}. In particular, for all stopping vehicles ($k=0$), Figs. 15, 16 show the deceleration zone and the acceleration zone which borders are defined by $|\Delta L_{AT}|=2$ dB.

Now assuming that during the time interval $T, k\cdot N$ vehicles cruise through the site, the increase of k makes the

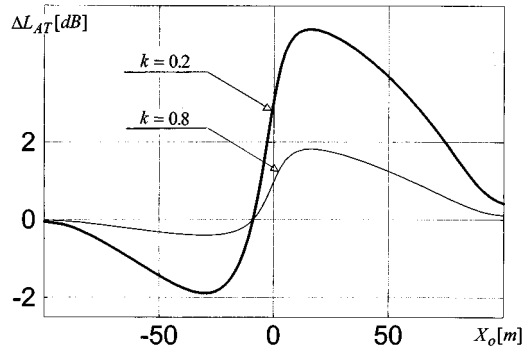


FIG. 17. Difference of the time-average sound levels, ΔL_{AT} [Eq. (77) with $d_g=140$ m, $l_1=l_2=90$ m, $\xi=4.0$, $D=7.5$ m] for a low ($k=0.2$) and high ($k=0.8$) content of cruising vehicles.

variations of $\Delta L_{AT}(x_0)$ more smooth. It is seen at Fig. 17: for a low number of cruising vehicles ($k=0.2$), the constant speed noise ($\widetilde{L}_{AT}$) is strongly affected by the vehicles stopping at $x=0$. The curve $\Delta L_{AT}(x_0)$ indicates a dip of about 2 dB ($x_0\approx-30$) and a maximum, which exceeds 4 dB ($x_0\approx 20$). For a larger number of cruising vehicles ($k=0.8$), the dip is no longer noticeable and the maximum falls down below 2 dB. The conclusion is that for $k>0.8$, the effect of traffic interruption can be neglected, and no corrections to the constant speed model (Sec. III) are needed.

Now let's find the zones of influence for any percentage of cruising vehicles ($0<k<1$). Making use of the difference of the time-average sound levels [Eq. (76)] we introduce the definition of the influence zone border:

$$|\Delta L_{AT}(x_0,D)|=2\text{ dB}. \tag{79}$$

Proceeding along the lines analogous to those above [see Eqs. (73)–(75)] we obtain the deceleration zone and the acceleration zone in terms of the time-average sound level. Figure 18 shows the borders of both zones, which were calculated for $k=0.2$ with $d_g=140$ m, $l_1=l_2=90$ m, and $\xi=4$. The same parameters have been used to determine the zones of influence shown in Fig. 16, where ΔL_{AT} may be replaced by ΔL_{AE}. However, for $k=0$ (no cruising vehicles) identity (78) holds, so Figs. 16 and 18 can be used for deter-

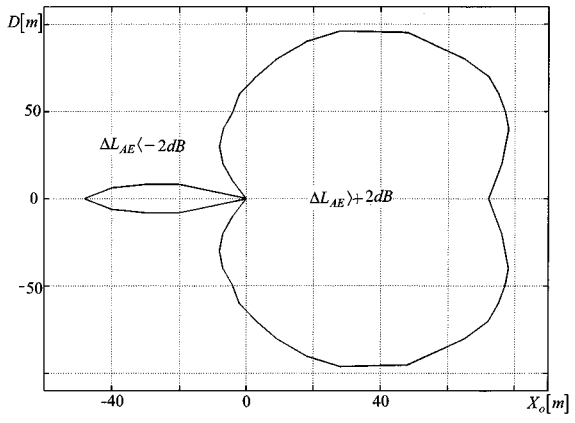


FIG. 18. The deceleration zone, $\Delta L_{AE}<-2$ dB, and the acceleration zone, $\Delta L_{AE}>2$ dB, for a low content of cruising vehicles, $k=0.2$ [Eqs. (72), (77) with $d_g=140$ m, $l_1=l_2=90$ m, $\xi=4.0$].

mining the dependency of the content of cruising vehicles (k) on the area of the influence zones. The increase of cruising vehicles decreases the area of the acceleration zone. The decrease of the deceleration zone area is less distinct.

When the area of the acceleration zone is larger than the area of the deceleration zone, the stop-and-go behavior of $(1-k)\cdot N$ vehicles deteriorates the noise condition. Using the concept of the total radiated energy, those vehicles produce energy [Eq. (65)],

$$e_n=(1-k)\cdot N\cdot S_0\cdot\left[\frac{l_1}{2}+l_2+\xi\frac{l_2}{2}\right]. \tag{80}$$

The contribution of cruising vehicles is [Eq. (66)],

$$\tilde{e}_n=k\cdot N\cdot S_0\cdot[l_1+l_2]. \tag{81}$$

Thus the traffic interruption increases the total noise energy when the following inequality is fulfilled:

$$e_n+\tilde{e}_n>N\cdot S_0\cdot[l_1+l_2]. \tag{82}$$

The right-hand side expresses the total radiated energy under the assumption that all vehicles are cruising. Surprisingly, solution to this inequality appears again to be expression (67). Thus using either the sound exposure level, L_{AE}, or the time-average sound level, L_{AT}, we arrive at the conclusion: the total increase of noise is related to the kinematics of the stop-and-go motion (l_1,l_2) and its dynamics (ξ). Remember that deceleration distance, l_1, is a function of the initial velocity, V_0, and acceleration distance, l_2, is related to the same ultimate velocity, V_0.

VIII. SUMMARY

The noise coming from constant speed traffic can be assesed by the cruise time average of sound level, $\tilde{L}_{AT}$ [Eq. (29)], where the contribution of the vehicles belonging to the ith category is [Eqs. (5), (21)],

$$L_{AT}^{(i)}=L_{WA}^{(i)}(V_0)+10\ \log\left\{\frac{N_i\cdot d_0^2}{2DV_0T}\left[1-\frac{D}{\sqrt{D^2+d_g^2}}\right]\right\},$$

$$D<200\text{ m}. \tag{83}$$

The A-weighted power level, L_{WA}, is determined by the cruise velocity, V_0 [Eq. (22)]. The quantity N expresses the number of vehicles, which pass the receiver during the time period T. The definition of the perpendicular distance from the road center is given in Fig. 2. Setting $d_g=20$, 140, and 1000 m, we get attenuation which is typical for the ''very soft ground'' with 6 dB per doubling of the distance, the ''soft ground'' with 4.5 dB per doubling of the distance, and the ''hard ground'' with geometrical divergence of 3 dB per doubling of the distance, while $D<200$ m.

When constant speed traffic is disturbed by a point of stop (Fig. 1), then the cruise time-average sound level, $\tilde{L}_{AT}$, has to be adjusted [Eqs. (4), (7)]. The correction term, ΔL_{AT} [Eq. (77)], contains the percentage of the cruising vehicles, k, and the difference of sound exposure levels, ΔL_{AE} [Eqs. (20), (36), (44), (50)–(52), (71)]. The latter accounts for the stop-and-go noise of a single vehicle.

The generation of that type of noise was discussed with the concept of the density of radiated energy, S [Eqs. (37), (38), (45)]. Because the scarcity of data regarding the process of noise generation during the stop-and-go motion, we were forced to make simplified assumptions: $S(x)$ varies linearly with the location of source, x, during the deceleration [Eq. (41), Fig. 10] and acceleration [Eq. (47), Fig. 10]. The slope of $S(x)$ during the acceleration, i.e., dynamics of acceleration, is quantified by the parameter of acceleration ξ. It appears in the formulae for ΔL_{AT} calculation (see above), and in the expressions which portray the energy changes which are caused by a traffic interruption [Eqs. (80)–(82)]. For any percentage of cruising vehicles, i.e., for any value of $0<k<1$, the point of stop increases the total radiated noise when [Eq. (67)],

$$\xi\frac{l_2(V_0)}{l_1(V_0)}>1. \tag{84}$$

Here the distances of deceleration, l_1, and acceleration, l_2, are functions of the cruise velocity, V_0. This remark regards all equations of this paper which contain l_1 and l_2.

To make the method of ΔL_{AT} prediction effective, the numerical value of ξ has to be known. Therefore the methodology of ξ determination has been worked out and illustrated (Sec. V). Our results indicate that the acceleration of automobiles within a town area (from $V=0$ to $V_0=55$ km/h) can be characterized by $\xi=1.3$ (experiment no. 2). One may expect larger values of ξ for trucks. This is the first task for additional measurements that can be easily fulfilled.

Measurements of L_{AT} of noise coming from interrupted automobile traffic have already been done. However, they are few in number and of partial nature. So far, general agreement between the measured and calculated results of L_{AT} have been obtained.

The problem of noise generation from congested flow situations, multistream roads and junction will be considered in the next papers.

ACKNOWLEDGMENTS

The authors are grateful to the anonymous reviewer for his valuable comments and to Miss Ewa Biskupska for technical assistance.

[1] B. Favre, ''Noise at the approach to traffic lights: results of the simulation programme,'' J. Sound Vib. **58**, 563–578 (1978).
[2] R. R. K. Jones, D. C. Hothersall, and R. J. Satler, ''Techniques for the investigation of road traffic noise in regions of restricted traffic by the use of digital computer simulation methods,'' J. Sound Vib. **75**, 307–322 (1981).
[3] W. Bowlby, R. L. Wayson, and R. E. Stammer, Jr., ''Predicting stop-and-go traffic noise levels,'' NCHRPR 311, Report No. 311, Washington, DC, 1989.
[4] L. Nijis, ''The increase and decrease of traffic noise levels at intersections measured with moving microphone,'' J. Sound Vib. **131**, 127–141 (1989).
[5] M. Hunt and S. Samuelson, ''Prediction of traffic noise at signalised inter-

sections: the australien experience,'' Inter Noise-92, Toronto, 1992, pp. 805–810.
[6] S. Radosz, M. Tracz, and J. Bohatkiewicz, ''Prediction of traffic noise at signalised using computer simulation method,'' Inter Noise-92, Toronto, 1992, pp. 811–816.
[7] S. Yamaguchi, Y. Kato, and S. Ishihara, ''A fundamental consideration on evaluating noise produced by road traffic controlled by traffic signals,'' Appl. Acoust. **42**, 55–73 (1994).
[8] K. Attenborough, ''Review of ground effects on outdoor sound propagation from continuous broadband noise,'' Appl. Acoust. **24**, 289–319 (1988).
[9] K. Yamamoto and M. Yamashita, ''A simple model for estimating excess attenuation of road traffic noise,'' J. Acoust. Soc. Jpn. (E) **8**, 13–22 (1987).
[10] H. Tachibana and M. Sasaki, ''ASJ prediction methods of road traffic noise,'' Inter Noise-94, Yokohama, 1994, pp. 283–288.
[11] W. Bowlby, ''Highway noise prediction and control,'' in *Acoustical Measurements and Noise Control*, edited by C. M. Harris (McGraw-Hill, New York, 1991), p. 48.12.
[12] Y. Watanabe, A study of sound power level radiated from an isolated starting truck,'' Inter Noise-94, Yokohama, 1994, pp. 419–422.

Simplified model of the ground effect

R. Makarewicz[a)]
Kyushu Institute of Design, 9-1 Shiobaru 4-chome, Minami-ku, Fukuoka 815, Japan

P. Kokowski
Institute of Acoustics, A. Mickiewicz University (UAM), 60-769 Poznan, Matejki 48, Poland

(Received 10 June 1996; accepted for publication 9 September 1996)

Noise interaction with the surface of the ground is considered to be the only wave phenomenon that disturbs geometrical spreading. A two-parameter model of the ground effect is developed and adjusted to account for industrial, road-traffic, and railroad noise. The model is founded on the Weyl–Van der Pol solution, which is treated as an exact theory of the ground effect. The adjustment of free parameters requires two simultaneous field measurements. © *1997 Acoustical Society of America.* [S0001-4966(97)01401-X]

PACS numbers: 43.50.Lj, 43.50.Vt [GAD]

INTRODUCTION

The geometrical spreading of noise above a plane ground surface, without buildings and any other interfering structures, is affected by air absorption, refraction, atmospheric turbulence, and ground effect. When studying problems close to the noise source, ground effect is the most important factor.

A number of computer programs are available for predicting industrial, road-traffic, and railroad noise. Although some of these programs generate precise prediction of the ground effect values, a tremendous expense in computing time is incurred. Subsequently, a need has arisen for methods that are practical for everyday use despite limited accuracy.[1–3]

A simple model of the ground effect is derived in Sec. I of this paper. The model appears simple because the estimates of free parameters are based on routine noise measurements. Additionally, the estimation procedure requires a relatively short computing time. The model presented here is valid for distances less than 150 m, since noise propagation is essentially independent of air absorption, refraction, and turbulence.[1]

Section II describes a procedure to predict the A-weighted sound-pressure level, L_A,[4] of industrial noise due to a stationary source. Section III A describes a prediction of the time-average sound level, L_{AT}, which is used to assess road-traffic noise. Noise generated by railroad trains is measured either by L_{AT}, or by the sound exposure level, L_{AE}. The procedure for L_{AE} prediction is developed in Sec. III B.

I. GROUND EFFECT

Under free-field conditions high above the ground surface, the squared A-weighted sound pressure may be written as

$$p_A^2=\frac{P_{\mathrm{A}}\cdot Q(\overline{n})\rho c}{r^2}, \tag{1}$$

where P_A is the A-frequency weighted sound power, ρc expresses the characteristic impedance of air, r denotes the source–receiver distance, and $Q(\overline{n})$ describes the noise radiation in the direction determined by the unit vector $\overline{n}$ (Fig. 1).

The law of energy conservation yields

$$\int\int_{4\pi} Q(\overline{n})d\Omega=1, \tag{2}$$

where $d\Omega$ represents the differential of a solid angle. Considering a nondirectional source, we find $Q=1/4\pi$. Thus close to a uniform and locally reacting surface (Fig. 2),

$$H_s\ll d, \quad H_0\ll d, \tag{3}$$

we can write,[5]

$$p_A^2=\frac{P_{\mathrm{A}}\rho c}{4\pi d^2}\ G(d,H_s,H_0,P_n,Z_n). \tag{4}$$

The ground factor G results from the Weyl–Van der Pol solution. The ground factor is influenced by the source–receiver location (d,H_s,H_0), the sound power (P_n), and the ground impedance in the nth frequency band (Z_n). The ground impedance is a continuous function of frequency f with four parameters,[6]

$$Z_n=Z(f_n,\sigma,n',s_p,\Omega). \tag{5}$$

To estimate G [Eq. (4)], numerical values of P_n and Z_n must first be determined. In some cases, such as fan noise, P_n can be established from anechoic or semianechoic data. Estimating G for a moving source (e.g., a moving truck) is not as simple.

A great deal of progress has been made in the determination of the impedance parameters, σ, n', s_p, Ω.[6] If these parameters are not available for a specific ground surface, we propose to replace the ''exact'' G from Eq. (4) by its approximation,

$$\tilde{G}=\frac{\beta}{1+\gamma\cot^2\Psi}, \quad \Psi>0, \tag{6}$$

where the grazing angle, Ψ, shown in Fig. 2, is defined by

[a)]On leave from the Institute of Acoustics (UAM), Poznan, Poland.

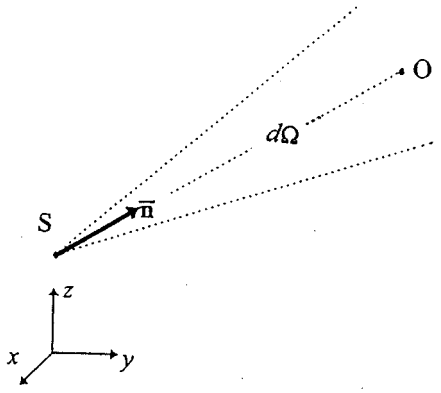


FIG. 1. Source (S) and receiver (O) high above the ground surface, i.e., the x,y plane.

$$\cot \Psi = d/(H_s + H_0). \tag{7}$$

As the ground coefficient γ approaches infinity above a very soft ground, i.e., $\gamma \to \infty$, the $\tilde{G}$ approaches zero. As the ground coefficient approaches zero above a very hard ground, i.e., $\gamma \to 0$, we obtain $\tilde{G} \to \beta$, for any angle $\Psi > 0$.

Results by Li *et al.*[7] indicate that the exact ground factor G is a function of the grazing angle, Ψ. Taking into account the asymptotic behavior of the exact ground factor,[5]

$$\lim_{d\to\infty} G = \lim_{d\to\infty} \tilde{G} \propto d^{-2}, \tag{8}$$

we propose the explicit form of $G(\Psi)$ [Eqs. (6) and (7)].

Close to the noise source, the approximation given by Eq. (6) yields

$$\lim_{d\to 0} \tilde{G} = \beta. \tag{9}$$

A sound-pressure level increase of 3 dB may be observed due to noise reflection from a hard ground. This increase corresponds to $\beta = 2$.

Using Eqs. (1), (6), and (7) for the source and receiver close to the ground [Eq. (3)], the approximated value of the squared A-weighted sound pressure can be calculated from

$$p_A^2 = \frac{\beta P_A \cdot Q(\Phi)\rho c}{d^2}\left[1 + \gamma \cdot \left(\frac{d}{H_s + H_0}\right)^2\right]^{-1}, \tag{10}$$

where the angle Φ remains in the x,y plane (Fig. 3).

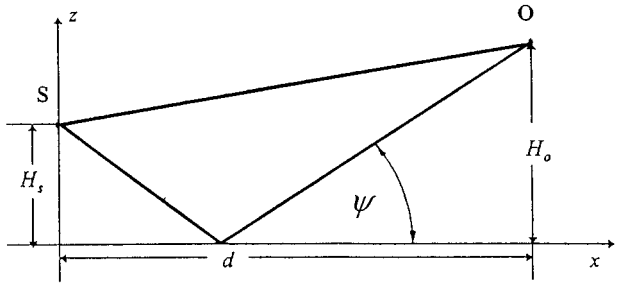


FIG. 2. Source (S) and receiver (O) close to the ground [Eqs. (3)], with grazing angle, Ψ [Eq. (7)].

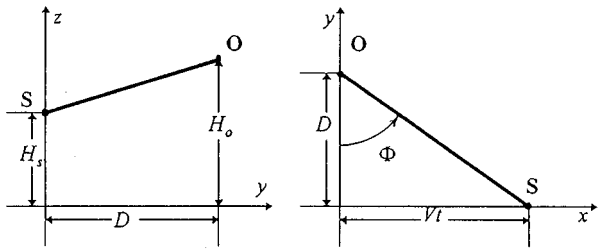


FIG. 3. Noise emitted by the source S(Vt,0,H_s) reaches the receiver O(0,D,H_0).

II. STATIONARY NOISE SOURCES

For a stationary and nondirectional source with $Q = 1/4\pi$, the A-weighted sound-pressure level is [Eqs. (3), (10)],

$$L_A = L_{WA} + 10 \log\left\{\frac{s_0}{4\pi d^2}\right\} + A(\sqrt{\gamma}\cdot d), \tag{11}$$

where $s_0 = 1\ \mathrm{m}^2$ and

$$L_{WA} = 10 \log\left\{\frac{\beta P_A}{P_0}\right\} \quad \text{with} \quad P_0 = 10^{-12} W, \tag{12}$$

denotes the effective sound power level, that is modified by the ground effect (βP_A). The ground attenuation is given by

$$A(\sqrt{\gamma}\cdot d) = -10 \log\left\{1 + \gamma \cdot \left(\frac{d}{H_s + H_0}\right)^2\right\}. \tag{13}$$

For very soft ($\gamma \to \infty$) and very hard ground ($\gamma \to 0$), the value of A approaches $-\infty$ and 0, respectively.

Equation (11) may now be considered a theoretical prediction with two adjustable parameters, L_{WA} and γ. If two measurements, [L_{A1} and L_{A2}], are performed at two different locations [(d_1,H_{01}) and (d_2,H_{02})], then the ground coefficient can be calculated from Eqs. (11) and (13) as follows:

$$\gamma = (m-1)/([d_2/(H_s + H_{02})]^2 - m[d_1/(H_s + H_{01})]^2), \tag{14}$$

where the quantity

$$m = \left(\frac{d_1}{d_2}\right)^2 \cdot 10^{(L_{A1} - L_{A2})/10} \tag{15}$$

is fully determined. By substituting the ground coefficient γ into Eq. (11), the effective sound power level, L_{WA}, is obtained.

Example 1

Noise was produced by a fan with an outlet opening at the height $H_s \approx 3$ m above the ground surface. Two microphones were installed at height $H_0 = 1$ m and at distances $d_1 = 50$ m and $d_2 = 100$ m, from the noise source. During a sunny day with no wind, ten measurements were taken and their averages were: $L_{A1} = 56$ dB and $L_{A2} = 49$ dB.

Using Eqs. (11), (14), and (15), the values $\gamma = 6 \times 10^{-4}$ and $L_{WA} = 101$ dB were obtained.

III. MOVING NOISE SOURCES

A. Road-traffic noise

The time-average sound level L_{AT} (Ref. 4) is widely used as a measure of road-traffic noise. For simplicity, a road with vehicles of one type (e.g., automobiles), is considered. In such a case, one can write

$$L_{AT}=L_{AE}+10\ \log\left\{\frac{Nt_0}{T}\right\}, \tag{16}$$

where $t_0=1$ s and the quotient N/T expresses the vehicle flow rate. The sound exposure level is defined by[4]

$$L_{AE}=10\ \log\left\{\frac{E}{p_0^2\ t_0}\right\}, \tag{17}$$

where $p_0=2\times10^{-5}$ Pa and the integral

$$E=\int_{-\infty}^{+\infty}p_A^2(t)dt, \tag{18}$$

defines the sound exposure.

Suppose the vehicle $S(x=Vt,y=0,z=H_s)$ is moving with steady speed, V, along a straight and horizontal road at a perpendicular distance D from the receiver $O(x=0,y=D,z=H_0)$ (see Fig. 3). Introducing the new variable, $\Phi=a\tan(Vt/d)$, the source–receiver distance can be approximated by, $d\approx D/\cos\Phi$, and the definition (18) takes the following form:

$$E=\frac{D}{V}\int_{-\pi/2}^{+\pi/2}\frac{p_A^2(\Phi)}{\cos^2\Phi}\ d\Phi. \tag{19}$$

A moving vehicle may be modeled by a nondirectional point source, so using Eqs. (10), (16), (17), and (19) we obtain

$$L_{AT}=L_{WA}+10\ \log\left\{\frac{Ns_0}{4DVT}\right\}+B(\sqrt{\gamma}\cdot D), \tag{20}$$

where L_{WA} denotes the effective sound power level [Eq. (12)] of a single vehicle that is moving with the steady speed, V. The ground attenuation is determined by

$$B(\sqrt{\gamma}\cdot D)=10\ \log\left\{1-\frac{\sqrt{\gamma}\cdot D}{\sqrt{\gamma\cdot D^2+(H_s+H_0)^2}}\right\}. \tag{21}$$

Note, that very soft ($\gamma\to\infty$) and very hard ground ($\gamma\to0$) are characterized by $B\to-\infty$ and $B\to0$, respectively.

Equation (20) may be used for traffic noise prediction when the width of the road is small as compared with the perpendicular distance D (Fig. 3).

To estimate the effective sound power level of a single vehicle, L_{WA}, and the ground coefficient, γ, two simultaneous measurements of L_{AT}, at perpendicular distances, D_* and $2D_*$, are needed. Equations (20) and (21) give the transcendental equation in respect to γ,

$$M_b(\xi)=L_{AT}(D_*)-L_{AT}(2D_*)-3. \tag{22}$$

The function

$$M_b=10\ \log\left\{\frac{\sqrt{1+\xi^2}-1}{\sqrt{1+\xi^2/4}-1}\cdot\frac{\sqrt{1+\xi^2/4}}{\sqrt{1+\xi^2}}\right\}, \tag{23}$$

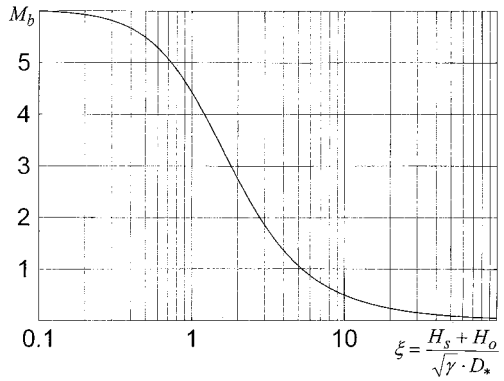


FIG. 4. Dependence of M_b upon ξ [Eq. (23)].

vs

$$\xi=\frac{H_S+H_0}{\sqrt{\gamma}\cdot D_*} \tag{24}$$

is shown in Fig. 4. Field measurements of $L_{AT}(D_*)$ and $L_{AT}(2D_*)$ provide the right-hand side (rhs) of Eq. (22). The numerical value of ξ can then be obtained from Fig. 4 by assigning the appropriate value of the rhs to the abscissa. For the given values of D_*, H_s (height of the vehicle), and H_0 (height of the microphone), the ground coefficient γ [Eq. (24)] is obtained. Finally, by substituting γ into Eq. (20), the effective sound power level, L_{WA}, may be calculated.

Example 2

Simultaneous measurements of the 1-average sound level, L_{1h}, during a clear day with no wind, were made at distances $D_*=25$ m and $2D_*=50$ m, both at the height $H_0=1.2$ m above grass. Traffic noise was produced by $N=523$ automobiles that passed the microphones during the 1-h period ($T=3600$ s) with the average speed $V=16$ m/s. The results of the measurements are as follows:

$$L_{1h}(25)=62\text{ dB}\quad\text{and}\quad L_{1h}(50)=58\text{ dB}.$$

Thus the right-hand side of Eq. (22) equals 1 dB, and Fig. 4 yields $\xi=5.2$. Since $H_s=0$ (height of the automobile), $H_0=1.2$ m, and $D_*=25$ m, the ground coefficient is found to be $\gamma=8\times10^{-5}$, and Eq. (20) yields the effective sound power level, $L_{WA}=103$ dB.

Road traffic normally consists of automobiles, medium trucks, and heavy trucks, so the total time-average sound level, L_{AT}, can be expressed as a logarithmic combination

$$L_{AT}=10\ \log\left\{\sum_{i=1}^{3}10^{L_{AT}^{(i)}}\right\}, \tag{25}$$

where $L_{AT}^{(1)}$, $L_{AT}^{(2)}$, and $L_{AT}^{(3)}$ correspond to the noise generated by automobiles, medium trucks, and heavy trucks, respectively. Vehicle types are distinguishable by the effective sound power level, $L_{WA}^{(i)}$, and the height, $H_s^{(i)}$, therefore Eq. (20) takes the form

$$L_{AT}^{(i)}=L_{WA}^{(i)}+10\ \log\left\{\frac{N_i\cdot s_0}{4DV_iT}\right\}$$
$$+10\ \log\left\{1-\frac{\sqrt{\gamma}\cdot D}{\sqrt{\gamma\cdot D^2+(H_s^{(i)}+H_0)^2}}\right\}, \tag{26}$$

where N_i/T and V_i characterize the traffic of each vehicle type.

The Federal Highway Administration (FHWA) has developed a highway traffic noise prediction model that is based on the maximum passby level, L_0, measured at a reference distance D_0.[8] For $D_0=15$ m (i.e., close to the road), Eq. (11) simplifies to the form,

$$L_0=L_{WA}+10\ \log\left\{\frac{s_0}{4\pi D_0^2}\right\}, \tag{27}$$

and Eqs. (20) and (21) allow the modification of the formula for L_{AT} calculation,

$$L_{AT}=L_0+\Delta_{\text{traffic}}+\Delta_{\text{distance}},\quad D<150\ \text{m}. \tag{28}$$

Here, the flow traffic adjustment and the distance adjustment are given by

$$\Delta_{\text{traffic}}=10\ \log\left\{\frac{\pi ND_0}{VT}\right\}, \tag{29}$$

and

$$\Delta_{\text{distance}}=10\ \log\left\{\frac{D_0}{D}\left[1-\frac{\sqrt{\gamma}\cdot D}{\sqrt{\gamma\cdot D^2+(H_s+H_0)^2}}\right]\right\}. \tag{30}$$

Similar to the procedure discussed above, Eq. (27) may be considered a theoretical prediction with two adjustable parameters, L_0 and γ. On a hard surface, such as a paved parking lot, the ground coefficient has a value of $\gamma<10^{-5}$. A soft surface, such as grass, is characterized by $\gamma>10^{-5}$.

C. Railroad noise

The time history of noise from a passing train indicates that the model of a homogeneous and finite line of incoherent point sources, with directivity proportional to $\cos^2\Phi$, can be used.[9–12] If the convection and Doppler effect are neglected[13] then the radiation pattern is given by $Q=3\cos^2\Phi/4\pi$ [Eq. (2)]. Thus high above the ground, the squared A-weighted sound pressure due to a line source of unit length, $l_0=1$ m, may be written as [Eq. (1), Fig. 5],

$$\tilde{p}_A^2=3\frac{P_Al_0}{4\pi\cdot r^2}\cos^2\Phi. \tag{31}$$

Close to the ground [Eq. (3)] the angle Φ remains in the x,y plane (Fig. 3), and from Eqs. (10) and (31) we get

$$\tilde{p}_A^2=3\frac{\beta P_Al_0\cdot\rho c}{4\pi d^2}\cos^2\Phi\cdot\left[1+\gamma\cdot\left(\frac{d}{H_s+H_0}\right)^2\right]^{-1}. \tag{32}$$

Similar to the equations described above we obtain the sound exposure as

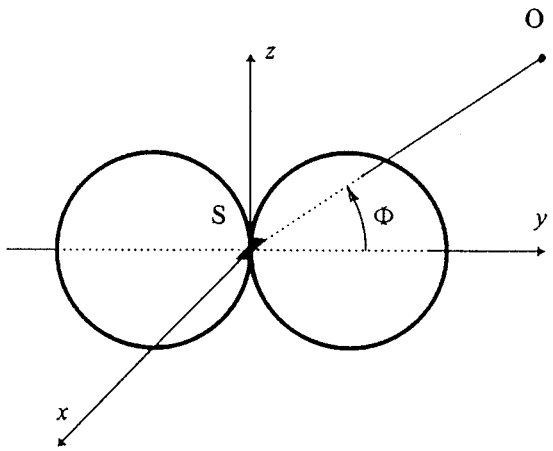


FIG. 5. Radiation pattern of the unit length line source [Eq. (31)].

$$E=\frac{Dl}{Vl_0}\int_{-\pi/2}^{+\pi/2}\frac{\tilde{p}_A^2(\Phi)}{\cos^2\Phi}\,d\Phi, \tag{33}$$

where l expresses the train's length. Substituting $\tilde{p}_A^2$ with Eq. (32), we arrive at the sound exposure level of noise that is emitted by a single train:

$$L_{AE}=L_{WA}+10\ \log\left\{\frac{3ll_0}{8DVt_0}\right\}+C(\sqrt{\gamma}\cdot D). \tag{34}$$

Here, $t_0=1$ s, D denotes the perpendicular distance to the track, the quantity

$$L_{WA}=10\ \log\left\{\frac{\beta P_Al_0}{P_0}\right\},\quad P_0=10^{-12}\ W, \tag{35}$$

expresses the effective sound power level of the unit length source, and

$$C(\sqrt{\gamma}\cdot D)=10\ \log\left\{1-2\left[\frac{\sqrt{\gamma}\cdot D}{H_s+H_0}\right]^2\cdot\left[1-\frac{\sqrt{\gamma}\cdot D}{\sqrt{\gamma\cdot D^2+(H_S+H_0)^2}}\right]\right\} \tag{36}$$

accounts for the ground attenuation.

Field measurements[14,15] indicate that the derived model can be applied at distances $D<150$ m.

In order to estimate the ground coefficient, γ, and the effective sound power level of the unit length source, L_{WA}, we have to perform two simultaneous measurements, $L_{AE}(D_*)$ and $L_{AE}(2D_*)$. Using Eq. (34) we obtain

$$M_c(\xi)=L_{AE}(D_*)-L_{AE}(2D_*)-3, \tag{37}$$

where

$$M_c=10\ \log\left\{\frac{(\xi^2-2)\sqrt{1+\xi^2}+2}{(\xi^2-8)\sqrt{1+\xi^2/4}+8}\cdot\frac{\sqrt{1+\xi^2/4}}{\sqrt{1+\xi^2}}\right\}, \tag{38}$$

with ξ defined by Eq. (24). The function $M_c(\xi)$ is plotted in Fig. 6.

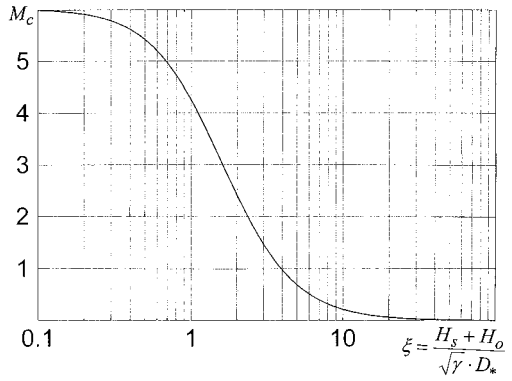


FIG. 6. Dependence of M_c upon ξ [Eq. (38)].

Proceeding as in the case of road-traffic noise (Sec. III A), we determine the ground coefficient, γ, and the effective sound power level of unit length source, L_{WA}.

Example 3

At a height of $H_0 = 1$ m above very soft ground, two simultaneous measurements of the sound exposure level, L_{AE}, at distances $D_* = 25$ m and $2D_* = 50$ m, have been made. The day was cloudy with no wind. The train of length $l = 200$ m was moving with the steady speed $V = 29$ m/s. The outcome of the measurements,

$$L_{AE}(25) = 90 \text{ dB} \quad \text{and} \quad L_{AE}(50) = 83 \text{ dB}$$

yields $\xi = 1.05$ [Eq. (37), Fig. 6]. Since $H_s \approx 1$ m (top of a rail) from Eqs. (24), (34), and (36) it follows that $\gamma = 7 \times 10^{-3}$ and $L_{WA} = 101$ dB.

IV. SUMMARY

Simplified models of propagation of noise generated by a stationary source [Eqs. (11), (13)], moving vehicles [Eqs. (20), (21)], and railroad trains [Eqs. (34), (36)], respectively, have been derived under following assumptions:

(1) the ground surface is plane, homogeneous, and uniterrupted by buildings or any other obstacles,

(2) the ground effect is the only factor affecting geometrical spreading,

(3) the Weyl–Van der Pol solution provides an exact theory of the ground effect,

(4) vehicles and trains are moving at steady speeds along straight lines.

The proposed models contain two empirically based parameters, the ground coefficient γ, and the effective sound power level L_{WA}, which require two simultaneous measurements of L_A (stationary source), L_{AT} (road traffic), and L_{AE} (railroad) for their determination.

Field measurements indicate that the models can be applied at distances $D < 150$ m from the noise source.

ACKNOWLEDGMENT

The study was partly supported by a grant from the Committee of Scientific Research-KBN (Poland).

[1] J. E. Piercy and G. A. Daigle, ''Sound propagation in the open air,'' in *Handbook of Acoustical Measurements and Noise Control*, edited by C. M. Harris (McGraw-Hill, New York, 1991).
[2] G. S. Anderson and U. J. Kurze, ''Outdoor sound propagation,'' in *Noise and Vibration Control Engineering*, edited by L. L. Beranek and I. L. Ver (Wiley, New York, 1992).
[3] K. Yamamoto, M. Yamashida, and T. Mukai, ''Revised expression of vehicle noise over ground,'' J. Acoust. Soc. Jpn. (E) **15** 233–241 (1994).
[4] ANSI S1.1-1994, ''Acoustical Terminology'' (Acoustical Society of America, New York, 1994).
[5] K. Attenborough, ''Review of ground effects on outdoor sound propagation from continuous broadband noise,'' Appl. Acoust. **24** 289–319 (1988).
[6] K. Attenborough, S. Taherzadeh, H. E. Bass, X. Di, R. Raspet, G. R. Becker, A. Gudesen, A. Chrestman, G. A. Daigle, A. L'Esperance, 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,'' J. Acoust. Soc. Am. **97**, 173–191 (1995).
[7] K. M. Li, K. Attenborough, and N. W. Heap, ''Source height determination by ground effect inversion in the presence of sound velocity gradient,'' J. Sound Vib. **145** 111–128 (1991).
[8] W. Bolwby, ''Highway noise prediction and control,'' in *Handbook of Acoustical Measurements and Noise Control*, edited by C. M. Harris (McGraw-Hill, New York, 1991).
[9] B. Hemsworth, ''Prediction of train noise,'' in *Transportation Reference Book*, edited by P. M. Nelson (Butterworth, London, 1987).
[10] D. Hohenwarter, ''Railway noise propagation models,'' J. Sound. Vib. **141** 17–41 (1990).
[11] R. Makarewicz, ''Prediction of train noise in open space,'' J. Acoust. Soc. Jpn. (E) **12** 195–202 (1991).
[12] Y. Okumura and K. Kuno, ''Multiple regression analysis of railway noise based on finite line source models including the excess attenuation,'' J. Acoust. Soc. Jpn. (E) **13** 161–169 (1992).
[13] R. Makarewicz, ''Influence of Doppler and convection effects on noise propagation,'' J. Sound Vib. **155** 353–364 (1992).
[14] ÖAL-Richtlinie Nr.30, ''Berechnung der Schalimmission durch Schienenverkehr,'' Österreichischer Arbeitsring fur Lärmbekämpfung, Wien (1990).
[15] Y. Okumura and K. Kuno, ''Statistical analysis of field data of railway noise and vibration collected in an urban area,'' Appl. Acoust. **33** 263–280 (1991).

Wavelet-based denoising of underwater acoustic signals

L. G. Weiss and T. L. Dixon
The Pennsylvania State University, Applied Research Laboratory, P.O. Box 30, State College, Pennsylvania 16804

(Received 20 October 1995; accepted for publication 22 July 1996)

Underwater environmental measurements of the ocean require signals that are free from unwanted backscatter and clutter. Removing these unwanted signal components usually amounts to applying some form of filtering technique such as a high pass filter, a bandpass filter, a Wiener filter, etc. These approaches however are limited in their abilities to remove acoustic returns that vary spectrally. This paper presents a multiresolution approach to removing unwanted backscatter from high-frequency underwater acoustic signals and compares it to high pass filtering of the same signals. The filtering approach presented applies wavelet transforms for signal recovery and denoising of high-frequency acoustic signals. It is shown that by computing a wavelet transform of the returned signals, applying a denoising technique, and then reconstructing the signal, additional unwanted backscatter can be removed.
[S0001-4966(97)03612-6]

PACS numbers: 43.60.Bf, 43.30.Sf, 43.30.Vh, 43.60.Cg [JLK]

INTRODUCTION

Accurate environmental acoustic measurements in the ocean are often difficult to accomplish because of unwanted backscattering in the acoustic return. Unidentified scatterers such as zooplankton often clutter the desired signal, while the density and compressibility of these biologics alter the distribution of inhomogeneities from location to location. Their scattering effects in acoustic returns have a varying effect on the signal. Removal of signal returns caused by dominant unwanted scatterers is therefore of significant importance when conducting underwater environmental measurements. See Fig. 1.

For high-frequency acoustic scattering, the dominant scatterers are often zooplankton such as euphasiids, copepods, pteropods, etc. that range in size from 0.1 mm to 1 cm. Goodman (1990) has reviewed the theory and applications of acoustic scattering from ocean variability and identifies one area of acoustic applications to oceanography that remains largely unexploited is the use of high-frequency ($\geqslant$20 kHz, usually 100's of kHz) acoustic scattering for environmental oceanographic measurements.

The most well known problems with operating at higher frequencies underwater are signal attenuation and transmission loss, so that long range measurements become difficult (Urick, 1983). If transmission loss and attenuation are acceptable for ranges of interest, operating at higher frequencies is desirable since higher frequencies offer the advantage of improved resolution. The problem with high-frequency acoustics when making environmental measurements is that large scatterers such as certain types of zooplankton may dominate the return and prevent an accurate acoustic environmental measurement (Holliday and Piper, 1980). Various filtering approaches have been applied to remove the dominant scattering, but with limited effectiveness. For example, high pass filtering approaches have been applied to the received signals to remove the unwanted backscattering since the backscattering typically concentrates in the low frequencies. However, it turns out that some of the backscattering appears across several frequencies so that a more flexible filtering approach is needed.

The filtering approach presented applies wavelet transforms for signal recovery and denoising of these high frequency acoustic signals. Wavelet transforms are applied since they perform a multiresolution decomposition in time and frequency and therefore are well suited to removing specific unwanted signal components. It is shown that by computing a wavelet transform of the returned signals, applying a denoising filtering technique, and then reconstructing the signal, some additional unwanted backscatter can be removed. This wavelet transform approach to noise reduction in high-frequency underwater acoustic signals is then compared to the standard high pass filter approach to removing unwanted backscatter.

The techniques presented are based on experience with in-water data. However, the results presented are from simulated data. Simulated data offers the advantage of providing ground truth to the analysis. In addition, simulated data offers the option of controlling the severity of the unwanted backscattering. This allows for evaluating the limits of the denoising algorithm.

This paper shows that signal recovery and noise reduction using wavelet transforms can be accomplished while retaining a significant amount of the signal energy and without significantly degrading performance. It is shown that this approach to wavelet based denoising of signals offers promise to the problem of signal recovery of acoustic signals contaminated with unwanted backscatter.

I. THE SIGNAL MODELS

The water column is insonified by a pulse with center frequency, f_c, and bandwidth, B. For the data in this paper, the received pressure waves are monitored in time at two (nonoverlapping) linear arrays orthogonal to each other.

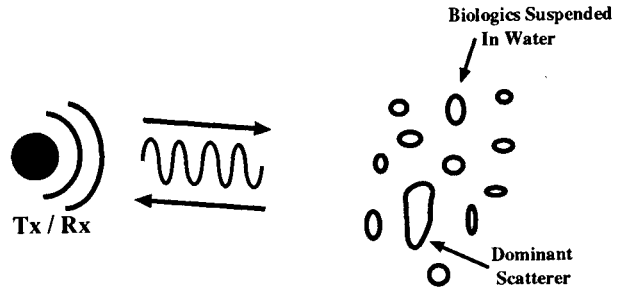


FIG. 1. Dominant scatterers often mask desired signal components.

There are several variables that affect the return signal: backscatter strength, type of scatterers, distribution of scatterers, platform motion, etc.

In any small range increment, Δr, the medium contains many scatterers moving with the same radial velocity whose superposed reflection can be characterized as a random variable. See Fig. 2. The medium response at time t to a scattering increment at range $r_0 = c\tau_0/2$ is given by (Ziomek, 1985)

$$\sqrt{E} b(\tau_0) f(t-\tau_0) e^{-j\omega_d(\tau_0)t} \Delta\tau, \tag{1}$$

where E is the energy of the transmitted signal, $f(t)$, $b(\tau_0)$ is a zero mean Gaussian random variable describing the reflection at time τ_0 and satisfying $E[b(\tau)b(\tau_0)^*] = \sigma_b^2 \delta(\tau-\tau_0)$, and $\omega_d(\tau_0)$ is the Doppler shift associated with this increment. Reflections at different ranges are assumed to be statistically uncorrelated.

By superposition, the medium's response, $s(t)$, to the transmitted signal is (Ziomek, 1985)

$$s(t) = \sqrt{E} \int_{-\infty}^{\infty} b(\tau) f(t-\tau) e^{j\omega_d(\tau)t} d\tau. \tag{2}$$

The signal received at the array also contains a noise component, $n(t)$, that models the noise sources present in the system and the environment. The received signal, $r(t)$, is then given by

$$r(t) = s(t) + n(t), \quad 0 \leq t \leq T. \tag{3}$$

where T is the length of the observation interval and $n(t)$ is a zero mean, stationary, white Gaussian process with $E[n(t)n^*(\tau)] = N_0 \delta(t-\tau)$. It is also assumed that the noise component $n(t)$ is uncorrelated with the signal component $s(t)$ so that $E[s(t)n^*(t)] = 0$.

II. NOISE CORRELATED ACROSS AN ARRAY

For a linear array with M elements, the acoustic returns at each element can be expressed as

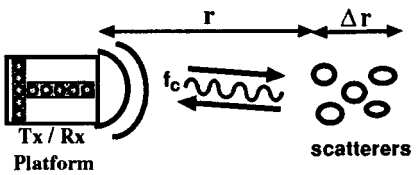


FIG. 2. An increment, Δr, of scatterers relative to the platform.

$$r_i(t) = s_i(t) + n_i(t), \quad 0 \leq t \leq T, \quad i = 1,\ldots,M. \tag{4}$$

In the absence of a dominant scatterer, the noise at any two array elements is uncorrelated:

$$E[n_i(t)n_j^*(t)] = N_0, \quad \text{if } i = j$$
$$= 0, \quad \text{if } i \neq j. \tag{5}$$

If a dominant scatterer is present, it may appear in several (but not necessarily all) of the returns at the array, depending on the scatterer's size and location relative to the array. In this case, the noise may correlate across the array. The noise term at element i then becomes

$$N_i(t) = n_i(t) + d_i(t), \tag{6}$$

so that

$$r_i(t) = s_i(t) + N_i(t). \tag{7}$$

Here, $d_i(t)$ is the acoustic return from the dominant scatterer. It is separated from the signal component $s_i(t)$ since it may not appear in all the acoustic returns at each element of the array, yet it is common to the elements in which it does appear. The size and properties of the dominant scatterer relative to the array size and location determine which elements of the array will receive the dominant component.

If two distinct array elements, $i = k,l$, receive a return from the same dominant scatterer, then the noise is correlated and can be characterized by

$$E[N_k(t)N_l^*(t)] = E[n_k(t)n_l^*(t)] + E[d(t-\tau_k)d^*(t-\tau_l)]$$
$$= E[d(t-\tau_k)d^*(t-\tau_l)], \tag{8}$$

where τ_k and τ_l denote the delay in the arrival time of the dominant scatterer's return at elements k and l of the array, and the dominant scatterer's return at element k is uncorrelated with the noise $n_k(t)$. By cross correlating each array element with each of the other array elements, one can determine if a dominant scatterer is present. This leads to the hypotheses at each array element being

$$H_0 : r_i(t) = s_i(t) + n_i(t), \text{ no dominant scatterer}$$
$$H_1 : r_i(t) = s_i(t) + n_i(t) + d_i(t), \tag{9}$$
$$\text{dominant scatterer present.}$$

If two linear arrays are used, then correlations can be computed between elements of each array. By correlating each element in one array with each element of the other array, a correlation matrix results.

Denote one array by $\overline{x}$ whose returns at each element at time t are $x_i(t)$, $i = 1,\ldots,M$, and denote the other array by $\overline{y}$ whose returns at each element are $y_j(t)$, $j = 1,\ldots,P$. To create a correlation matrix, the received signals at each of the elements of $\overline{x}$ are correlated with the received signals at each of the elements of $\overline{y}$, with the correlation denoted by ρ_{ij}, $i = 1,\ldots,M$, $j = 1,\ldots,P$. See Fig. 3.

The value of ρ_{ij} is computed using

$$\rho_{ij} = \frac{\int_0^T x_i(t) y_j^*(t) dt}{\left[\int_0^T |x_i(t)|^2 dt \int_0^T |y_j(t)|^2 dt\right]^{1/2}}, \quad 0 \leq |\rho_{ij}| \leq 1, \tag{10}$$

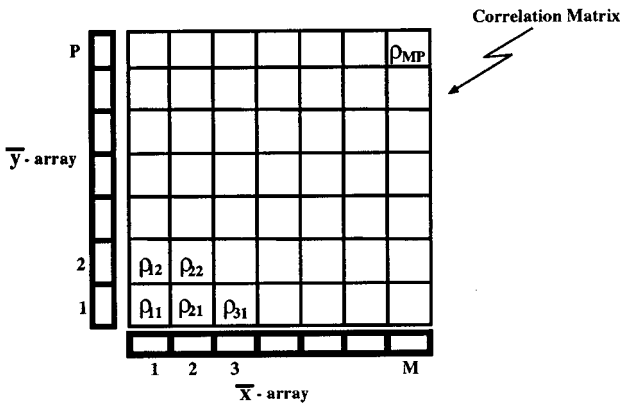


FIG. 3. Correlation matrix for two linear arrays.

where T is the pulse length of the signals, $x_i(t)$ and $y_j(t)$, $i=1,\ldots,M$, $j=1,\ldots,P$ and $x_i(t)$ and $y_j(t)$ are assumed to be mean zero. To display the resulting correlations, a color coded correlation matrix is generated with red denoting the highest correlation value (1) and blue the lowest. If an unwanted dominant scatterer is illuminated, its return generally appears in several of the array elements and increases the cross correlation values in those cells of the correlation matrix.

For our application, the array elements are omnidirectional and arranged in a ''$\perp$'' shape (but do not overlap at their intersection). They are then time adjusted (with a signal time delay) to correspond to a ''+'' shape. The time adjustment affects which elements a scatterer appears in, only in the sense that the peak has been moved from being at the bottom of the matrix to the center. With the + configuration, uncorrupted signals have a solid background with one central red peak whose value is nominally 1. Since the two arrays do not actually intersect, signal differences cause this peak amplitude to decrease. Received signals containing unwanted backscatter have correlation matrices saturated with red and yellow striations. These striations are indicative of a dominant unwanted scatterer corrupting the received signals. Figure 4 shows correlation matrices for two sets of actual in-water data collected at ranges from 5 to 80 m from the array.

The goal is to identify returns that contain dominant scattering and filter those signals so that further environmental measurements can be obtained from these returns. The correlation matrix is used to evaluate the level of corruption in the acoustic returns at the array.

III. APPROACH

In this paper, correlation matrices of corrupted signals are computed both before and after a wavelet denoising scheme is applied. Correlation matrices created after high pass filtering of the signal are also presented for comparison. The wavelet transform filtering is shown to reduce the dominant scattering effects while retaining a significant amount of the signal energy. The two features of the correlation matrix that are of interest in determining the improvements of the denoising algorithms on the reconstructed (filtered) signals are the amplitude of the correlation peak and the location of the peak

A. High pass filtering

A signal acquired by a sensing device is usually degraded by the environment through which it passes. The signal is often recovered by inverting the convolution integral

$$y(t)=\int_a^b h(t-\tau)x(\tau)d\tau+n(t), \tag{11}$$

where $x(t)$ is the original signal, $y(t)$ is the degraded signal, $n(t)$ is zero mean white Gaussian noise, and $h(t)$ is the convolution kernel.

The convolution kernel $h(t)$ can be designed to take one of several forms: a low pass filter, a bandpass filter, a high pass filter, etc. For the data in this paper, the unwanted corruption concentrates in the low frequencies so that we wish to retain the high-frequency acoustic data. Such data is often filtered with a high pass filter, and consequently, we apply a high pass filter to the data so that a comparison can be made to the wavelet denoising algorithm.

For our application, the received signal is basebanded and normalized to a sampling rate of $f_s=1$ Hz. The high pass filter applied has a hard cutoff at 0.125 Hz, i.e., rectangular in frequency.

B. Denoising with wavelet transforms

The idea behind denoising signals with wavelet transforms is to compute a wavelet transform of a noisy signal and apply a soft threshold that is SNR dependent (DeVore and Lucier, 1992; Weaver *et al.*, 1991). The soft threshold has two cutoffs. The first is a frequency cutoff chosen to agree with the cutoff frequency of the high pass filter so that a fair comparison can be made. The second cutoff is in the wavelet transform domain where wavelet coefficients exceeding a specific value are filtered. Both criteria must be met in the denoising algorithm for a coefficient to be removed prior to signal reconstruction.

This approach usually outperforms linear smoothing techniques such as high pass filtering since the multiscale nature of wavelets analyzes a signal with a pair of high pass and low pass filters. See, for example, Mallat (1989). This approach also avoids inversion of the convolution integral.

The *wavelet transform*, $W_g f(s,\tau)$, of a signal $f\in L^2(R)$ with respect to the mother wavelet $g(t)$ is (Grossmann and Morlet, 1984)

$$W_g f(s,\tau)\overset{\text{def}}{=}\frac{1}{\sqrt{|s|}}\int_{-\infty}^{\infty} f(t)g^*\left(\frac{t-\tau}{s}\right)dt, \tag{12}$$

where $*$ denotes complex conjugation, and $g(t)$ is assumed to be admissible [i.e.,

$$c_g\overset{\text{def}}{=}\int_{-\infty}^{\infty}\frac{|G(\omega)|^2}{|\omega|}\,d\omega<\infty,$$

where $G(\omega)$ is the Fourier transform of $g(t)$].

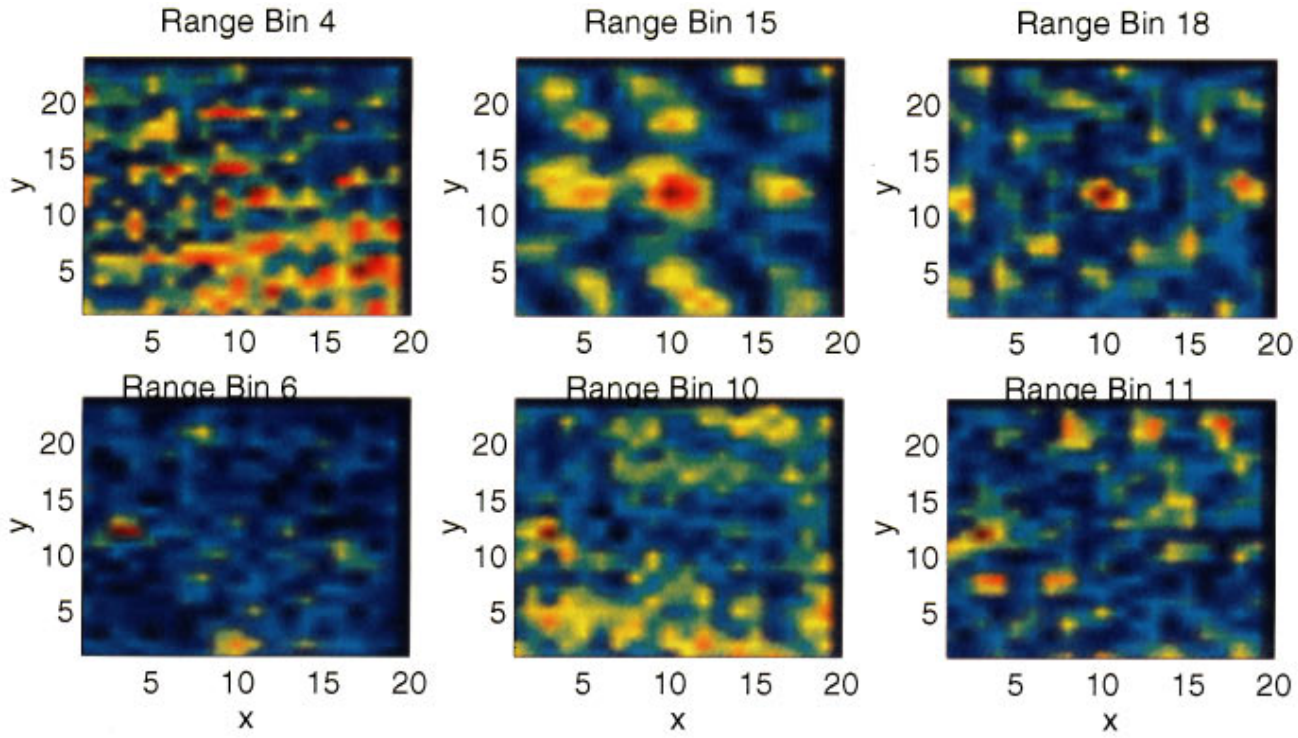


FIG. 4. Correlation matrices for two in-water data sets.

Recovery of the signal $f(t)$ from its wavelet transform $W_g f$ is given by the inversion formula for wavelets:

$$f(t)=\frac{1}{c_g}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_g f(a,b)\,\frac{1}{\sqrt{|a|}}\,g\left(\frac{t-b}{a}\right)\frac{da\;db}{a^2}. \tag{13}$$

The wavelet denoising scheme amounts to selecting an appropriate mother wavelet computing the wavelet transform of the received signal with respect to this mother wavelet, applying a soft threshold (one that depends on time location, frequency location, and level of the wavelet coefficients), removing the threshold crossing coefficients, and reconstructing the signal using the wavelet inversion formula. By doing this, a spectrally varying signal can be appropriately denoised and filtered.

C. The correlation matrix

To measure the amount of corruption in the received signals, a color correlation matrix scheme for signal analysis is generated. A correlation matrix is computed from the $\overline{x}$ and $\overline{y}$ channels of the array for each range bin of interest. An example of a severely corrupted in-water correlation matrix is shown in Fig. 4 in range bin 4 of pulse pair 19. An ex-

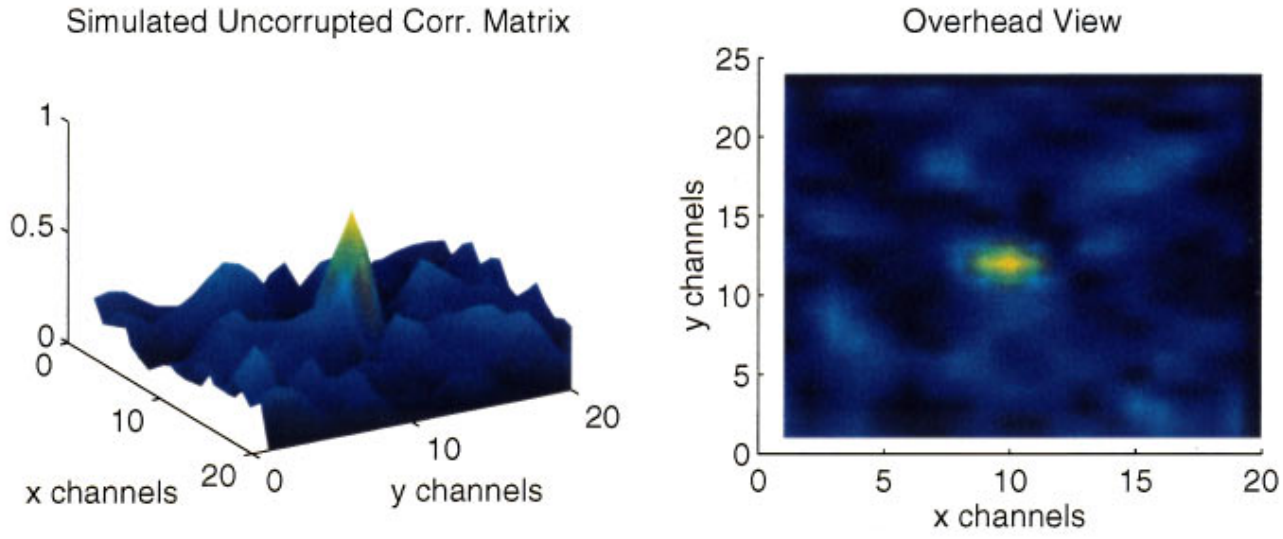


FIG. 5. Correlation matrix of simulated data with no unwanted backscatter.

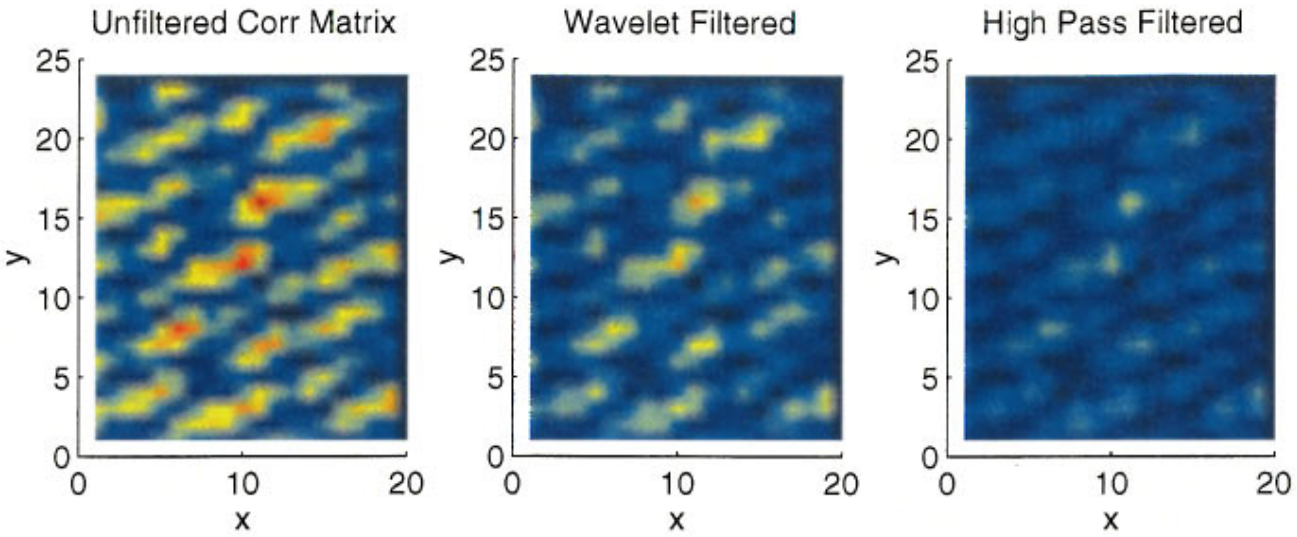


FIG. 6. Simulated data: correlation matrix, denoised matrix, and HPF matrix.

ample of a correlation matrix with no unwanted backscatter is shown in Fig. 5. This matrix was computed from simulated data with no noise. The plot on the left is a 3-D view of the matrix while the plot on the right shows an overhead view. The central peak's amplitude is less than one because of signal differences in the $\perp$ configuration of the two arrays. Since the arrays do not actually intersect, no two elements receive identical signals.

To measure the effectiveness of wavelet denoising for removal of unwanted backscatter from the signals, several simulated data sets were analyzed, two of which are presented to demonstrate the method. The data consists of corrupting the signals used to generate the matrix in Fig. 5. Two fairly high levels of corruption are processed. In particular, the amounts of red and yellow striations are severe when compared to the levels of corruption that usually occur with in-water data. These extreme levels allow us to evaluate the limits of the denoising technique. The results are presented below.

IV. RESULTS

Two pulse pairs of corrupted data are now presented to exemplify the denoising process. For each range bin, overhead views of the unfiltered and filtered correlation matrices are shown. Figure 5 is the original uncorrupted matrix. Its peak is located at $(x,y)=(10,12)$, and its amplitude is 0.6793. The data simulates returns from signals scattered off zooplankton at ranges from 5–80 m with $f_c=300$ kHz. This figure is used as a basis to compare the results from wavelet denoising and high pass filtering of the corrupted data.

The mother wavelet used for the analysis is a Daubechies-4 (Daubechies, 1992). Selection of this mother wavelet was not optimized in any way. Additional performance gains are likely to be achieved if one optimizes the mother wavelet selection. In particular, matching characteristics of the mother wavelet with expected characteristics of the signal usually yields improved performance.

The unfiltered correlation matrix in Fig. 6 has a false

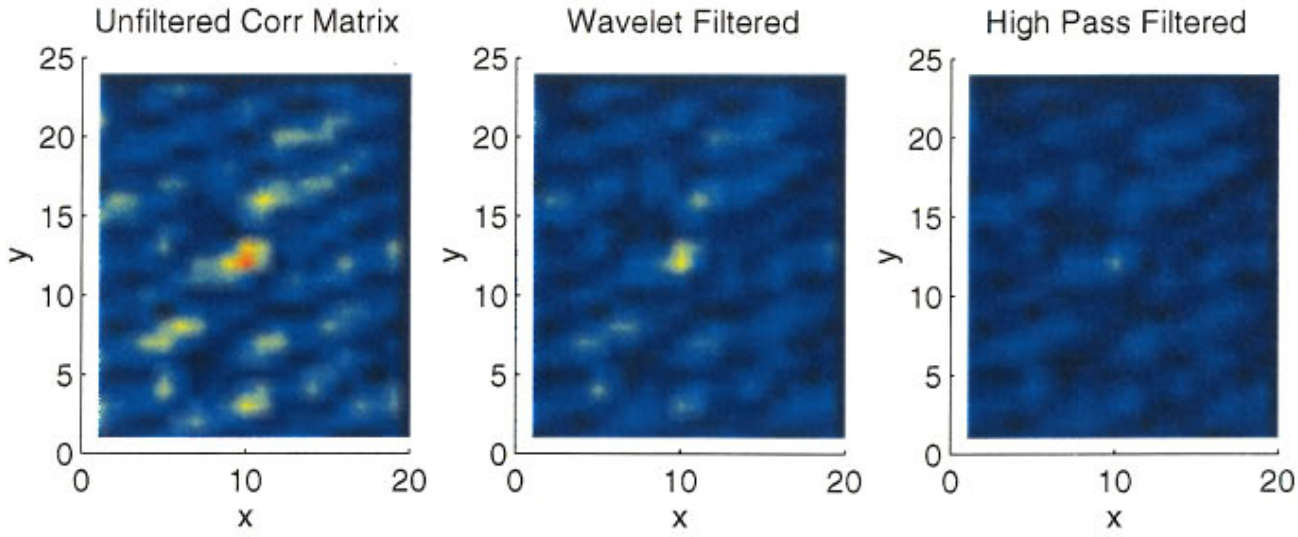


FIG. 7. Simulated data: correlation matrix, denoised matrix, and HPF matrix.

TABLE I. Wavelet denoising.

Data set	Before wavelet denoising: Peak amplitude	Before wavelet denoising: Peak location	After wavelet denoising: Peak amplitude	After wavelet denoising: Peak location	Signal energy retained	Peak amplitude relative error
Fig. 5 (no corruption)	0.6793	(10,12)	…	…	…	…
Fig. 6	0.9175	(11,16)	0.7435	(11,16)	82.54%	9.451%
Fig. 7	0.8146	(10,12)	0.6188	(10,12)	81.35%	8.906%

peak at (11,16) as well as several other high amplitude regions. After wavelet filtering, these amplitudes are greatly reduced. More or less backscatter can be removed depending on how much of the received signal energy must be retained. Wavelet denoising is able to remove several high amplitude regions, but it does not remove the false peak. The second largest peak is at (10,12).

The high pass filter is also unable to remove the false peak. In addition, it suffers from removing too much of the signal energy to render it useful for processing. As seen in the correlation matrix, nearly all remnants of the signal are removed. If less signal energy were filtered (the cutoff frequency were lowered), the corruption in the correlation matrix would not be reduced enough.

Figure 7 shows a second correlation matrix. Here, the wavelet denoised correlation matrix closely resembles that in Fig. 5. High pass filtering of this data removes too much of the signal energy, and the peak amplitude is significantly decreased. Important signal information is lost along with the noise.

Tables I and II summarize the results. The column labeled ''Signal energy retained'' is the amount of energy retained in the corrupted signal after filtering, where the cutoff frequency for the high pass filter is chosen to agree with the cutoff frequency of the soft threshold in the denoising technique. In Table I Fig. 6, for example, the remaining 82.54% of the signal energy still contains unwanted backscatter energy as evidenced by the level of the peak amplitude. The column labeled ''Peak amplitude relative error'' is the percent difference between the level of the peak after filtering and the level of the peak in Fig. 5. These values are discussed in the next section.

V. ANALYSIS

Both high pass filtering and wavelet denoising remove several unwanted components of the corrupted signal. However, the high pass filter suffers since it removes too much of the signal energy to be rendered useful. For it, less filtering is required to retain the peak amplitude, but at the expense of not removing enough signal corruption. In addition, the peak value is often filtered out with the high pass filter while amplitude errors with wavelet denoising remain less than 10%. As shown in Table II, 16%–32% of the peak amplitude is lost with the high pass filter. Wavelet denoising therefore allows for a more accurate estimate of the peak, which is crucial to environmental measurements.

VI. CONCLUSIONS

This paper presents a multiresolution approach to removing unwanted backscatter from high-frequency underwater acoustic signals and compares it to high pass filtering of the same signals. Since the unwanted backscatter typically concentrates in the low frequencies, high pass filters are often applied but with limited effectiveness. It turns out that some of the backscattering actually appears across several frequencies, and so a more flexible filtering approach is needed.

The wavelet denoising approach presented applies wavelet transforms for signal recovery and denoising of high-frequency acoustic signals. Wavelet transforms are applied since they perform a multiresolution decomposition in time and frequency and therefore are well suited for removing specific unwanted signal components that may vary spectrally. It is shown that by computing a wavelet transform of the returned signals, applying a denoising technique, and then reconstructing the signals, additional unwanted backscatter can be removed while preserving important aspects of the signal. In particular, after wavelet denoising, the peak amplitude is within 10% of the actual value while high pass filtering causes errors of 16%–32%.

TABLE II. High pass filter.

Data set	Before HPF: Peak amplitude	Before HPF: Peak location	After HPF: Peak amplitude	After HPF: Peak location	Signal energy retained	Peak amplitude relative error
Fig. 5 (no corruption)	0.6793	(10,12)	…	…	…	…
Fig. 6	0.9175	(11,16)	0.5678	(11,16)	61.41%	16.41%
Fig. 7	0.8146	(10,12)	0.4645	(10,12)	62.07%	31.62%

ACKNOWLEDGMENTS

The authors thank Dr. Leon H. Sibul at Pennsylvania State University's Applied Research Laboratory for many helpful comments and Dr. Douglas Lake at the Office of Naval Research.

Daubechies, I. (**1992**). *Ten Lectures on Wavelets* (SIAM, Philadelphia).

DeVore, R. A., and Lucier, B. J. (**1992**). ''Fast Wavelet Techniques for Near-Optimal Image Processing,'' *IEEE-ICASSP-93* (IEEE Military Comm. Conf. New York), pp. 48.3.1–48.3.7.

Goodman, L. (**1990**). ''Acoustic Scattering From Ocean Microstructure,'' J. Geophys. Res. **95**(c7), 11557–11573.

Grossmann, A., and Morlet, J. (**1984**). ''Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape,'' SIAM J. Math. Anal. **15**(4), 723–736.

Holliday, D. V., and Piper, R. E. (**1980**). ''Volume Scattering Strength and Zooplankton Distributions at Acoustic Frequencies Between 0.5 and 3 mHz,'' J. Acoust. Soc. Am. **67**, 135–146.

Mallat, S. G. (**1989**). ''Multiresolution Approach to Wavelets in Computer Vision,'' in *Wavelets: Time-Frequency Methods and Phase Space*, edited by J. M. Combes, A. Grossman, and Ph. Tchamitchian (Springer-Verlag, New York), 2nd ed., pp. 313–327.

Urick, R. (**1983**). *Principles of Underwater Sound* (McGraw-Hill, New York), 3rd ed.

Weaver, J. B., Xu, Y., Healy, Jr., D. M., and Cromwell, L. D. (**1991**). ''Filtering Noise From Images with Wavelet Transforms,'' Magn. Resonance Med. **21**, 288–295.

Ziomek, L. J. (**1985**). *Underwater Acoustics: A Linear System Theory Approach* (Academic, New York).

Shallow water beamforming with small aperture, horizontal, towed arrays

James H. Wilson[a)]
Neptune Sciences, Inc., 4711 Viewridge Avenue, Suite 150, San Diego, California 92123

Robert S. Veenhuis[b)]
Naval Postgraduate School, Department of Oceanography, Monterey, California 93943

(Received 30 May 1995; accepted for publication 18 July 1996)

This paper shows that the performance of plane-wave beamformers is significantly degraded in shallow water for horizontal arrays, even for short aperture arrays. Horizontal towed arrays must have short apertures for practical reasons, and the plane-wave beamformer needs to be augmented to optimize performance in shallow water. Matched-field processing (MFP) is a beamforming technique which allows for the detection and localization of an acoustic source in range and depth using passive sonar. The performance of MFP to augment plane-wave beamforming in shallow water environments using a short aperture horizontal line array is investigated. It is realized that MFP performance is enhanced when vertical arrays are used in deep water, or large aperture horizontal arrays are available in shallow water. However, this paper addresses the realistic case of operational, tactical horizontal arrays which are practical to deploy in shallow water. This paper attempts to integrate inverse beamforming (IBF) [A. H. Nuttall and J. H. Wilson, J. Acoust. Soc. Am. **90**, 2004–2019 (1991); J. H. Wilson, J. Acoust. Soc. Am. **98**, 3250–3261 (1995)], a plane-wave beamformer, with MFP so that the strengths of each are utilized. Results of test cases using synthetic data are presented to evaluate the effects of frequency, bottom type, relative target bearing, and sound-speed profile on MFP and IBF performance. IBF consists of three algorithms: the Fourier integral method (FIM) beamformer; the eight nearest neighbor peak picker (ENNPP); and a sophisticated M of N tracker. It is shown that MFP performance is significantly better than plane-wave beamforming for higher frequencies, for more reflective bottom types in shallow water, and for relative target bearings away from the broadside beams. IBF, or plane-wave beamforming, performs well at very low frequencies (VLF), in mud/silt-clay bottoms, and at beams near broadside. It is also shown that the performance of any plane-wave beamformer, including IBF, is severely degraded in shallow water for relative bearings away from the broadside beams, particularly at higher frequencies. Finally, the MFP algorithm in conjunction with the ENNPP and M of N tracker are shown to perform well in the real, deep water ocean environment. No appropriate shallow water measured data is available to evaluate the integrated IBF/MFP algorithm in shallow water. [S0001-4966(97)03701-6]

PACS numbers: 43.60.Gk, 43.30.Wi [JLK]

INTRODUCTION

The original intent of this research was to investigate the feasibility of utilizing matched field processing (MFP)[1–15] as a range and depth estimation technique for sonars with short aperture horizontal arrays operating in shallow water. As such, MFP supplements plane-wave beamforming's detection and tracking capability with single heading target range and depth estimations using a short aperture horizontal line array. However, the results of the analyses showed that in shallow water environments, plane-wave beamformers for horizontal line arrays suffer a significant fundamental performance degradation. Since all of the Navy's operational sonars for horizontal line arrays use plane-wave beamforming, this is considered a significant problem for anticipated shallow water operations. MFP should therefore be regarded as not just a supplement to plane-wave beamforming, but as an absolutely necessary complement to conventional plane-wave beamformers with horizontal line arrays in shallow water. The use of MFP in deep water at tactical ranges is also addressed.

The performance of a plane-wave beamformer is degraded simply because propagation in shallow water environments produces a received signal that is not estimated well by a single plane wave in those cases where the acoustic frequency is near design frequency or when the signal arrival is off broadside. The problem is illustrated in Fig. 1 for a 48-element, equally spaced horizontal line array with design frequency f_0, where beamformer output is plotted as a function of range for a range independent, isospeed medium with a hard, reflective bottom and a water depth of 100 m. The magnitude of the beamformer output for this plot is normalized to unity for a single plane wave. Figure 1(a) shows that for a relative target bearing of 90° (i.e., broadside), the plane-wave approximation is very good. This is

[a)]Work performed while occupying ONR sponsored Arctic Chair in Marine Sciences, Naval Postgraduate School, Department of Oceanography.
[b)]Currently stationed at the Naval Oceanographic Office (NAVOCEANO), Code N22, Stennis Space Center, MS 39522-5001.

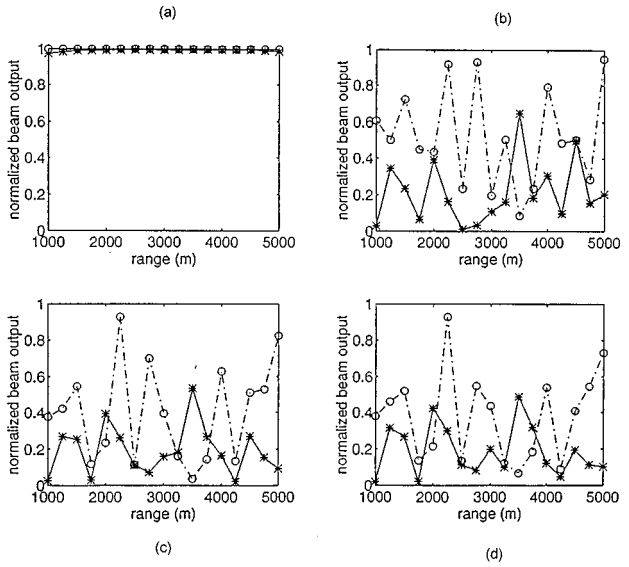


FIG. 1. Normalized plane-wave beamformer output versus target range. (a) 90°; (b) 60°; (c) 30°; (d) 0° = end fire. ∗ = design frequency; 0 = 21% of design frequency.

because the distance from the source to each hydrophone is very nearly the same for all of the array hydrophones when the array is broadside to the source, thus causing the acoustic signal to be in phase over the entire array aperture. In terms of normal mode modeling, the modal wavefront at equal range is approximated well by a plane wave. However, as the target relative bearings get closer to 0° (i.e., forward endfire) as shown in plots (b)–(d), the coherence between the signal and plane-wave model decreases significantly. This is consistent with the fact that the propagation modeled by the normal mode solution to the wave equation is significantly different from a plane wave when the range from each hydrophone to the source varies across the array. The phase of the interfering mode functions is not approximated well by a plane wave when the range changes significantly relative to the acoustic wavelength (thus the degradation near array design frequency and off broadside beams). Based on this result, the performance of plane-wave beamformers in shallow water is expected to be fundamentally degraded, thereby promoting research into alternative beamforming methods to augment plane-wave beamforming for shallow water environments. Mismatch is the term normally used to describe the difference between modeled and measured environments for MFP applications, but is not used in this paper because the modeled signal is simply a plane wave.

Target motion analysis (TMA) with passive sonar has traditionally been accomplished by conducting a series of maneuvers while measuring the target's bearings and bearing rates, and then estimating the target's course, speed, and range by utilizing Eklund Ranging and other algorithms. TMA has been highly successful against targets operating in the open ocean where the ray theory assumption is valid, thereby facilitating the use of plane-wave beamformers. However, the end of the cold war and the increased threat of diesel submarines in shallow water from third world countries has stimulated interest in passive ranging methods on a single heading. As was previously discussed, this is a complex acoustic problem due to the propagation characteristics of shallow water environments, which severely and fundamentally degrade the performance of plane-wave beamformers. Also, restricted maneuverability in shallow water makes conventional TMA difficult, particularly when a towed array is utilized.

Matched-field processing (MFP) is a promising solution to this problem because the MFP algorithm requires no course changes and it utilizes the multiple acoustic arrival paths of shallow water environments to estimate target bearing, range, and depth. Although the intensive computer processing requirements made this method impractical when it was first developed, recent improvements in computer hardware and software technology have made MFP both possible and relatively easy to perform in real time. A microprocessor is fast enough to perform the MFP calculations in real time, and is also sufficiently compact and portable for at-sea tests. These advancements in technology have made MFP a visible beamforming method to supplement IBF for shallow and deep water environments.

The objective of this paper is to assess the performance of the IBF version of MFP with a short aperture horizontal line array in deep and shallow water as a necessary supplement to plane-wave beamforming. The MFP/IBF algorithm will be evaluated in a variety of shallow water environments and in a typical deep water environment. The shallow water test cases allow for an evaluation of MFP performance in the absence of measured shallowed water data. MFP/IBF is evaluated with simulated and measured data in a typical deep water environment.

I. ACOUSTIC PROPAGATION CHARACTERISTICS

A. Deep water acoustic propagation

In deep water environments signal transmission loss (TL) and horizontal signal coherence are measured, analyzed, and modeled as separate acoustic parameters. In the 1970s and 1980s, horizontal signal coherence proved to be sufficient to support the use of very long arrays, as evidenced by a successful experiment with a towed array having an aperture of over one mile. Array signal gain (ASG) can usually be assumed as ideal ($20 \log M$) in deep water near broadside, where ''M'' is the number of elements in an equally spaced line array.

TL modeling has received a great deal of attention in deep water, with the development of very sophisticated and successful TL models such as the parabolic equation (PE), ASTRAL, and the finite element parabolic equation (FEPE). The bottom interacting energy is of relatively minor importance for long horizontal arrays in deep water because this energy is dominant only at ranges between the direct path and first convergence zone (CZ) and also between successive CZs. The TL from bottom interacting paths is very large beyond the first CZ and is most often significantly greater than figure of merits (FOM) for typical threats of

interest. Therefore, the bottom interacting paths do not generally impact sonar performance significantly in deep water. Sonar systems could still benefit from the MFP's range and depth prediction, especially if the array has a vertical aperture. However, short aperture horizontal tactical arrays are impacted by bottom and other multipath energy at tactical ranges (0–40 km).

B. Shallow water acoustic propagation

In shallow water the previous discussion does not apply. Bottom interacting energy dominates signal propagation below 1 kHz in most shallow water areas, and may impact propagation well above 1 kHz. Both TL and horizontal signal coherence show high spatial variability and are impacted greatly by the geoacoustic properties of the bottom and subbottom in shallow water.[16–18] Recent shallow water analyses[19] have shown that TL amplitude fluctuations and signal coherence phase fluctuations are highly correlated as a function of range, and closely related to the bottom and subbottom geoacoustic properties. Furthermore, for ranges up to 25 km, TL amplitude fluctuations occur in the interval of current FOMs for shallow water diesel threats. The test cases in this paper will show that signal coherence phase fluctuations significantly impact ASG for plane-wave beamformers, thus degrading the performance of plane-wave beamformers in shallow water. As a result, signal coherence and TL must be evaluated jointly in diverse shallow water environments, unlike the deep water case. Although MFP/IBF does not provide the solution in all circumstances, this paper will show that it is a promising method of shallow water beamforming in certain frequency intervals and geoacoustic environments where plane-wave beamformers are fundamentally degraded. Conversely, plane-wave beamformers, such as IBF, have enhanced performance where MFP/IBF performance is degraded. Therefore, it is natural to integrate IBF and MFP/IBF into a signal beamformer and to use each in frequency intervals and acoustic environments where their performance is not degraded. Plane-wave beamforming is covered in detail elsewhere[20–22] and is not described in general in this paper.

II. INVERSE BEAMFORMING (IBF)/PLANE-WAVE BEAMFORMING

In this paper, IBF[23–28] is a processing method which consists of three separate algorithms: a beamforming algorithm known as the Fourier integral method (FIM), a data thresholding algorithm called the eight nearest neighbor peak picker (ENNPP), and a post-processing algorithm called the M of N tracker. The Fourier integral method (FIM) is a beamforming method with 3 dB less area under its beam pattern than conventional beamforming (CBF) and thus has 3 dB more array gain than CBF for a line array. IBF is described in many recent references[23–28] and its performance as a plane-wave beamformer has shown over 10 dB of detection and tracking gain with respect to CBF. The ENNPP and M of N tracker provide the remainder of the 10 dB when used in frequency/azimuth (FRAZ) space as a plane-wave post-processing algorithm. In addition, they are described in the next section since they are used in the MFP/IBF algorithm in range/depth space in the same way they are used in FRAZ space for IBF plane-wave beamforming.

Conventional beamforming methods were developed for acoustic fields consisting of perfectly coherent plane wave signals in totally incoherent noise. These plane-wave beamformers are the ideal choice for this simplistic acoustic environment, and in general, they perform well in those environmental conditions and array configurations where the plane-wave approximation is accurate.

However, in shallow water environments, the plane-wave approximation may be poor, since correlated multiple signals (multipaths) from different normal mode arrivals are present, and the ambient noise field is far from being spatially incoherent. Plane-wave beamforming algorithms, including FIM, multiple signal classification (MUSIC),[29] and the minimum variance method,[30] have shown improved performance over conventional (CBF) time delay and phase delay beamformers. But to estimate range on a single heading in shallow water, the preferred method is one which takes advantage of these multiple arrival paths and also eliminates the need for course changes. MFP/IBF is discussed in the next section.

III. MATCHED-FIELD PROCESSING

A. Introduction

MFP determines target location by ‘‘matching’’ the measured acoustic pressure field at the hydrophone outputs to a predicted pressure field based on an assumed source location.[1–15] The predicted acoustic field at the hydrophones is determined by using a suitable acoustic propagation model, such as the generic sonar model (GSM) for deep water, or a fully coupled, normal mode model for shallow water. These predictions, which vary with source location, are compared to the measured signal and noise coherence

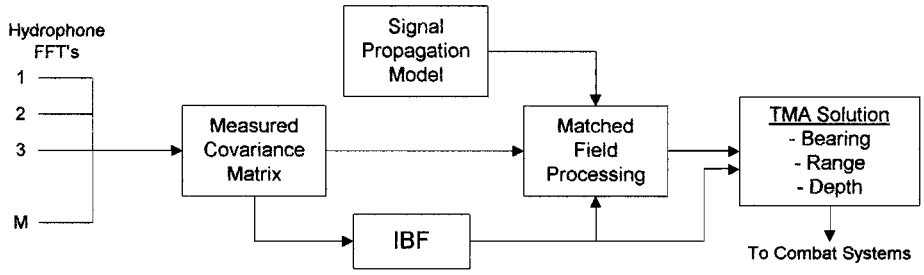


FIG. 2. IBF/MFP processing schematic.

	Depth 1	Depth 2	Depth 3
Range 1	0.2	0.1	0.2
Range 2	0.1	0.7 Peak	0.3
Range 3	0.4	0.2	0.1

FIG. 3. Eight nearest neighbor peak picker.

measurements to find the highest correlation. The basic processing schematic for the integrated MFP/IBF beamformer is shown in Fig. 2 . For optimum performance over the entire frequency range, IBF plane-wave beamforming is performed in conjunction with MFP/IBF. The target location is resolved by correlating the measured hydrophone outputs to a series of predicted values at several different range/depth cells. The correlation coefficient (or beamformer output) is calculated for each cell and plotted on a correlation surface. A spike in the correlation surface indicates the most probable target location. This method is similar to covariance matrix plane-wave beamforming, except that instead of using plane-wave steering vectors, MFP uses predicted complex pressure vectors obtained from an appropriate propagation model of the environment with an assumed source location. These predicted fields are more realistic than the plane-wave fields assumed by plane-wave beamformers which, among other things, assume a source at infinite range.

B. MFP/IBF algorithm

1. MFP/IBF beamforming

MFP utilizes steering vectors which are generated from a model with the source located at a specified range/depth cell instead of at infinite range, as is the case with plane-wave beamforming. The MFP beam pattern (or correlation output) is determined by

$$B(\theta,f,R,D,t)=\frac{1}{M^2}[\mathrm{SV}^t]\cdot[\mathrm{COV}(f)]\cdot[\mathrm{SV}], \tag{1}$$

where $B(\theta,f,R,D,t)$ is the MFP beamformer output as a function of azimuth, frequency, range, depth and time, $\mathrm{COV}(f)$ is the covariance matrix which can be conventionally weighted or FIM weighted,[26] and SV is the steering vector. The steering vectors are determined from the acoustic propagation model. The correlation output is controlled by the array geometry (which affects the covariance matrix) and the sound propagation environment (which affects the steering vectors and the covariance matrix).

2. Eight nearest neighbor peak picker (ENNPP)

As correlation coefficients are generated, the output is evaluated for the presence of relative maxima on the correlation surface. However, location of a true source is sometimes ambiguous due to high range/depth sidelobes, and when the signal level is very weak. To overcome this problem, an eight nearest neighbor peak picker (ENNPP) algorithm[23–25] is used to identify relative peaks in the correlation surface. This algorithm, which is illustrated in Fig. 3 compares each correlation coefficient with those in the eight adjacent range/depth cells, and defines a ''peak'' as a correlation coefficient which is greater than all eight of the correlation coefficients in the surrounding range/depth cells. The example in Fig. 3 shows a correlation peak located in range/depth cell (2,2).

TABLE I. M of N tracker parameters and settings.

Parameter	Setting
M of N	M/N≅0.67 with N matched to system averaging time
Range tolerance	To be determined
Range variation fix	Normally ''off''
Depth tolerance	To be determined
Depth variation fix	Normally ''on''
Peak time averaging	Usually set to M or less
Threshold	''On'' or ''off''
Zoom	

3. M of N tracker

The M of N tracker[23–25] is a three-dimensional tracker (beam level versus range, depth, and time) which operates in conjunction with the ENNPP to track persistent peaks (or relative maxima) on the FRAZ surface and reduce false target detections that do not satisfy the M of N tracker criteria. The tracker parameters include range and depth tolerances which establish a neighborhood of range/depth cells within which a correlation peak is allowed to move in a set of N time epochs. If the peak falls within the preset tolerances in ''M'' of ''N'' epochs, the successive peaks are considered a target track. There are also range and depth variation fixes which are set ''on'' (to fix the tolerance in space) or ''off'' (to allow the tolerance to move with time). Table I lists the tracker parameters. The ENNPP and M of N tracker have performed extremely well in the IBF beamformer in tracking very low SNR targets which had persistent peaks near the same FRAZ cell in M of N time epochs. The same ENNPP and M of N tracker will be used in the MFP/IBF algorithm, except that range/depth space will replace frequency/azimuth space.

The basic operation of the tracker is illustrated in Fig. 4. The range tolerance (ΔR) is established at time zero around the first peak validated by the ENNPP. Peaks which appear in successive time epochs are considered as peaks in the track only if they fall within the preset range tolerance. With the range variation fix set ''off,'' the range tolerance moves with each time epoch to center itself around the most recent peak. If two peaks appear within the range tolerance, two separate range tolerances are established in the time epoch, one around each peak, thereby generating two possible tracks. A track is designated as a target only if there are M peaks in N time epochs that satisfy the tracker settings. Since M of N is preset to 4 of 5 in Fig. 4, there must be 4 valid peaks in 5 successive time epochs to generate a track.

While targets are expected to have opening or closing range rates, depth is expected to be constant for significant periods of time. The depth variation fix is therefore set ''on'' to keep the depth tolerance fixed. In general, the depth versus time display is used as a classification tool to distinguish

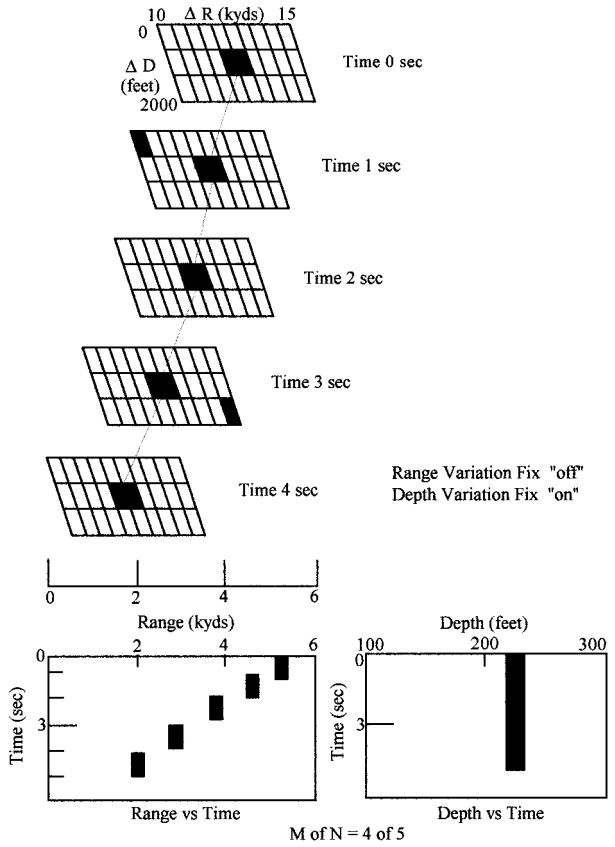


FIG. 4. M of N tracker and displays.

submerged targets from surface ships, and to alert the tracking party to a change in depth by the target vessel.

The *threshold* setting is available to further reduce false target tracks. When it is set ''on,'' peak values which fall below the preset correlation threshold are ignored by the tracker. The threshold level is determined by the operator based on the correlation characteristics of the acoustic environment.

Gray shading provides video enhancement to distinguish between ''strong'' and ''weak'' targets. Normally, tracks with high correlation values are displayed as bright white, while tracks with very low correlation values are displayed as dark gray. Color enhancement is also available to aid the operator in identifying the depth of any particular range peak or track. For example, range tracks may be designated as blue to indicate surface tracks, red to indicate submerged tracks from 100 to 300 ft in depth, yellow to indicate 300 to 500 ft, etc.

There are 14 total tracker settings and the remaining parameters not described above are used to distinguish or classify submerged from surface targets by the stability of their signal.

C. Acoustic propagation models

The complex steering vectors used in MFP must be calculated using an acoustic model which takes into account the propagation characteristics of the waveguide (i.e., normal mode model for shallow water and/or low-frequency signals, and ray theory for high-frequency signals in deep water). In this paper, the generic sonar model (GSM) or CASS (no acronym), both of which are ray models, are used in deep water and the source is assumed to be stationary during the time period over which each FFT is calculated. In shallow water, the normal mode model KRAKEN[31,32] is used to predict steering vectors assuming a moving source. Such a method was recently developed to determine the acoustic field of a moving source in a range independent environment.[33] This theory has also been extended to estimate the acoustic field from an arbitrary moving source in a range-dependent environment.[34] For the case of a slowly moving source in the horizontal plane, these results reduce to the well-known equation derived by Hawker.[35] The solution for the acoustic field of a source with an arbitrary velocity and frequency ω_0 in a range independent environment (expressed in retarded time) is given by

$$P(r,z,t) = -\frac{1}{8\pi\rho}\sum_{m=1}^{\infty}\int_{-\infty}^{+\infty} e^{-i\omega_0 t'}\,dt' \times \int_{-\infty}^{+\infty}\frac{Z_n(z)Z_n(z_s)}{(\omega-\omega_0)}e^{-i\omega(t-t')}\,\frac{dH_0(\kappa_m^0 r)}{dt'}\,d\omega. \tag{2}$$

For the case of a slowly moving source, Eq. (2) simplifies to

$$P(r,z,t) \approx \frac{i}{\rho(z_s)\sqrt{8\pi r(t')}}e^{-i(\omega t+\pi/4)} \times \sum_{m=1}^{\infty} Z_m(z_s)Z_m(z)\frac{e^{i\kappa_m r}}{\sqrt{\kappa_m^0}}, \tag{3}$$

where $t' = t - r(t')/c_m$ is the retarded time, c_m is the modal group speed, and where the superscript ''0'' indicates functions evaluated at the source frequency, ω_0.

Consider now the special case of a slow-moving source with constant horizontal velocity, as in Hawker's model. Define the angle, α, as

$$\sin\,\alpha(t) \equiv -\hat{v}\cdot\hat{r}(t), \tag{4}$$

where $\hat{v}(t)$ is the unit vector in the direction of target motion, and $\hat{r}(t)$ is the unit vector from the source to the receiver at time t. The expression for $r(t')$ now becomes

$$r(t') \approx r(t)\left[1-\frac{v\,\sin\,\alpha(t)}{c_m}\right], \tag{5}$$

where v is the speed of the source. The term $v\,\sin\,\alpha(t)$ gives the velocity component in the line of sight and corrects for the pressure wave phase shift for a moving target. Equation (5) can be substituted into Eq. (2) to obtain the acoustic field in contemporary time:

$$P(r,z,t) \approx \frac{i}{\rho(z_s)\sqrt{8\pi r(t)}}e^{-i(\omega_0 t+\pi/4)} \times \sum_{m=1}^{\infty}\frac{Z_m^0(z_s)Z_m^0(z)}{\sqrt{\kappa_m^0}}e^{i\kappa_m^0 r(t)[1-(v/c_m\sin\,\alpha(t)]}. \tag{6}$$

This is Hawker's solution[35] for the acoustic field of a slowly moving source with constant velocity. This solution will be utilized in this paper for calculation of steering vectors, but the fast-moving target case can easily be solved by using Eq. (2) if a target such as a torpedo had a high range rate.

Although the previous discussion of normal mode theory was limited to range independent environments, Lim and Ozard[34] have extended the calculations to sources moving in *weakly* range-dependent environments (i.e., environments in which the sound speed and bathymetry are weak functions of the horizontal displacement). Even though the solutions for the acoustic field in cases of arbitrary range dependence are unknown, the equations can be solved for some specific ocean geometries. In addition, Lim and Ozard developed a perturbation approach to provide a reasonably accurate model of sound propagation over the continental shelf. The mathematical details for these cases will not be discussed further in this paper, but the theoretical basis exists for extending the IBF version of MFP to fast-moving targets or weakly range-dependent environments.

D. Calculation of steering vectors

The steering vectors used for MFP in shallow water are determined using Eq. (6) to predict the acoustic pressure at each hydrophone location with the acoustic source at an assumed position. From the law of cosines, the range from the assumed source location to the nth hydrophone, $r_n(t)$, is given by

$$r_n(t)=\{r^2(t)+[(24.5-n)d]^2 + 2r(t)[24.5-n]d\cos\theta_r\}^{1/2}, \tag{7}$$

where $r(t)$ is the range from the predicted source location to the center of the array, θ_r is the relative bearing to the source, d is the hydrophone spacing, and n is the hydrophone number (1 to 48). For each range/depth cell (see Fig. 3), the acoustic pressure at the nth hydrophone can be calculated using Eq. (6):

$$P_n(r_n,z,z_s,t)\approx\frac{i}{\rho(z_s)\sqrt{8\pi r_n(t)}}e^{-i(\omega_0 t+\pi/4)} \times\sum_{m=1}^{\infty}\frac{Z_m^0(z_s)Z_m^0(z)}{\sqrt{\kappa_m^0}} \times e^{i\kappa_m^0 r_n(t)[1-(v/c_m)\sin\alpha(t)]},$$

where each value of $r_n(t)$ is calculated from Eq. (7), and the normal mode values (Z_m) are determined from the SACLANTCEN normal mode acoustic propagation model (SNAP)[37] based on inputs of $c(z)$ and water depth. The velocity of the source is assumed to be zero initially, and then updated when range rate estimates are obtained from the MFP algorithm. For the 48-element line array used to obtain the data in this paper, the steering vector for each range/depth cell is given by

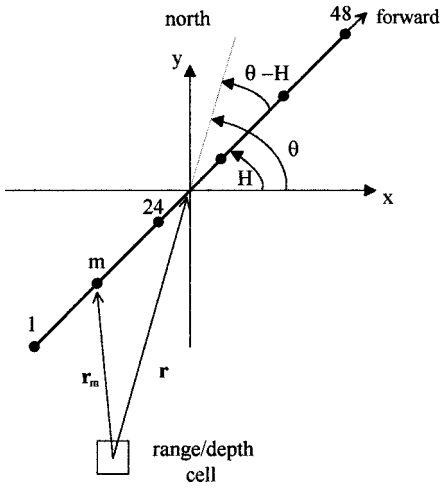


FIG. 5. Geometry for calculating complex pressures with CASS.

$$\mathrm{SV}=\begin{bmatrix} P_1(r_1,z,z_s,t) \\ P_2(r_2,z,z_s,t) \\ \vdots \\ P_{48}(r_{48},z,z_s,t)\end{bmatrix}. \tag{8}$$

Note that the variables r and z_s represent the range and depth, respectively, of each range/depth cell. Each value of r is applied to Eq. (7) to calculate the variable r_n, which, along with the values of z_s, are used to determine the acoustic pressures at the hydrophones.

E. Overview of CASS

In deep water, range-independent environments it is preferred to calculate the acoustic pressures using a ray model such as CASS.[38] In deep water, CASS is much more efficient than the normal mode model w'hich requires calculations of hundreds of mode functions for each range/depth cell. However, it is important to remember that CASS assumes the target to be stationary between FFT intervals. This may be of no consequence when dealing with slow-moving targets, but could be a significant source of error for faster targets such as torpedoes. The selection of the appropriate acoustic propagation model must therefore be based on the tradeoff of efficiency in calculations versus the accuracy of the model. To use CASS for MFP tracking of a moving source, the following assumptions must be made:

(i) the environment is range independent; and
(ii) the acoustic source does not move a significant distance in one averaging time (i.e., slow-moving target).

The principle advantage of using CASS for MFP is that the eigenrays can be specified by amplitude, conical angle, and relative phase within the CASS model.

As with the normal mode model, steering vectors are determined from the complex pressures provided by the propagation model. Figure 5 shows the geometry utilized by CASS in calculating the pressure at each hydrophone.

For any given range/depth cell, the complex pressure is given by

$$p_m(f) = \sum_{i=1}^{I_{\max}} A_i e^{i(2\pi f/c|r_m|\cos\Psi_i + \Phi_i)}, \quad (9)$$

where A_i is the amplitude, Φ_i is the relative phase, and ϕ_i is the elevation angle of the ith eigenray. The conical angle, Ψ_i is determined from the relationship

$$\cos\Psi_i = \cos(\theta - H)\cos(\phi_i). \quad (10)$$

For each eigenray below a user specified threshold of the maximum eigenray level (measured by A_i), the values of A_i, ϕ_i, and Φ_i are provided by CASS and utilized in Eq. (9) to calculate the value of $p_m(f)$. The steering vectors for an M-element, equally spaced line array are then determined from Eq. (8), and the MFP correlation, or beamformer output is calculated using

$$B(f,R,D,t) = [\mathrm{SV}^t]\cdot[\mathrm{COV}(f)]\cdot[\mathrm{SV}]/M^2 \quad \text{for CBF,} \quad (11)$$

or

$$B(f,R,D,t) = [\mathrm{SV}^t]\cdot[\mathrm{COV}^{FIM}(f)]\cdot[\mathrm{SV}]/(2m-1) \quad \text{for FIM,} \quad (12)$$

where the covariance matrix (COV) is measured at the hydrophone outputs, and the steering vector (SV) and its transpose SV^t) are modeled as a function of range, frequency, source depth, receiver depth, and sound-speed profile (SSP). Due to the nature of the covariance matrix, the MFP beamformer output (B) can be used to assess complex pressure amplitude and signal coherence simultaneously. For CBF/MFP, all elements of the covariance matrix are weighted equally, while in FIM/MFP, all diagonal and off diagonal averages are weighted equally. Since the complex pressure amplitudes are determined in part by the TL, the signal covariance matrix contains information on both TL and signal coherence, each of which affect the magnitude of the beamformer output.

It is shown in the next section that certain shallow water environments cause large increases in the covariance matrix amplitude whether or not a signal is present and is the source of high MFP range/depth sidelobes. The highest beamformer outputs will therefore occur in those cases where the combination of amplitude and signal coherence is highest for a particular shallow water environment.

The performance of MFP is tested in the following section using simulated targets to evaluate various test cases in shallow water. The targets are simulated by utilizing Hawker's formula [Eq. (6)] to calculate the appropriate acoustic pressures at the desired target range and depth, after which a signal covariance matrix is assembled from the calculated pressures. The performance of the IBF/MFP algorithm is evaluated in Sec. V along with the ENNPP and M of N tracker data.

IV. SHALLOW WATER TEST CASES

All of the test scenarios were constructed assuming a 48-element equally spaced horizontal line array. Unless otherwise noted, each case assumed a hard, reflective bottom, a constant water depth of 100 m, a source depth of 40 m, a receiver depth of 80 m, and a target relative bearing of 45°. The IBF/MFP beamforming algorithm was tested to evaluate the effects of normalization, acoustic frequency, bottom type, SSP, source and receiver depth, and target relative bearing.

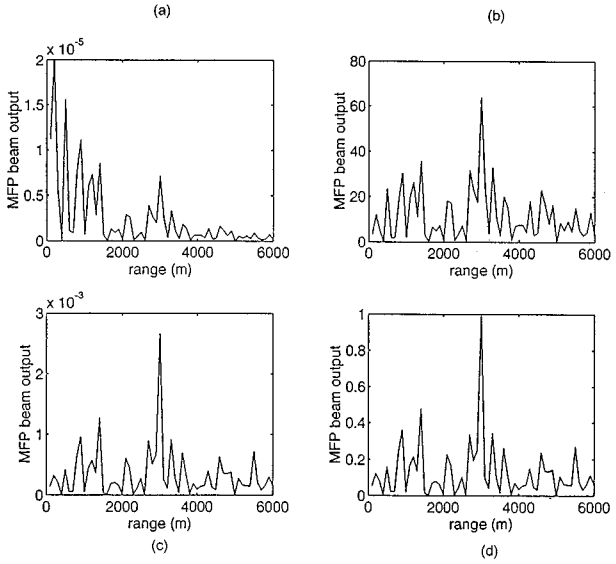


FIG. 6. Beamformer output versus range. (a) no normalization; (b)–(d) various normalization methods.

A. Normalization techniques

To effectively evaluate MFP, the MFP beamformer output must be normalized to account for the effects of cylindrical spreading. The impact of TL on the beamforming process is illustrated in Fig. 6(a), which shows a typical plot of MFP beamformer output versus range (i.e., correlation curve) at design frequency. The target for this example has a range of 3000 m. While this curve does show a moderate peak at a range of 3000 m (corresponding to the target), it also shows some larger, secondary peaks (or side lobes) at ranges less than 2000 m. These side lobes, which are undesirable because they ''mask'' the target, are a result of the cylindrical spreading term in Eq. (6) ($1/r^{1/2}$) which causes the steering vectors to have high complex pressure amplitudes at close ranges. To correct the beamformer output for these TL effects, three different methods of normalization were tested. These three methods are illustrated in Fig. 6(b)–(d) and summarized in the following sections.

1. Method 1

Each steering vector was normalized by removing the cylindrical spreading term from Eq. (6), thus removing the contribution of transmission loss to the amplitudes of the predicted pressures. As can be seen in Fig. 6(b), this method greatly reduced the side lobes at close ranges while amplifying the relative magnitude of the primary peak corresponding to the target.

2. Method 2

This method was tested as a means of adjusting the scale of the vertical axis such that changes in the beamformer outputs would result from changes in the measured covariance

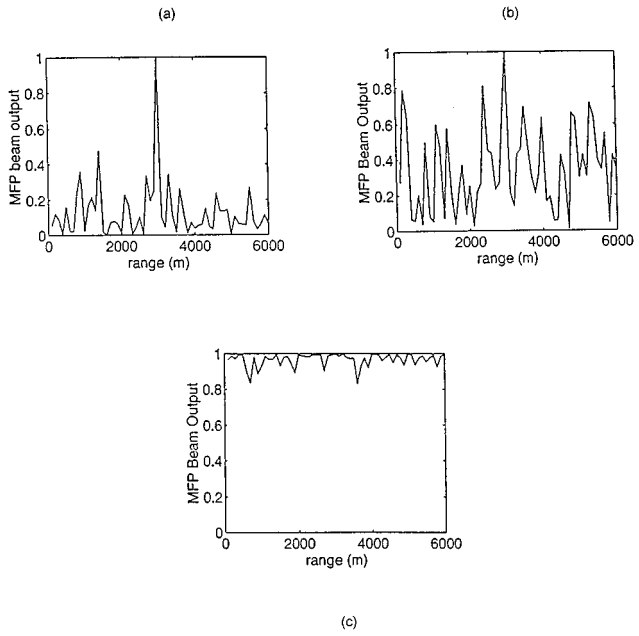


FIG. 7. MFP beamformer output versus range. (a) $f/f_0=1$; (b) $f/f_0=0.21$; (c) $f/f_0=0.08$

matrix, thus allowing for a more objective comparison of beamformer outputs at all depths. The steering vectors in this case were normalized by dividing each of the complex pressures by the average pressure magnitude for the entire vector. As Fig. 6(c) shows, the scale of the vertical axis did change, thus illustrating the power in the measured covariance matrix.

3. Method 3

The scale adjustment which was desired in method 2 was finally achieved by normalizing the signal covariance matrix in addition to the steering vector normalization of method 2. The primary peak [see Fig. 6(d)] now has a value of 1 (representing perfect correlation), while the side lobes have values between 0 and 1. This is the preferred method of normalization for synthetic data (constant covariance matrix) and is used in all of the test cases discussed in the remainder of this chapter.

B. Effects of variations from design frequency (f_0)

MFP correlation curves were plotted for acoustic frequencies of $f/f_0=1.0$, 0.21, and 0.08 as shown in Fig. 7(a)–(c), where f_0 is design frequency. Each figure shows a correlation curve with target range at 3000 m. Similar results were obtained at ranges from 2000 m to 5000 m. The $f/f_0=1.0$ plot in Fig. 7(a) shows a clearly defined primary peak corresponding to the target, and much smaller side lobes. The $f/f_0=0.21$ and 0.08 curves, however, clearly show a progressive increase in the side lobe height with decreasing frequency. This is not surprising, considering that at $f/f_0=0.08$ the acoustic wavelength is 24 times the hydrophone spacing, while at $f/f_0=1.0$ the acoustic wavelength is two times the hydrophone spacing. MFP is clearly most effective at $f/f_0=1.0$ and perhaps marginally effective at $f/f_0=0.21$. At $f/f_0=0.08$, the side lobes are so high that it is impossible to distinguish them from the peak.

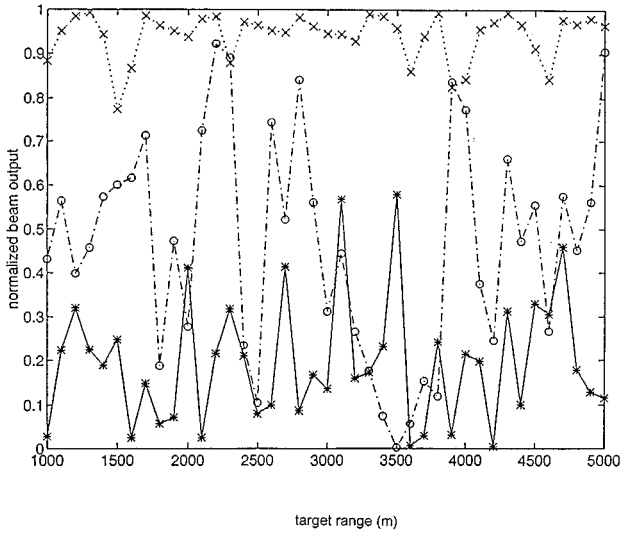


FIG. 8. Normalized beamformer output versus range. ∗: $f/f_0=1$; ○: $f/f_0=0.21$; ×:$f/f_0=0.08$.

The degradation of the plane-wave beamformer near design frequency in shallow water is illustrated in Fig. 8, which shows normalized beam outputs versus target range for the case of plane-wave steering vectors. For these curves, a beamformer output close to 1 indicates good correlation between the signal and the plane wave, while an output close to 0 indicates poor correlation with a plane wave. As the plot shows, the $f/f_0=0.08$ signals correlate very well with the plane wave, but the $f/f_0=1.0$ signals correlate very poorly. This further explains why the MFP side lobes are so high at $f/f_0=0.08$, but much smaller at higher frequencies. At $f/f_0=0.08$, the array aperture is acoustically very small and ''sees'' the acoustic signals as plane waves. The beamformer is therefore unable to resolve range, since plane-wave signal coherence does not vary as a function of range. A plane-wave beamformer in this environment would be expected to perform quite well at $f/f_0=0.08$, but the MFP beam former would perform poorly at this frequency. At $f/f_0=1.0$, the situation is reversed. Since the acoustic wavelength is much shorter, the array *is* able to resolve range and it no longer sees the acoustic signals as plane waves. At this frequency, we expect a plane-wave beamformer to perform poorly in this environment, while the MFP beamformer would perform quite well. Thus, the plane-wave version of IBF[23–28] has proven to perform exceptionally well at VLF and can supplement MFP as a beamformer in shallow water so that either MFP or IBF performs well over a wide range of frequencies.

C. Target relative bearing

Figures 9 and 10 show correlation curves at $f/f_0=1.0$ for target relative bearings ranging from 0° (endfire) to 90° (broadside). In Fig. 9, the normalized beamformer outputs are calculated using plane-wave steering vectors. The plots in this case show that the beamformer outputs are very high (greater than 0.99) at a relative bearing of 90°, but the values decrease rapidly as the relative bearings get closer to endfire.

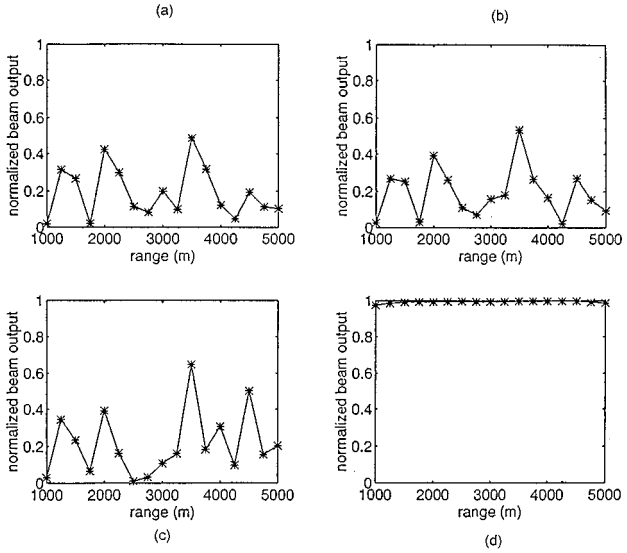


FIG. 9. Normalized plane-wave beamformer output versus target range. (a) 0°; (b) 30°; (c) 60°; (d) 90° (broadside).

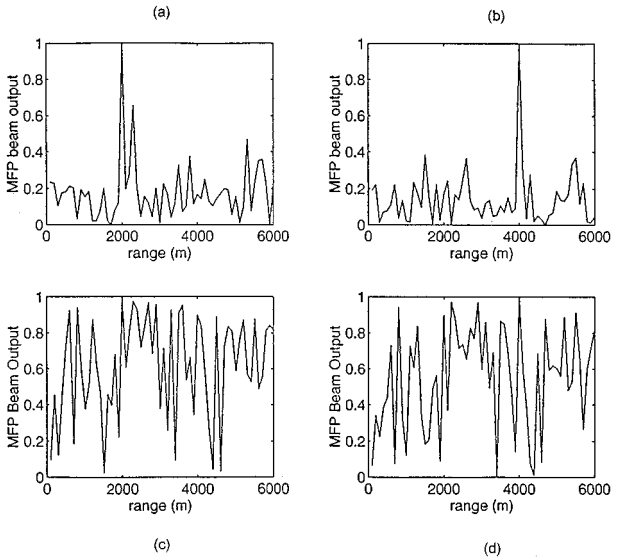


FIG. 11. MFP beamformer output versus range—hand bottom. (a) 0°; (b) 30°; (c) 60°; (d) 90° (broadside).

A plane-wave beamformer should therefore be expected to perform well only if the relative bearing is at or close to broadside. At angles closer to endfire, there is a significant degradation in plane-wave beamformer performance.

In Fig. 10, MFP beamformer outputs are plotted, again for target relative bearings ranging from endfire to broad side. These plots clearly show that the MFP beamformer, in contrast to the plane-wave beamformer, performs much better at or near endfire than it does near broadside. For relative bearings at or near broadside, MFP is not able to resolve range at all, creating so many high sidelobes that the correlation curve appears flat. On the other hand, for bearings at or near endfire, the sidelobes are much smaller.

D. Bottom type

Figures 11–13 show the contrast between a hard, reflective bottom type and a softer silt-clay bottom which absorbs much more acoustic energy. The MFP plots of Figs. 11 and 12 illustrate that for the same frequency, the MFP side lobes are significantly larger for the soft bottom than they are for the hard bottom. This is because the softer bottom type ''strips'' away some of the higher modes in the shallow water waveguide, thereby reducing the complexity and uniqueness of the modal interactions at various ranges and causing the larger range/depth sidelobes. In the hard bottom cases, the acoustic energy is comprised of more energy and more modes, causing a more unique modal interference pattern with range and thereby reducing the sidelobe level.

The effect of acoustic energy absorption by the bottom is further illustrated in Fig. 13, which shows normalized beam outputs at $f/f_0 = 1.0$ using plane-wave steering vectors. As expected, the soft bottom plot shows higher values of beam output on the average than the hard bottom plot, indicating that the acoustic signals in the soft bottom environ-

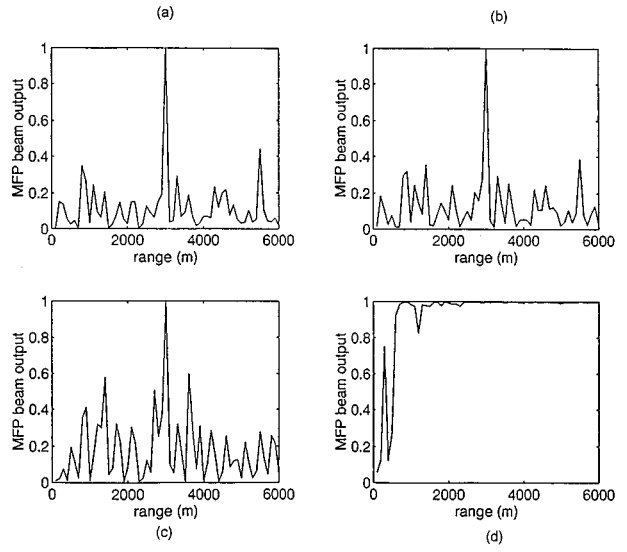


FIG. 10. MFP beamformer output versus range. (a) 0°; (b) 30°; (c) 60°; (d) 90° (broadside).

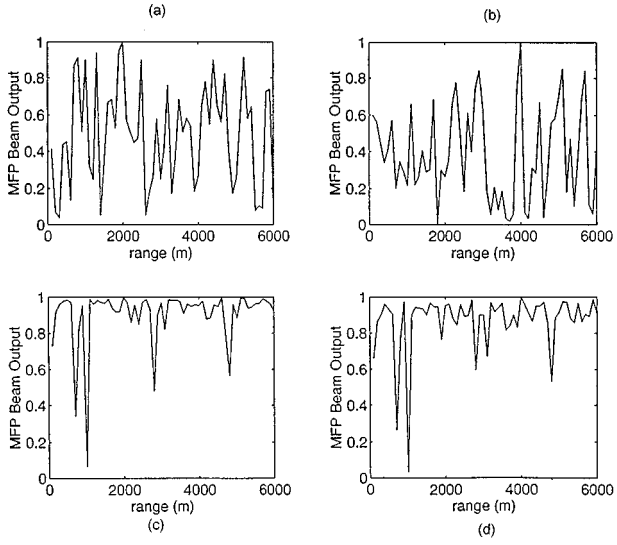


FIG. 12. MFP beamformer output versus range—soft bottom. (a) 0°; (b) 30°; (c) 60°; (d) 90° (broadside).

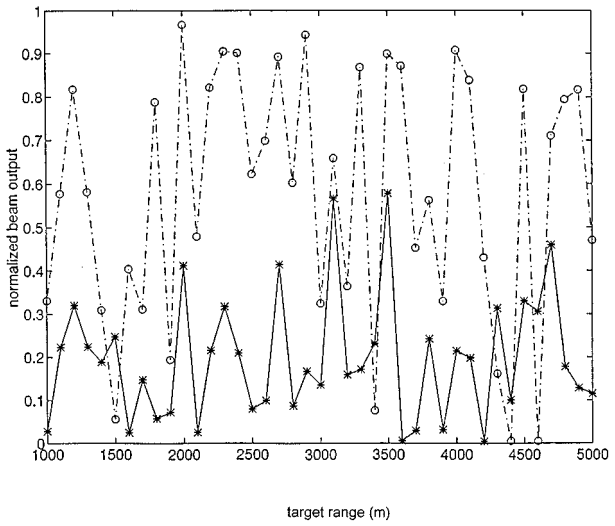


FIG. 13. Normalized beamformer output versus target range—$f/f_0=1$. (○: soft bottom; ×: hard bottom.

ment are more ''plane wave'' in nature. Again, this is due to modal ''stripping'' by the soft bottom which reduces the uniqueness of the modal interactions and causes the acoustic signals to have characteristics more like plane waves.

It should be emphasized that the geoacoustic bottom properties of real ocean environments would generally lie between the two extremes discussed above. However, in those areas where there is significant variability in geoacoustic bottom properties, passive sonar performance may vary noticeably. Specifically, an MFP beamformer would be expected to perform better in hard bottom rather than soft bottom environments, while a plane-wave beamformer would be expected to perform better where the bottom is less reflective. It is also noteworthy that at $f/f_0=1.0$, the plane-wave beamformer outputs are, at many ranges, quite low even with a soft bottom, indicating that plane-wave beamformer performance in shallow water may be questionable at higher frequencies, regardless of the bottom type.

E. SSP and source/receiver depth

The MFP algorithm was tested for various combinations of source and receiver depth with a variety of different SSPs. From this series of test cases, the effect of SSP and source and receiver depth on MFP performance is not well defined. None of the cases show a clear advantage over the other, and there is no discernible relationship between SSP and MFP sidelobe height.

V. DEEP WATER TEST CASE

A deep water data set collected in the Pacific Ocean was utilized to test the IBF version of the MFP beamforming algorithm along with the ENNPP and M of N tracker in a measured ocean environment. Although deep water acoustic propagation differs significantly from shallow water propagation, these measured data are nevertheless useful in testing the effectiveness of the ENNPP and M of N tracker in differentiating the actual acoustic source from the MFP sidelobes. This is important because as the test cases in the previous section showed, MFP sidelobes are a significant problem in MFP beamforming. This measured data set will also demonstrate the potential usefulness of the M of N tracker displays for real time shipboard applications.

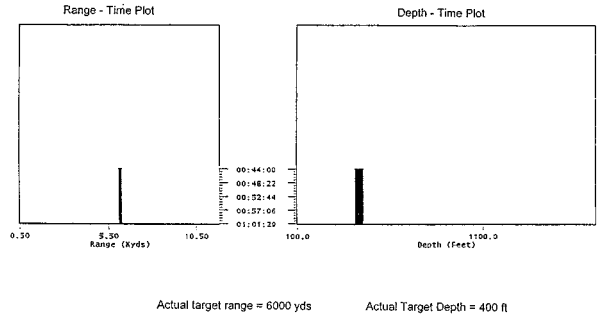


FIG. 14. M of N tracker displays for a simulated target.

The ENNPP and M of N tracker were first tested against a simulated target in a deep water environment (water depth 5000 m) with an isospeed profile and utilizing CASS as the acoustic propagation model. The covariance matrix was constructed from acoustic pressures which were calculated from the CASS eigenrays. Figure 14 shows the range versus time and depth versus time displays produced by the M of N tracker for a target at a range of 6000 yd and a depth of 400 ft. Note that both plots show a single trac
e corresponding to the target's range and depth. In this ideal situation, the tracker values correspond exactly to the actual target range and depth, and the operator can easily read the displays in real time.

Figure 15 shows the tracker displays for a 26-min portion of measured data from an at-sea experiment. This portion of data was selected because the target relative bearing varied by only 3°, thereby providing for a relatively constant bearing during the calculation of FFTs. Referring first to the range–time display, the actual target track is denoted by the solid line, while the estimated MFP track is denoted by the dotted line. As the plot shows, the MFP track differs from the actual target track by approximately 1000–1500 yd. This is an impressive result considering that CASS models the bottom in a very simplistic fashion, treating it as a single layer with a constant speed profile. The range information provided by this display would be a very useful input to track a target, especially considering the fact that range solutions obtained by conventional methods are seldom within 1000 yd of actual target range. Furthermore, the torpedo firing

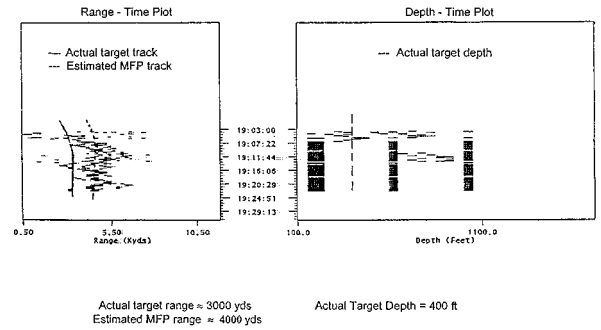


FIG. 15. M of N tracker displays for measure data.

solutions used in many successful exercise torpedo launches have had range errors of as much as 5000 yd.

The depth estimate, as expected, is not as accurate as the range estimate for this short aperture array, as illustrated by the depth versus time display in Fig. 15. While the actual target depth was 400 ft, the tracker shows three traces, none of which are any closer than 200 ft from the actual depth. This demonstrates that for this deep water environment, the MFP beamformer is much better at range discrimination than at depth discrimination for a horizontal array. The poor depth discrimination is due to poor vertical sampling by the short aperture horizontal line array.

VI. CONCLUSIONS

(1) Plane-wave beamforming with short aperture horizontal line arrays will be significantly degraded for relative bearings near endfire and away from broadside by the multipath/multimode acoustic environment of shallow water, particularly at higher frequencies. Degradation is even greater in environments with a hard, reflective bottom.

(2) MFP performance for horizontal line arrays is significantly affected by the choice of normalization technique, the geoacoustic properties of the bottom and subbottom, and the acoustic frequency (below design frequency).

(3) Range/depth sidelobes are a significant problem for short horizontal arrays at low frequencies.

(4) The IBF ENNPP and M of N tracker show great potential for reducing the impact of false targets due to the high range/depth sidelobes characteristic of MFP.

(5) MFP with horizontal line arrays shows potential as a means of single heading range solutions for shallow water environments.

[1] R. M. Hamson and R. M. Heitmeyer, ''Environmental and system effects on source localization in shallow water by the matched-field processing of a vertical array,'' J. Acoust. Soc. Am. **86**, 1915–1959 (1989).

[2] S. M. Jesus, ''Normal mode matching localization in shallow water: Environmental and system effects,'' J. Acoust. Soc. Am. **90**, 2034–2041 (1991).

[3] C. A. Zala and J. M. Ozard, ''Matched-field processing for a moving source,'' J. Acoust. Soc. Am. **92**, 403–417 (1992).

[4] H. Y. Chen and I. T. Lu, ''Matched-mode processing schemes of a moving point source,'' J. Acoust. Soc. Am. **92**, 2039–2056 (1992).

[5] J. Candy and E. Sullivan, ''Model-based passive ranging,'' J. Acoust. Soc. Am. **85**, 2472–2480 (1989).

[6] C. A. Zala and J. M. Ozard, ''Matched-field processing in a range-dependent environment,'' J. Acoust. Soc. Am. **88**, 1011–1019 (1990).

[7] T. C. Yang and C. W. Bogart, ''Matched-mode processing for sparse three-dimensional arrays,'' J. Acoust. Soc. Am. **95**, 3149–3166 (1994).

[8] A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, ''Matched-field processing: Source location in correlated noise as an optimum parameter estimation problem,'' J. Acoust. Soc. Am. **83**, 571–587 (1988).

[9] G. B. Smith, C. Feuillade, and D. R. del Balzo, ''Matched-field processing enhancement in a shallow water environment by incoherent broadband averaging,'' J. Acoust. Soc. Am. **91**, 1447–1455 (1992).

[10] D. E. Dudgeon, ''Fundamentals of digital array processing,'' Proc. IEEE **65**, 898–904 (1977).

[11] H. P. Buker, ''Use of calculated sound fields and matched-field detection to locate sound sources in shallow water,'' J. Acoust. Soc. Am. **59**, 368–373 (1976).

[12] A. Tolstoy, *Matched Field Processing for Underwater Acoustics* (World Scientific, Singapore, 1993).

[13] T. C. Yang, ''Effectiveness of mode filtering: A comparison of matched-field and matched-mode processing,'' J. Acoust. Soc. Am. **87**, 2072–2084 (1990).

[14] C. W. Bogart and T. C. Yang, ''Comparative performance of matched mode and matched-field processing in a range-dependent environment,'' J. Acoust. Soc. Am. **92**, 2051–2068 (1992).

[15] W. Bogart and T. C. Yang, ''Source localization with horizontal arrays in shallow water: Spatial sampling and effective aperture,'' J. Acoust. Soc. Am. **96**, 1677–1686 (1994).

[16] E. L. Hamilton, ''Geoacoustic modeling of the sea floor,'' J. Acoust. Soc. Am. **68**, 1313–1340 (1980).

[17] E. L. Hamilton, ''Sound velocity and related properties of marine sediments,'' J. Acoust. Soc. Am. **72**, 1891–1904 (1982).

[18] A. C. Kibblewhite, ''Attenuation of sound in marine sediments: A review with emphasis on new low frequency data,'' J. Acoust. Soc. Am. **86**, 716–738 (1989).

[19] S. Duarte, ''A comparative study of acoustic models in a shallow water range independent environment,'' Masters thesis, Naval Postgraduate School, Monterey, CA, 1994.

[20] Naval Undersea Warfare Center Detachment, New London, CT, Report TR 10003, ''Sonar Beamforming—An Overview of Its History and Status,'' by R. K. Knelpfer, 7 April 1992.

[21] W. S. Burdic, *Underwater Acoustic System Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1991), pp. 192–203.

[22] R. J. Urick, *Principles of Underwater Sound*, 3rd ed. (McGraw-Hill, New York, 1983), pp. 54–58.

[23] A. H. Nuttall and J. H. Wilson, ''Estimation of the acoustic field directionality by use of planar and volumetric arrays via the Fourier series method and the Fourier integral method,'' J. Acoust. Soc. Am. **90**, 2004–2019 (1991).

[24] J. H. Wilson, Memorandum from WAR, Inc., Tech. Rep. #1, May 1988.

[25] J. H. Wilson and P. Schey, ''Beamforming and Spatial Prewhitening using the Fourier Series Method,'' Tech. Rep. #81, Polar Research Lab, Inc., 13 January 1988.

[26] J. H. Wilson, ''Applications of inverse beamforming (IBF),'' J. Acoust. Soc. Am. **98**, 3250–3261 (1995).

[27] J. Paquin Fabre and J. H. Wilson, ''Minimum detectable level (MDL) evaluation of inverse beamforming (IBF) using outpost sunrise data,'' J. Acoust. Soc. Am. **98**, 3262–3278 (1995).

[28] J. H. Wilson, ''Signal Detection and Localization Using the Fourier Series Method (FSM) and Cross-Sensor Data,'' J. Acoust. Soc. Am. **73**, 1648–1656 (1983).

[29] R. O. Schmidt, ''Multiple emitter location and signal parameter estimation,'' IEEE Trans. Antennas Propag. **AP-34**(3), (1986).

[30] J. Capon, ''High-resolution frequency-wave number spectrum analysis,'' Proc. IEEE **57**, 1408–1418 (1969).

[31] M. B. Porter and E. L. Reiss, ''A numerical method for bottom interacting ocean acoustic normal modes,'' J. Acoust. Soc. Am. **77**, 1760–1767 (1985).

[32] M. B. Porter, ''The KRAKEN normal mode program,'' Naval Research Laboratory Memorandum Report 5120-92-6920, Chap. 2, 1992.

[33] P. H. Lim and J. M. Ozard, ''On the underwater acoustic field of a moving point source. I. Range-independent environment,'' J. Acoust. Soc. Am. **95**, 131–137 (1994).

[34] P. H. Lim and J. M. Ozard, ''On the underwater acoustic field of a moving point source. II. Range-dependent environment,'' J. Acoust. Soc. Am. **95**, 138–151 (1994).

[35] K. E. Hawker, ''A normal mode theory of acoustic Doppler effects in the ocean waveguide,'' J. Acoust. Soc. Am. **65**, 675–681 (1979).

[36] F. B. Jensen and M. C. Ferla, ''SNAP: The SACLANTCEN NORMAL-MODE ACOUSTIC PROPAGATION MODEL,'' SACLANT ASW Research Centre, San Bartolomeo, Italy, 1979.

[37] Naval Undersea Warfare Center Detachment, New London, CT, ''CASS Sonar Manual,'' in preparation.

Development of distortion product emissions in the gerbil: "Filter" response and signal delay

David M. Mills[a] and Edwin W Rubel
Virginia Merrill Bloedel Hearing Research Center, University of Washington, Seattle, Washington 98195

(Received 3 April 1996; revised 20 August 1996; accepted 10 September 1996)

Amplitude and phase responses of distortion product otoacoustic emissions as a function of stimulus frequency ratio were measured for frequencies between 2 and 48 kHz, in Mongolian gerbils (*Meriones unguiculates*) aged 15 to 30 days after birth. After baseline measurements, furosemide was administered to distinguish active from passive emissions. At all ages, structure in the form of multiple peaks was observed in the amplitude responses of specific odd-order emissions. This structure depended on the *emission* frequency, not the stimulus frequency ratio, and did not generally depend on the stimulus amplitude. Nor was it dependent on the functioning of the cochlear amplifier: At moderate stimulus levels, the observed emission distribution simply shifted to lower amplitudes when the cochlear amplifier was made temporarily dysfunctional by furosemide injection. The center frequencies and widths of the peaks in the amplitude response did not generally change with age, except that the relative amplitudes of the higher-frequency peaks were increased in younger animals. At 2 kHz, however, the distribution showed other evidence of maturation, with the frequency of maximum emission moving downward with age. The phase responses yielded estimates of the round trip signal (group or traveling wave) delay. At a given frequency, the active signal delay typically decreased substantially with increasing stimulus level. However, there was a rapid variation in delay as the stimulus level passed the normal active–passive crossover level. At stimulus levels measured *relative* to the active–passive crossover level, i.e., either 20 or 30 dB lower, the active signal delay decreased only slightly with age. Overall, both filter response and signal delay characteristics were found to be essentially mature near the onset of hearing. [S0001-4966(97)04601-8]

PACS numbers: 43.64.Jb, 43.64.Kc [RDF]

INTRODUCTION

The variation of distortion product emissions as a function of stimulus frequency ratio has been important in the investigation of cochlear mechanics. There have been two main applications, one associated with the interpretation of the amplitude response of the emissions, the other with the interpretation of the phase angle response. Both applications have obvious relevance to the study of the mechanics in the developing cochlea.

It is well known that there is structure in the amplitude response of specific odd-order emissions when the stimulus frequency ratio is varied (Brown and Gaskill, 1990b; Gaskill and Brown, 1990). This structure appears to be associated with the *emission* frequency, not with the stimulus frequency ratio itself (Brown and Gaskill, 1990a). That is, for constant stimulus levels the amplitudes of these emissions appear to peak when the emission frequency is about one half-octave below the higher-frequency stimulus (Brown and Gaskill, 1990b; Brown *et al.*, 1992). This discovery has been used to support the idea that there is a ''second filter'' in the cochlea, which filters the emissions *after* generation (Brown and Williams, 1993). Specific micromechanical structures have been proposed to account for the filtering (Allen and Fahey, 1993a, b).

It seems useful to extend these investigations into the developing mammalian cochlea, and to more completely characterize the emission amplitude response as a function of stimulus amplitude, frequency, and age of the animal. If the filtering is due to specific micromechanical structures, the characteristics of this filtering may change as these structures develop. Overall, emission amplitude responses should be related to developmental changes in the interaction of waves of different frequencies in the cochlea. These interactions must depend, for example, on the extent of the region of active amplification along the basilar membrane (BM), and on the sharpness of the peak BM response.

Separate from the amplitude response, the phase angle response has primarily been employed as a measure of the round trip travel time. This time has been variously called the group, signal, or traveling wave delay (Brown and Kemp, 1985; Kimberley *et al.*, 1993; Plonsey and Collin, 1961). Contrasted with the phase velocity, the signal velocity is the speed associated with energy or information transmission in a medium (Plonsey and Collin, 1961). For distortion product emission measurements, the signal delay specifically is the round trip time associated with a stimulus signal passing the microphone and going into the cochlea, traveling down the BM, generating an emission, and the emission traveling back out of the cochlea and being detected by the microphone. With certain assumptions, the round trip signal delay time can be related to the measured change in phase in the emission at the microphone location relative to that of the stimu-

[a]Electronic mail: dmmills@u.washington.edu

lus, the phase changes occurring when the ratio of the stimulus frequencies is changed.

The major difficulty with the interpretation of the signal delay as measured by emissions is the fact that the derived signal delays depend strongly on stimulus level, typically decreasing with increasing stimulus level (Brown and Kemp, 1985). It is difficult to know how to relate the wide span in measured values to physical distances in the cochlea. It has been suggested that the *differences* between signal delays at the same stimulus intensities but at different frequencies represent a valid comparison between traveling wave delays at different frequencies (Kimberley *et al.*, 1993). However, the situation cannot be resolved so easily. Because of possible differences in passive conduction into the cochlea at different frequencies and ages, the equivalent stimulus levels in the cochlea at different frequencies cannot be *assumed* to occur with equal stimulus levels in the ear canal. There is, after all, no *a priori* reason to chose ''equal'' stimulus levels at any particular point and using any particular measure. That is, the quantity to be ''equalized'' at different frequencies or ages could equally well be stimulus pressure, displacement, volume velocity, or energy flow (Keefe *et al.*, 1993), and the point of equalization could be any point along the input transmission path to the cochlea.

This is a particularly vexing problem in interpreting changes in signal delay during development, given the known changes in passive conductance into the mammalian cochlea during development (Mills *et al.*, 1994; Mills and Rubel, 1996). For example, there have been two recent measurements of the signal delays in humans, comparing term infants and adults (A. M. Brown *et al.*, 1994; D. Brown *et al.*, 1994). Both measured delays at about the same, constant stimulus pressure levels in the ear canal. At mid-frequencies, the first study found a decrease in the mean delay with age and the second study an increase; the reasons for the difference are not known.

At this point, it seems important to establish valid procedures for the measurement of delays using experimental animals; for these animals distortion product emissions are much easier to measure than for human infants and the results can be compared to independent measures of cochlear travel time.

I. METHODS

A. Animal preparation

Young gerbils (*Meriones unguiculates*) were obtained from breeding pairs maintained in our colony, originally purchased from a commercial supplier (Tumblebrook Farms, Brookfield, MA). Pairs were checked for births daily, and the date the birth was first observed was denoted 0 days after birth (dab). All animal preparation and recording were performed in an I.A.C. double walled acoustic booth. Animals were initially anesthetized with a subcutaneous injection of a mixture of ketamine hydrochloride (Ketaset: 15 mg/kg) and xylazine (Rompun: 5 mg/kg). A surgical depth of anesthesia was maintained by subsequent injections, as needed, of either ketamine alone or the ketamine-xylazine mixture, at about one half the initial dosage. The pinna, surrounding skin, and outer third of the ear canal were removed on the left side, along with much of the scalp. The skull was attached to a head holder with cryanoacrylic adhesive (Borden), and a thermocouple placed in the rectum. An automatic heating pad kept the internal temperature at 36–37 °C. Tissue was removed over the bulla immediately posterior to the ear canal, and a hole (1–2 mm diameter) drilled into the bulla to equalize static pressure in the outer and middle ears. The bulla hole was left open, to improve low-frequency sound conduction through the middle ear in these young animals (Cohen *et al.*, 1993; Mills *et al.*, 1994).

B. Equipment and procedures

The equipment and basic procedures employed for the high-frequency distortion product measurements were the same as previously reported (Mills and Rubel, 1996). Briefly, a coupler was employed which incorporated two microphones and two sound delivery tubes connected to a 2-mm diameter central cavity. One microphone was a 1/4-in. high-frequency microphone (Larson & Davis 2530) calibrated as a probe microphone, with the probe tip located at the coupler opening. The other was a removable ER-10B low noise microphone (Etymotic). Using a micromanipulator and operating microscope, the coupler was joined to the ear canal opening. A wide band noise signal was introduced into the ear canal through one of the sound delivery tubes, and the output of the probe microphone used to calibrate the sound delivery system and the ER-10B microphone response *in situ*, to 50 kHz.

A set of measurements of normal emissions, taking 1–2 h, was then completed. For each single emission measurement, two tones (frequencies f_1 and f_2) were introduced into the ear canal through the two coupler tubes, and the ER-10B microphone output synchronously averaged, typically for 4 s. The lower-frequency stimulus amplitude, L_1, was always 10 dB higher than the higher-frequency stimulus amplitude, L_2 (Mills *et al.*, 1993; Whitehead *et al.*, 1995a, b). For this report, the basic measurement sequence involved a sequence of stimulus frequency ratios. The upper frequency, f_2, and both amplitudes of the two stimulus tones were fixed for these sequences, while the lower stimulus frequency, f_1, was incremented after each single measurement. For the initial data set for all animals, the basic sequence included measurements at 17 frequency ratios from $f_1/f_2=0.714$ to 0.925. There were additional measurements for some animals at some frequencies, particularly including lower ratios for $f_2=2$ kHz. The basic frequency ratio sequence was conducted for a number of different stimulus levels, starting near the noise floor for the cubic distortion tone (CDT, $2f_1-f_2$) emission. Stimulus levels were incremented in 10-dB steps, to a maximum of 80–100 dB SPL depending on age and f_2 frequency. This set of measurements was made for each of the following f_2 frequencies: 2, 8, 16, 32, and 48 kHz. Note that only for the older animals could emissions be detected above the noise floor for the higher frequencies (Mills and Rubel, 1996).

After the initial data set was completed, the animal was given an intraperitoneal (I.P.) furosemide injection to establish the ''passive'' emission response. Dosages were set on the basis of previous experiments (Mills *et al.*, 1993, 1994;

TABLE I. Number of animals in experimental groups. Age of each group is given in days after birth (dab) with the furosemide dosage used for that age group.

Age (dab)	Number	Furosemide (mg/kg)
15	6	60
17	6	80
20	6	100
25	5	150
30	4	200
42–46	3	300

Mills and Rubel, 1996) so that the cochlear amplifier would be rendered essentially nonfunctional for a brief time, usually about 5–7 min. The required dosages varied with age as noted in Table I. During the time that the cochlear amplifier was nonfunctional, a brief series of frequency ratio sequences was taken. Because of time constraints, these were limited to five or six different frequency ratios, and at stimulus amplitudes near the upper limits. The passive measurements were conducted for the same f_2 frequencies as the initial, active measurements, and 10-dB intervals were also used for the stimulus amplitudes. For the younger animals (up to 20 dab) the passive emissions were so weak that additional filtering was required to successfully detect them. A programmable low-pass filter was employed on the ER-10B output, to reduce the magnitude of the stimulus frequency components in the microphone signal, as was done previously (Mills and Rubel, 1996). This was successful in enabling detection of the amplitude of the weak passive signals but the phase response was not then usable, due to the additional phase shifts caused by the filtering.

Each animal was monitored for one half hour after injection to assure that there was at least partial recovery of emissions during this time. The total number of animals measured in each age group is given in Table I.

Procedures for the care and use of the animals reported on in this study were approved by the University of Washington Animal Care Committee (*re:* Grant No. NIH DC 00395, Ontogeny of Sensory Processes).

C. Data analysis: Signal delay

For this report, the variation in phase angle with stimulus frequency ratio (f_1/f_2) was generally converted into an equivalent ''signal delay.'' For an emission at the frequency given by $f_{mn}=(mf_1-nf_2)$, the round trip signal delays associated with that emission, T_{mn}, have been determined using the equation (Kimberley *et al.*, 1993; Mahoney and Kemp, 1995)

$$T_{mn}=-\Delta\Phi_{mn}/\Delta f_{mn}, \tag{1}$$

where the *corrected* emission phase angle, relative to the stimuli phases, is given by $\Phi_{mn}=\phi_{mn}-(m\phi_1-n\phi_2)$. The quantity ϕ_{mn} is the measured emission phase angle, and ϕ_1 and ϕ_2 the measured phase angles of the two stimuli at f_1 and f_2. All emission phase angles reported here are corrected emission phase angles.

The procedure employed has been to include a limited range of f_1 frequencies, typically 5–7 points, located around the amplitude maximum and spaced closely enough that the typical change in emission angle between neighboring ratios was securely less than 180°. The emission angles were then unwrapped by a computer program and the corrected, unwrapped phase angles were calculated and displayed. A linear, least-squares fit was then made to these points, and the round trip signal delay was calculated from the slope of this line, according to Eq. (1).

II. RESULTS

A. Signal delay, active emissions

We operationally define active emissions as those emissions which are emitted by a normal mammalian cochlea and which are essentially eliminated when the endocochlear potential is sharply reduced (Mills *et al.*, 1994; Norton *et al.*, 1991; Norton and Rubel, 1990). In practice, this means that these are certain odd-order emissions found at relatively low stimulus levels, although care must be employed in defining what stimulus levels are ''low'' for a given frequency, species, and developmental stage (Mills *et al.*, 1994). Representative results for amplitude and phase responses of active emissions at the cubic distortion tone (CDT, $2f_1-f_2$) frequency are shown in Fig. 1, for an adult gerbil. For each panel, the f_2 frequency and the stimulus levels L_1 and L_2 were kept fixed, and the f_1 frequency stepped in small increments. The amplitude of the emission is shown in the lower panel of each pair, and the phase angle response in the upper panel. Note that the horizontal axis in these figures is the ratio f_1/f_2. This is the most useful form for the axis for experiments in which f_2 is kept fixed, and f_1 varied. That is, a simple group delay leads to a linear relationship between phase and the f_1 frequency, or between phase and the ratio f_1/f_2.

The responses shown are typical of adult gerbils at relatively low stimulus levels. The amplitude responses had maxima at f_1/f_2 frequency ratios near 0.75–0.8, and the phase response was approximately linear across these middle f_1/f_2 ratios. There appeared to be a trend for the frequency ratio at maximum amplitude to move to lower ratios, for lower f_2 frequencies. This trend appeared stronger in younger animals, as the representative example of a 15 dab gerbil in Fig. 2 illustrates. Emissions are shown in Fig. 2 only for $f_2=2$ and 8 kHz, as adequate phase responses for stimuli with $f_2=16$ kHz and above were not obtainable in the 15 dab neonates, due to emission amplitudes being at or below the noise floor.

The range of frequency ratios chosen for the estimate of the signal delay in this experiment is also illustrated in Figs. 1 and 2, by the filled-in phase data points in the upper panels, and the horizontal bars in the lower panels. This range was chosen, on the basis of responses such as shown in Figs. 1 and 2, to be the best compromise across the range of f_2 frequencies from 2 to 32 kHz. This range, from $f_1/f_2=0.746$ to 0.82, includes the regions where the amplitude is maximum for nearly all ages, the exception being the younger animals for $f_2=2$ kHz. After the phase response was un-

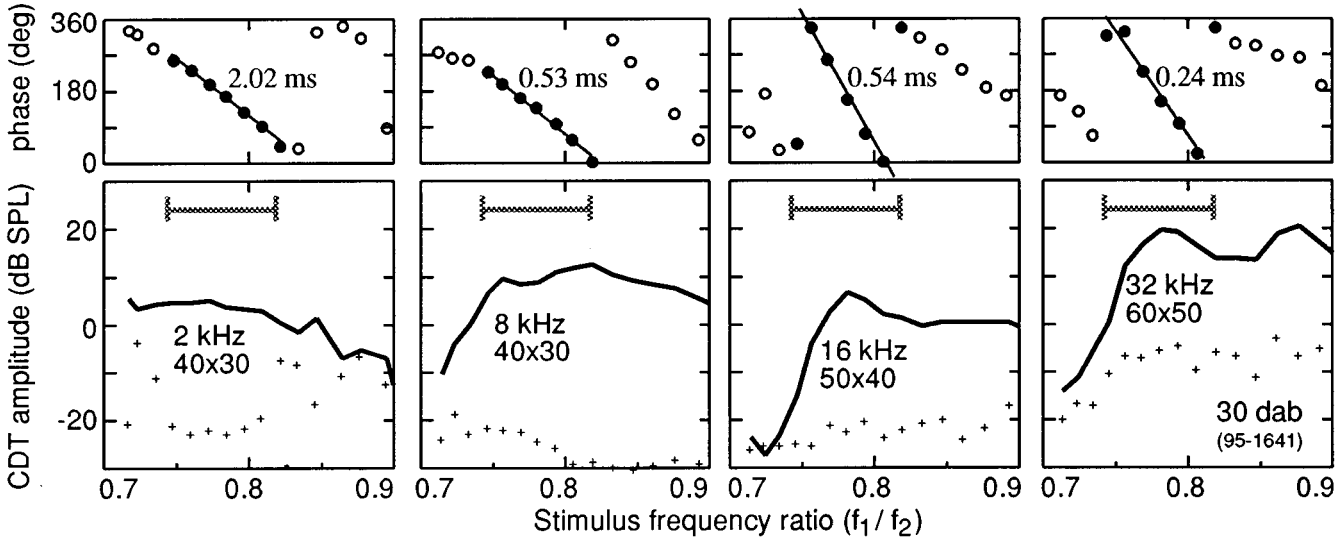


FIG. 1. Individual data: Amplitude and phase for the cubic distortion product (CDT, $2f_1-f_2$) emission are presented as a function of frequency ratio, for representative responses in a 30-day-old animal at low stimulus levels. The parameters listed in each of the lower panels are the upper stimulus frequency, f_2, and the stimulus levels, $L_1\times L_2$ (dB SPL). The crosses in the lower panels indicate the measured noise levels. The horizontal bars and the filled-in phase symbols indicate the phase responses that were used to calculate the round trip signal delay to be associated with that f_2 frequency and stimulus level pair. These are the seven measurements taken from $\phi_1/\phi_2=0.746$ to 0.820 in approximately 0.012 increments. The CDT phase angles listed here and elsewhere are those referenced to the measured stimulus phases, that is, by subtracting the quantity $(2\phi_1-\phi_2)$ from the measured CDT phase, where ϕ_1 is the measured phase angle of the $f1$ stimulus at the microphone, and ϕ_2 the phase angle for the f_2 stimulus. The phase zero reference in each panel is arbitrary. The least-squares best linear fit to the 7 phase points are shown by the lines, the figures give the derived signal delays in milliseconds (ms). Note that the phase angles are shown folded into a single 360° interval here, but were unwrapped before fitting the straight line shown.

wrapped, least-squares linear fits to the phase variation with frequency were made, as illustrated in Figs. 1 and 2. Values of signal delay from Eq. (1) corresponding to the slopes shown are also listed. These values are typical of the range found in these experiments, from 2 to 3 ms for $f_2=2$ kHz down to 0.2 ms for $f_2=32$ kHz. Note that, in contrast to the results for humans (Kimberley *et al.*, 1993), the phase angle response in gerbils was not always linear with frequency ratio, particularly outside the region of the maximum amplitude for the emission (Fig. 1).

The signal delays found varied not only with f_2 frequency, but with stimulus amplitude, $L_1\times L_2$, at a given f_2 frequency. Figure 3 presents the observed variation for the same two animals as in Figs. 1 and 2. Stimulus levels were incremented in 10-dB steps, with L_2 always 10 dB below L_1. As in these two examples, the derived signal delay typically decreased by a factor of 2 as the stimulus level increased

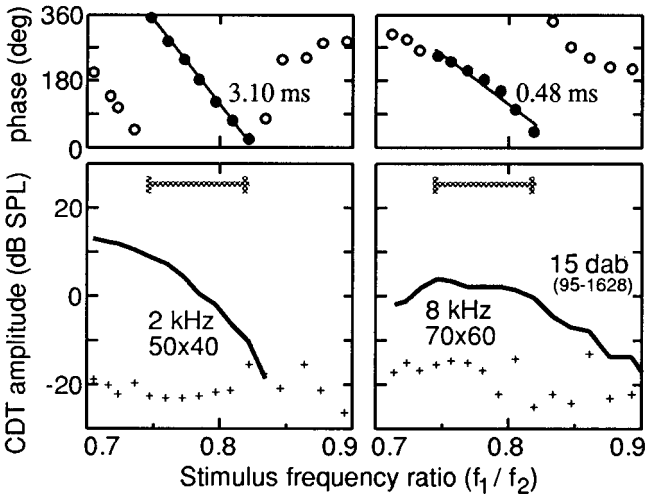


FIG. 2. Individual data: Amplitude and phase of CDT emissions are plotted versus frequency ratio, for representative responses in a 15-day-old animal at relatively low stimulus levels. Same conventions as Fig. 1. For these younger animals, adequate phase determinations for stimulus frequencies $f_2=16$ kHz and above were not obtainable.

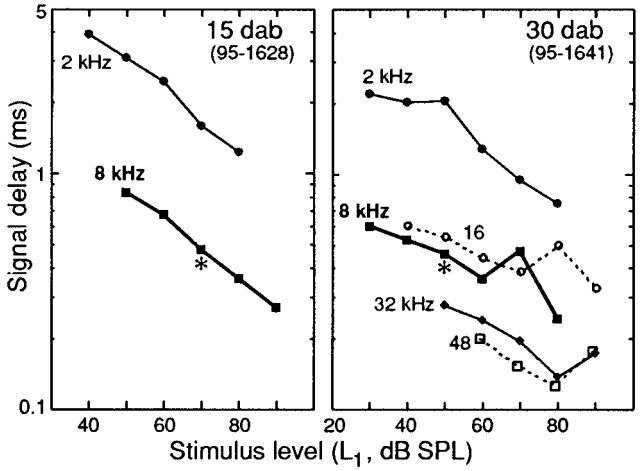


FIG. 3. Individual data: Round trip signal delay is presented as a function of stimulus level for the same individual animals presented in Figs. 1 and 2. The parameter listed is the higher stimulus frequency, f_2. Signal delays were calculated using the phase responses noted in Figs. 1 and 2. The asterisks denote ''equivalent'' stimulus levels for $f_2=8$ kHz (see text).

from 40 to 80 dB SPL. Further, there was no obvious tendency for the derived signal delay to reach an asymptote as lower or higher stimulus levels were approached. Finally, at higher stimulus levels were typically unexpected, systematic variations in the derived signal delays. For example, note the rapid variation in the derived signal delays for the 30 dab animal in Fig. 3, occurring between L_1=60 and 80 dB SPL for f_2=8 kHz, and between 70 and 90 dB SPL for f_2=16 kHz. These rapid minima and maxima were quite consistently found at higher stimulus levels, and the total variation in signal delay typically spanned about an octave.

These intrinsic variations with intensity make direct developmental comparisons very difficult. For example, the 8 kHz delay was typically larger for the 15 dab animal than it was for the 30 dab animal. This would seem to imply a *decrease* in signal delay at 8 kHz during development. However, it is known that higher stimulus levels are generally required in the ear canal at 15 dab compared to 30 dab to provide the same, equivalent levels in the cochlea at 8 kHz (Mills *et al.*, 1994; Mills and Rubel, 1994). Suppose this were a 20-dB difference, which is typical. The delay at 15 dab for, say, L_1=70 dB SPL should then be compared with the delay at 30 dab with L_1=50 dB SPL. These ''equivalent'' locations are marked with an asterisk in Fig. 3. It can be seen that, rather than showing a marked decrease with maturation, the comparison of such equivalent levels would imply that there was little change with maturation. There is a similar potential problem in comparisons between different f_2 frequencies at the same age, due to frequency variations in input transmission. For example, for the individual animal in Fig. 3, the 16-kHz delays were larger than the 8 kHz at the same levels (SPL). If middle ear effects accounted for a 20-dB difference at 16 kHz, however, one might correct the 16-kHz curve by shifting it to the left by 20 dB. The delays at 16 kHz at these ''equivalent'' levels would then be *smaller* than at 8 kHz.

In order to properly compare results, the stimulus levels employed for each animal were *normalized* by referencing them to the crossover stimulus level, L_X. This has been defined to be the stimulus level in the ear canal at which the transition from primarily active response to primarily passive response occurs in the cochlea (Mills *et al.*, 1994; Mills and Rubel, 1994). Values of L_X were estimated for each animal at each f_2 frequency as illustrated in Fig. 4.

The pre- and post-injection growth functions were obtained (with 10-dB stimulus intervals) for the three f_1/f_2 ratios noted, and were plotted as shown. The crossover stimulus level, L_x (and the shift, ΔC, used later) were calculated for each ratio as illustrated, and averaged over the three f_1/f_2 ratios.

The results in Fig. 3 can now be replotted in Fig. 5, with the stimulus level at each f_2 frequency now plotted relative to its crossover level, L_x, determined for that animal and f_2 frequency. The estimated ''threshold'' crossover levels for 2 kHz changed little between the 15 and 30 dab animals. This lack of improvement with age can be attributed to the open bulla condition of these experiments: In closed bulla conditions, the smaller bulla of younger animals typically causes an increase in the threshold measure, L_x, for frequencies

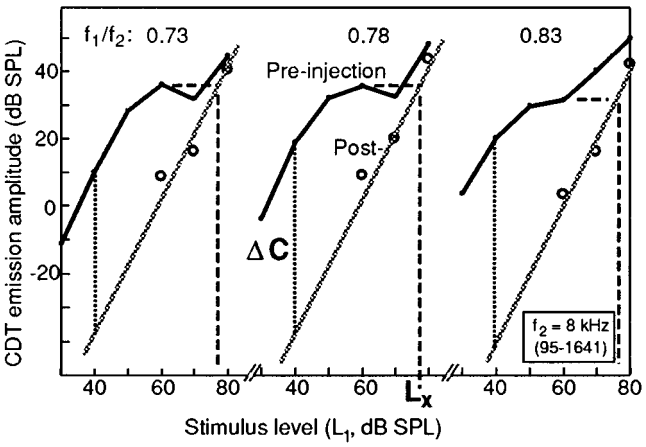


FIG. 4. Individual input–output, or ''growth functions,'' i.e., variation of cubic distortion tone (CDT, $2f_1-f_2$) emission amplitude with stimulus level. The solid lines are the normal, pre-injection growth functions. The post-injection responses, taken at the time of the ''flat minimum'' (Mills and Rubel, 1994), are indicated by the open circles, with the ''best-fit'' straight line with a slope of 2:1 shown. The three growth functions are all for f_2=8 kHz, but different ratios of stimulus frequencies, f_1 and f_2, as noted. Each of these growth functions yields an estimate of the active–passive transition level, L_x, and the shift in emission amplitude at low signal levels, labeled ΔC. The shift, ΔC, is the difference between the CDT emissions with and without a functional cochlear amplifier, and is therefore related to the gain of the cochlear amplifier. For the example shown, L_x was 77 dB SPL for all three ratios, and the average shift, ΔC, was 55 dB.

below 4 kHz (Mills *et al.*, 1994). At 8 kHz, there was a decrease of nearly 20 dB for the older animal compared to the 15 dab animal.

The relative maxima and minima seen in the 30 dab animal in Fig. 3, which did not occur at the same absolute levels, are now observed in Fig. 5 at approximately the same levels *relative* to the crossover stimulus level. These rapid variations occurred largely in the 10-dB range for which the stimulus levels L_1 were 10 dB below the crossover level, L_x,

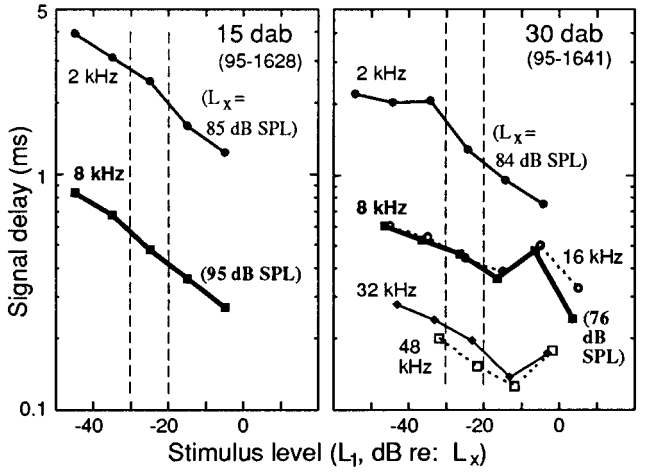


FIG. 5. Individual data: Round trip signal delay is plotted as a function of *relative* stimulus level. This is the same data as in Fig. 3, but now the stimulus levels are plotted relative to the active–passive transition levels, L_x. These were determined for each animal at each f_2 frequency using growth function plots as illustrated in Fig. 4. The estimated values of L_x are shown, in parentheses, for f_2=2 and 8 kHz. The vertical dashed lines simply note the relative levels 20 and 30 dB below L_x.

and equal to it (i.e., between $L_1=L_x-10$ dB and $L_1=L_x$). It seems likely, therefore, that the rapid variation observed was associated with the transition between ''active'' and ''passive'' responses. That is, the rapid variation could be attributed to a phase interaction between these two components. This interpretation was further supported by the observation that, among individual animals of the same age, the rapid variation always occurred near the transition level L_x, but the form varied, sometimes dominated by a very sharp minimum, and at other times there was a more moderate maximum (i.e., like that shown in Figs. 3 and 5).

It seems obvious that, to study the developmental variation of the signal delay in the normal, *active* cochlea, this active–passive transition region should be avoided. For this report, we have therefore chosen to report results in signal delay primarily for the stimulus level which is 30 dB below L_x. This level is far enough from the active–passive transition that it seems little affected by it, yet the stimulus level is typically still high enough to yield adequate signal to noise ratios. The level 20 dB below L_x would work as well; we have made all the calculations presented below for this case as well, and the results are virtually identical.

In Fig. 6(A) and (B) we present the variation with age for the mean signal delay, measured at the stimulus level estimated to be ''constant'' in the cochlea. That is, for each individual animal the signal delay was estimated at a stimulus level 30 dB below the crossover level, L_x, as in Fig. 6. The estimated delays were then averaged across animals within an age group. For comparison, in Fig. 6(C) we present the mean delays measured at a constant stimulus level as measured in the ear canal, in this case for $L_1\times L_2=50\times 40$ dB SPL.

The overall trend is obvious: there was a modest decrease in signal delay between 15 and 20 dab, at all f_2 frequencies. The decrease was not as large, however, when comparison was made with stimulus levels at a constant offset to characteristic levels in the cochlea [Fig. 6(A) and (B)] compared to delays measured at the same absolute levels in the ear canal [Fig. 6(C)]. The reason is that the *relative* stimulus levels corresponded to higher ear canal stimulus levels in the younger animals than the older (except at 2 kHz where there was little change with age). Note that there were no measurable emissions available for stimulus levels of 50 dB SPL for younger animals at higher frequencies, because of the increase in passive threshold (Mills *et al.*, 1994).

Note also in Fig. 6(A) that there was an interesting minimum in the derived signal delay that occurred early in development for most f_2 frequencies. This minimum was observed first at 2 kHz, at about 17 dab, and then at 8 kHz at about 20 dab, and finally at 32 kHz at about 25 dab. In all cases, the minimum occurred several days to a week after the emissions first become measurable at the corresponding frequency.

The dashed line in Fig. 6(A) displays the difference between the mean signal delay for stimulus frequency $f_2=2$ kHz and that for $f_2=8$ kHz. This difference is obviously dominated by the variations in the signal delay for $f_2=2$ kHz.

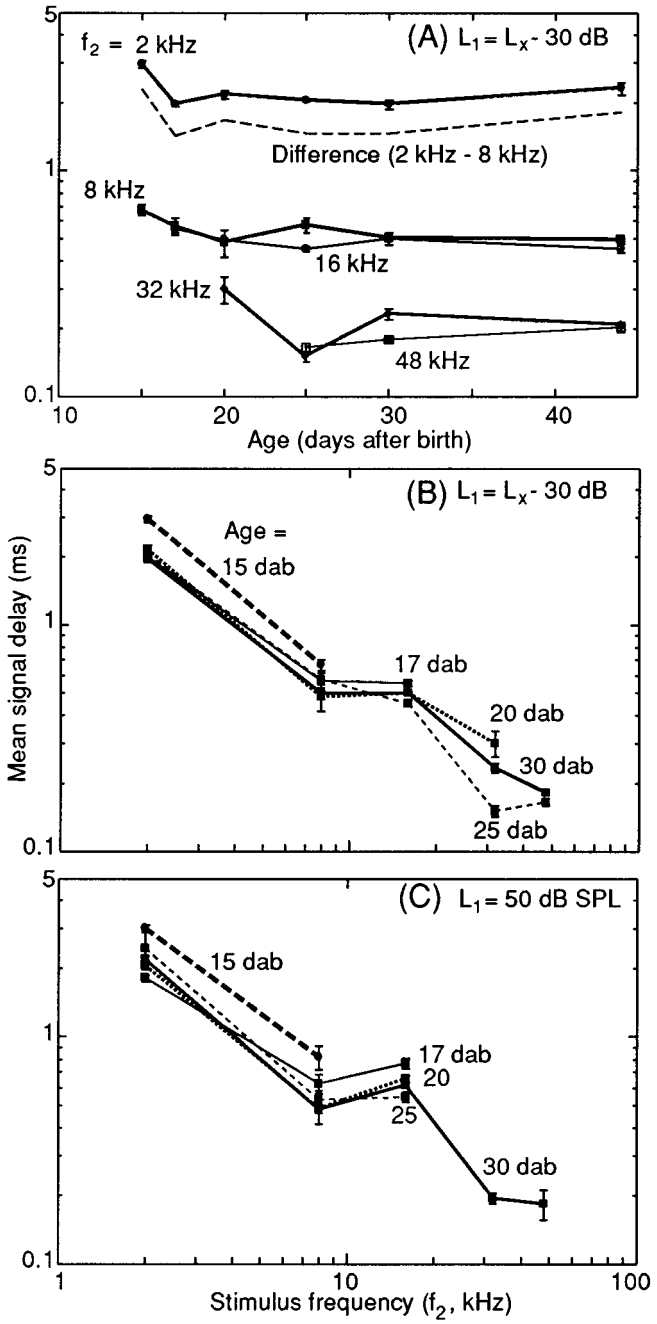


FIG. 6. Group data, giving the variation with age of the mean signal delay. Bars indicate standard error of the mean. (A) Delay estimated for each individual at the stimulus level L_1 which was 30 dB below the active–passive level, L_x, for that individual, and then averaged. The parameter is the stimulus frequency, f_2. The dashed line shows the *difference* in the mean delay between that for $f_2=2$ kHz and that for $f_2=8$ kHz. (B) Same data as (A), plotted to show the variation with f_2 frequency, with age as the parameter in days after birth (dab). (C) Mean delay measured at the stimulus level $L_1=50$ dB SPL, for all frequencies and ages. Note that emission amplitudes for higher frequencies at this stimulus level were not above the noise floor in the younger animals.

B. Variation of amplitude with f_1/f_2: The ''filter'' response for active emissions

Measurements were made of the normal, pre-injection emission amplitudes as a function of stimulus frequency ratio for fixed frequency f_2 of 2, 8, 16, and 32 kHz. Typical

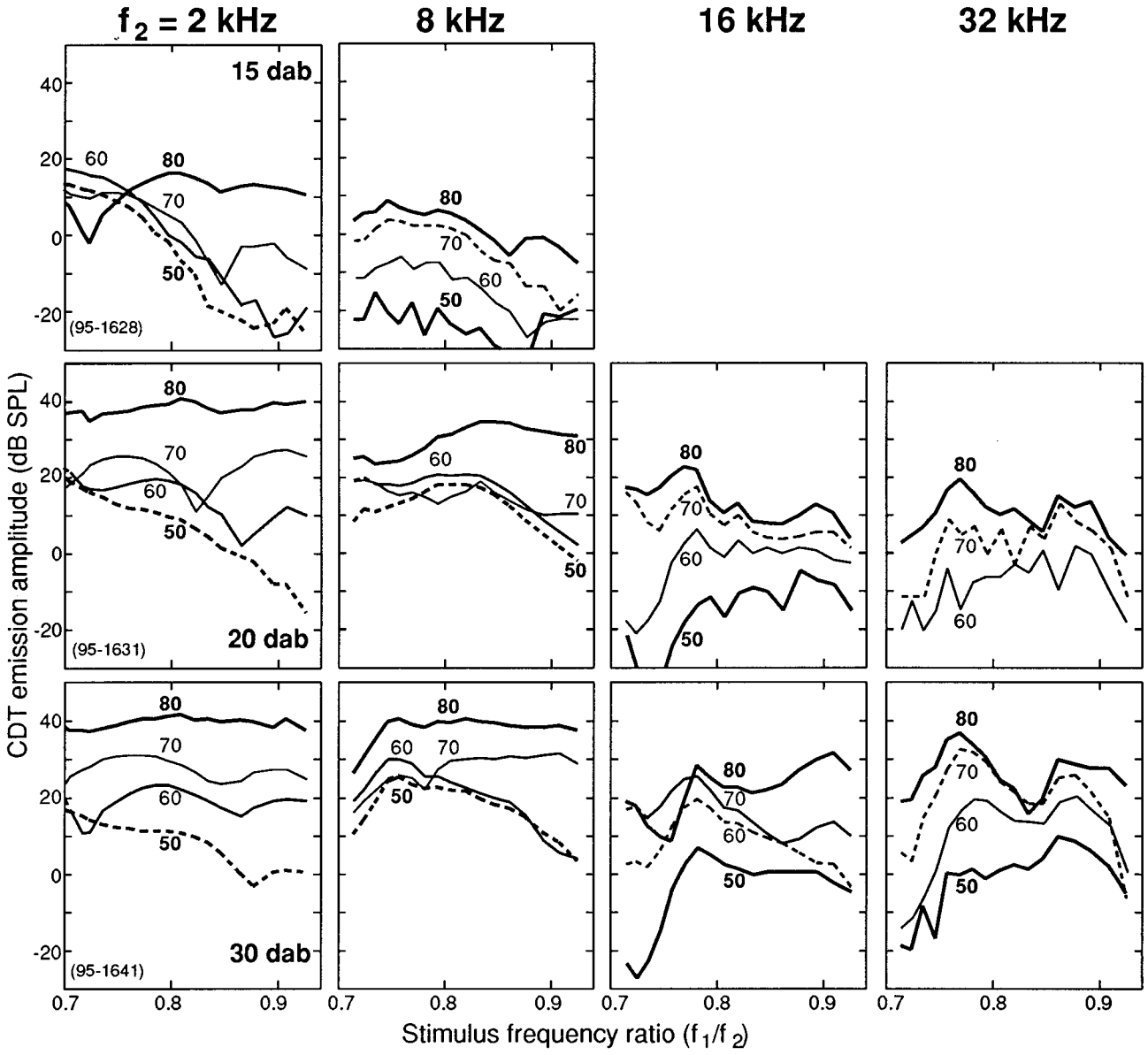


FIG. 7. Variation of cubic distortion tone (CDT, $2f_1-f_2$) emission as a function of stimulus frequency ratio (f_1/f_2), for individual animals at three different ages. The parameter listed is the stimulus level, L_1, in dB SPL. For all data, the higher-frequency stimulus level, L_2, was 10 dB below L_1. For clarity, only results for stimulus levels for L_1=50 to L_1=80 dB SPL are shown, and the extremes are shown in heavier lines. The line in each panel which is dashed indicates the emission for the stimulus level which falls between 20 and 30 dB below the active–passive transition level, L_x, illustrated in Fig. 4.

results are presented for the CDT emission for three individual animals in Fig. 7. For clarity, only the emission distributions for stimulus levels from L_1=50 to 80 dB SPL are shown.

Over the frequency ratio covered, there were often two main peaks in the emission amplitude. There was no obvious, consistent trend for the emission amplitude to ''flatten'' with increasing stimulus level, i.e., the amplitude did not tend to become constant with frequency ratio at high stimulus levels. There also was no consistent change with stimulus level for the frequency ratios associated with the amplitude maxima, i.e., the ''peak center frequency.''

This ''filter'' behavior is examined more closely in Fig. 8, where the mean CDT ($2f_1-f_2$) and fifth order amplitudes ($3f_1-2f_2$) are presented. For this plot, the horizontal axis employs the *emission frequencies* (*re*: f_2) rather than the stimulus frequency ratio. This is done to better compare these two odd-order terms, because it is well known that at least some of the peaks in the odd-order emission amplitudes coincide when plotted against emission frequency, rather than against stimulus frequency (Brown and Gaskill, 1990b; Fahey and Allen, 1986; Gaskill and Brown, 1990). To obtain the mean filter responses, individual filter functions were first defined by normalizing the emission components of each individual animal to the maximum values occurring in the interval observed. The normalized values were then averaged over the age group, and the mean results presented in Fig. 8. For comparison between different ages, at each age we present only the emission distribution for one stimulus level pair, choosing in each case the stimulus level falling between

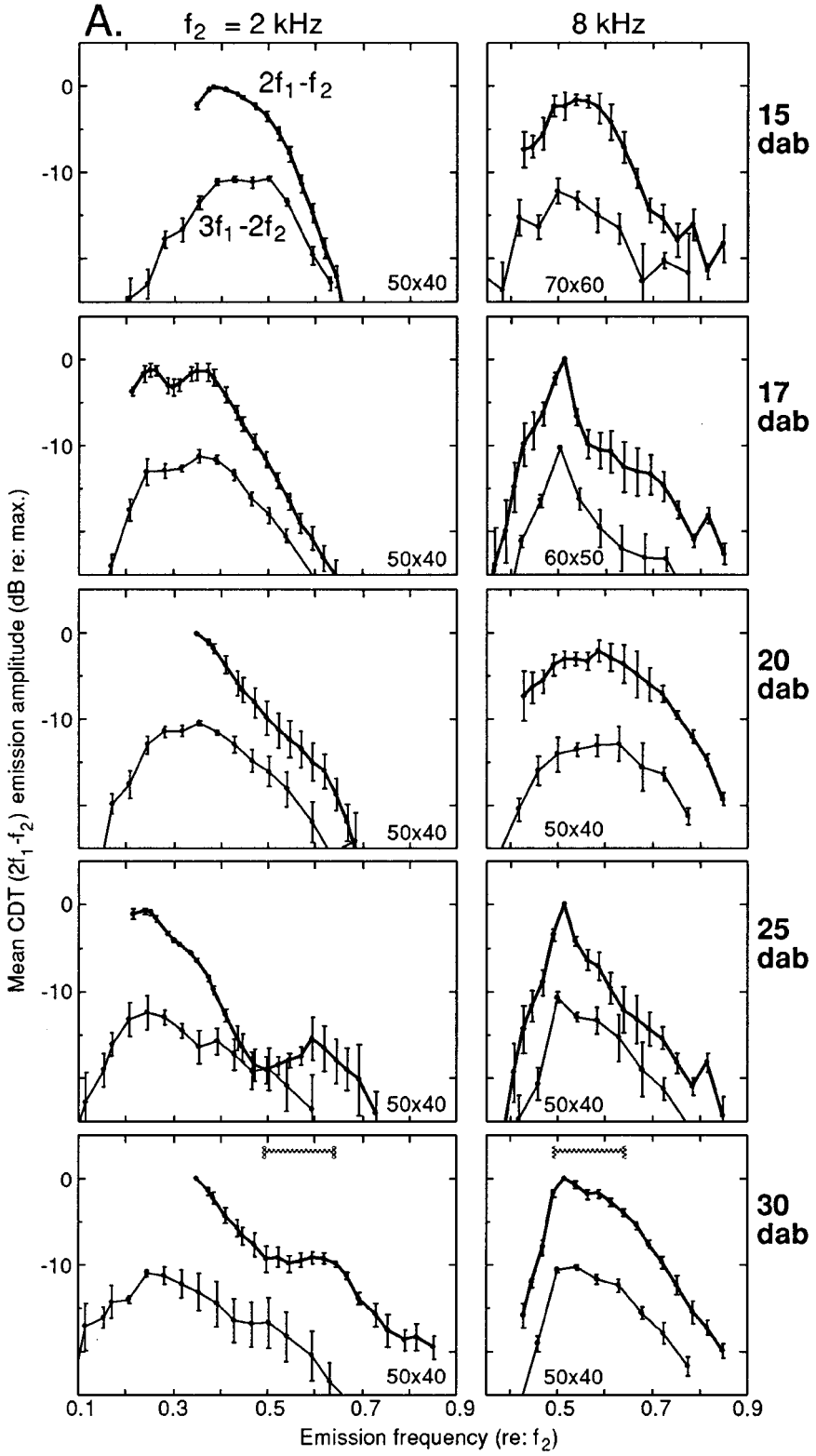


FIG. 8. A, B. Group data. Mean ''filter'' responses for emissions at $2f_1-f_2$ and $3f_1-2f_2$: Variation of emission amplitude when stimulus frequency ratio is changed, for the stimulus levels ($L_1\times L_2$, dB SPL) listed in the lower part of each panel. Stimulus levels L_1 were chosen which were between 20 and 30 dB below the mean active–passive level, L_x, for that age group. (See Fig. 5.) In each panel, the vertical axis is the response (in dB) relative to the maximum emission. That is, individual amplitude responses were normalized to the maximum response for each animal over the measured frequency interval at a given f_2, before averaging over the age group. The 0-dB reference levels are indicated: note the emission for $3f_1-2f_2$ is displaced 10 dB below $2f_1-f_2$ for clarity. The horizontal axis is the *emission* frequency relative to the f_2 frequency, not the stimulus frequency ratio as in Fig. 7. For the $2f_1-f_2$ emission, the horizontal axis is chosen to be equivalent to that for Fig. 7; note, however, that *different* values of f_1/f_2 are associated with the responses for $3f_1-2f_2$ compared to $2f_1-f_2$ at the same position on the horizontal axis. Vertical bars represent standard errors of the mean. For reference, the horizontal bars in the row for the 30 day after birth (dab) animals indicate the frequency range which was employed to determine the signal delays (for all ages) from the phase change of the $2f_1-f_2$ emission, at approximately the same stimulus levels, in the previous section.

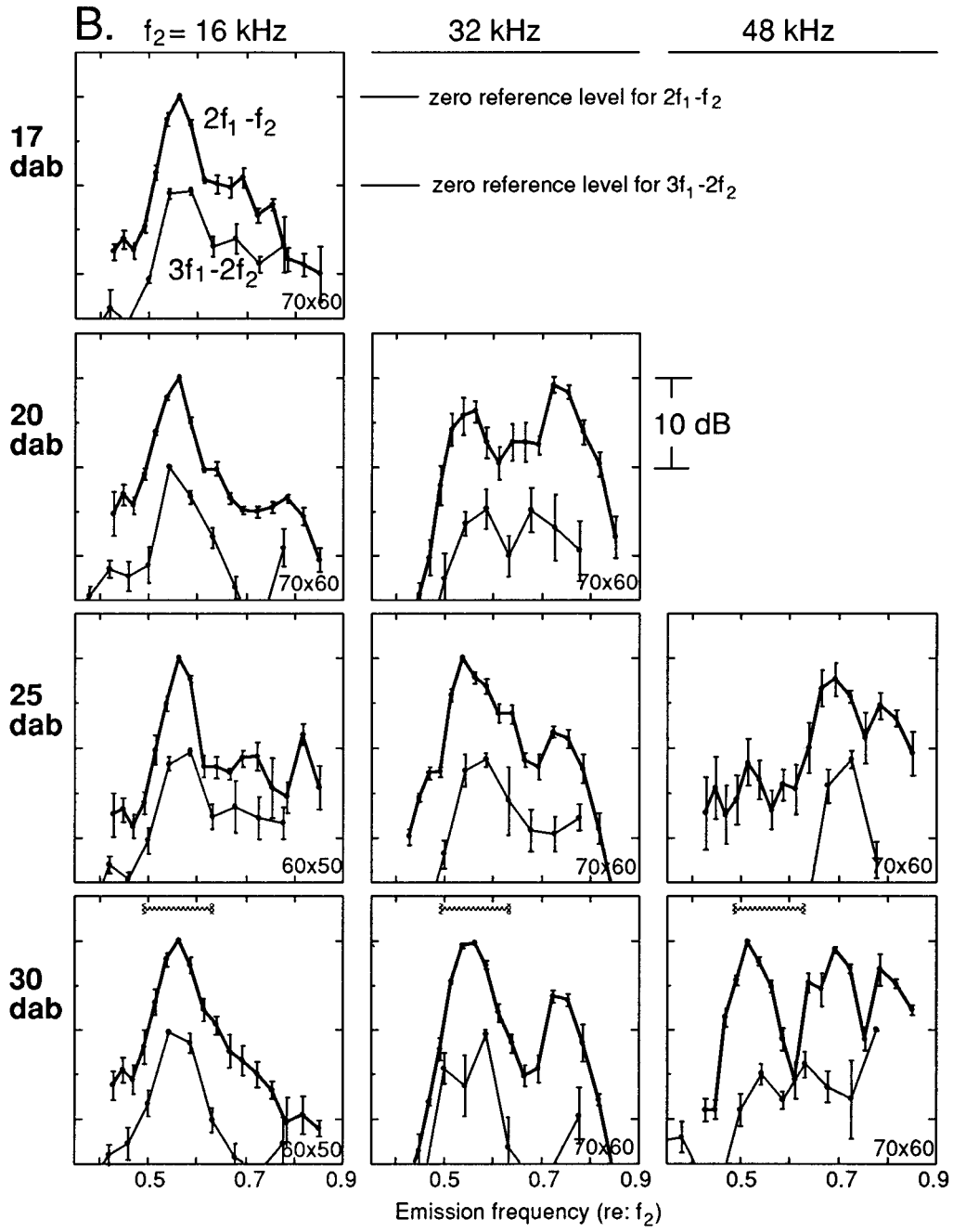


FIG. 8. (*Continued.*)

the mean values for (L_x-30 dB) and (L_x-20 dB) for the age group. That is, we display in Fig. 8 approximately the same stimulus levels that were used in estimates of signal delays in the previous section, so that these amplitude results can similarly be considered characteristic of active processes. The stimulus level pairs chosen are listed at the bottom of each panel. The horizontal bar in the lowest panels indicates for reference the frequency ratios which were employed to obtain the signal delay information from the CDT emission in the first section, for all age groups. The upper, heavier line in each panel denotes the $2f_1-f_2$ component. The lighter line is the $3f_1-2f_2$ component, with its zero reference shifted 10 dB downward for clarity.

The ''second filter'' effect is dramatically displayed in this presentation. The higher order term, $3f_1-2f_2$, had peaks at exactly the same *emission* frequencies as did the cubic term, $2f_1-f_2$. Further, at the stimulus levels chosen, there was generally only one primary peak in the *mean* emission for the adult animals. This primary peak location and shape was typically very consistent from animal to animal in a given age group. This is evidenced by the very small variances typically found around the peak values. Note that if the normalized mean amplitude at the primary peak was 0 dB, i.e., exactly equal to the reference level with zero variance, it means that the maximum amplitude occurred at the same emission frequency for *every* animal in that group. Many of the mean primary peak values were equal to or very close to

0 dB, illustrating the remarkable consistency of the frequency of maximum emission within each age group. The ''center frequency'' of the primary peak (frequency of emission maximum) was close to $0.5f_2$ for all f_2 frequencies, except for f_2=2 kHz.

Developmentally, there was no significant change in the mean center frequencies for these emission peaks across age groups from 15 dab to 30 dab. There was also no obvious trend for the peaks to become sharper as the animal matured.

In addition to these peaks, there were other peaks observed at most ages. These other peaks were most obvious for stimulus frequencies above 2 kHz, where they were typically found with emission frequencies of $0.6f_2$ to $0.7f_2$. These peaks were often relatively stronger at younger ages, frequently strong enough to be the dominant peak.

The mean results for f_2=8 kHz were more complex. There was an apparent broad maximum at 15 dab, which became quite sharp at 17 dab, but then broadened again at 20 dab, only to change again into a sharp peak at 25 dab. The emission frequency at the maximum amplitude also appeared to change with age in a complex manner. Figure 9 presents the individual observations which made up these mean distributions, for more detailed analysis.

The individual distributions show that the broad peaks in the mean data could generally be considered to consist of the combination of two close, partly ''resolved'' peaks, rather than a single broad peak. These typically consisted of a single peak at an emission frequency of $0.5f_2$, and another peak at about $0.6f_2$. At most ages, the first peak was the dominant one. At 20 dab, however, the first peak appeared to be relatively and absolutely weaker than at other ages, to the extent that it was not detectable in comparison to the second peak in three of the six animals in this age group. At 25 dab, all of the animals had very similar amplitude distributions, having a single sharp maximum at about $0.5f_2$. There were at least some animals in every age group with a peak of similar sharpness occurring at nearly the same frequency ($0.5f_2$). The exception was the youngest age: At 15 dab, no single animal had a peak that sharp, although most animals did have a peak with a relative maximum at an emission frequency near $0.5f_2$.

The overall behavior of the peak responses are summarized in schematic form in Fig. 10. Here, we include the mean peak frequencies, relative amplitudes, and peak widths for all animals in this study. Each panel represents one stimulus frequency, f_2, in which the age of the animal increases downward. The frequencies at maximum amplitude have been determined for individual responses and then averaged over each age group. The ''center'' frequency variances were quite small, typically about $0.01f_2$. The relative mean peak amplitudes, essentially the same as in Fig. 8, are indicated by the size of the symbol at the peak frequency (see key). The horizontal bars represent the widths of the peak, measured 10 dB below the peak maximum. Because of the interference in the response caused by neighboring peaks (e.g., Fig. 9) the two ''half-widths,'' i.e., the distances from the emission frequency at maximum amplitude to the frequencies where the emission was 10 dB lower, were determined for each individual peak. The *smaller* half-width was

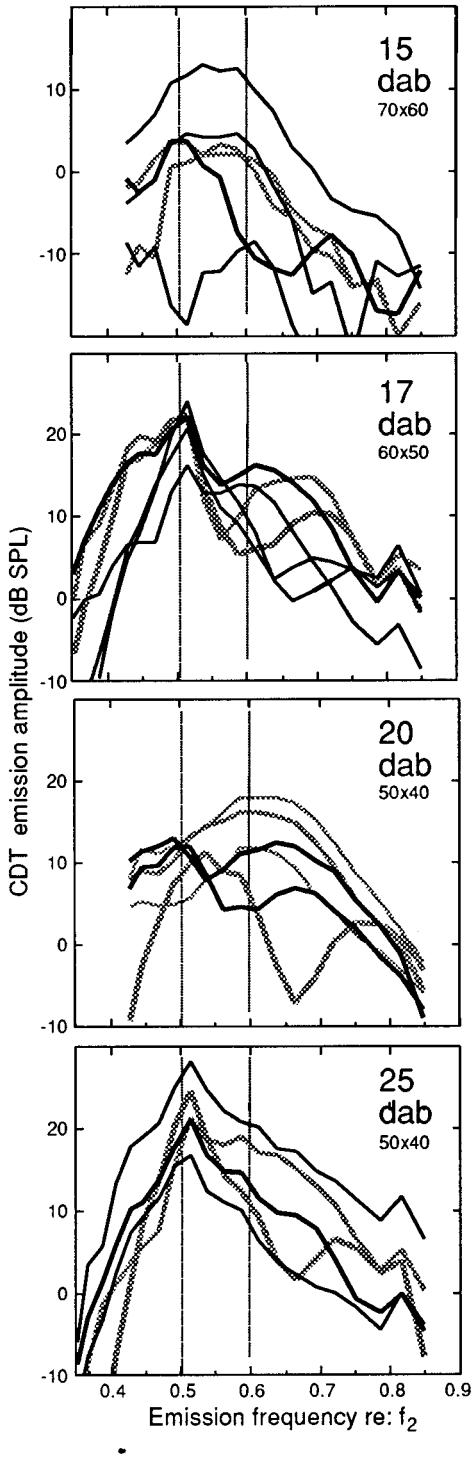


FIG. 9. Individual responses for all animals in each age group for f_2=8 kHz. The stimulus levels chosen were the same as Fig. 8 (i.e., L_1 was between 20 and 30 dB below L_x) and are indicated in each panel ($L_1 \times L_2$, dB SPL). For clarity, some individual responses are emphasized by different line intensities. The two vertical lines, at $0.5f_2$ and $0.6f_2$, are for reference.

used to obtain the means shown. The stimulus levels were the same as in Fig. 8, that is, between 20 and 30 dB below the crossover level, L_x.

There was a consistent developmental trend for f_2=2 kHz. There was one prominent peak at relatively low emission frequencies, which shifted to even lower emission frequencies as the animals matured. The peak width, however,

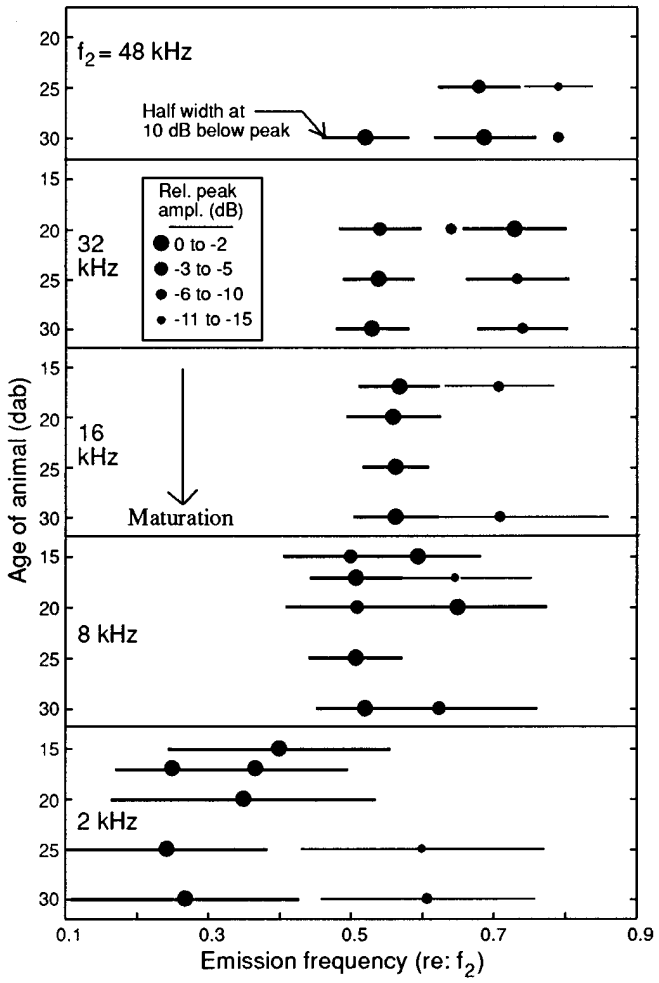


FIG. 10. A schematic illustrating the development of mean peak amplitudes, center frequencies, and half widths. Center frequencies of each peak were determined for the individual responses, and averaged to obtain the mean peak frequencies shown. The relative mean amplitude of each peak was determined from Fig. 8, and is denoted by the size of the symbol, according to the key. The horizontal bars are not the variances of the center frequencies, which were much smaller. Rather, the bars represent the width of the amplitude response at the point 10 dB down from the maximum. Strictly, since the peaks were usually asymmetrical, often because one side was obscured by the presence of a neighboring peak, the *smaller* of the two estimated ''half'' widths of each individual peak was used to obtain the mean half widths. Widths are not shown for two of the peaks, because the response was too confused by noise or by stronger neighboring peaks.

changed little during this time. There was also a secondary peak at about $0.60f_2$ which appeared as the animals matured. It should be noted that this peak could have been ''present'' earlier, but obscured by the much stronger, lower-frequency peak.

For higher frequencies, there was no general trend noted in the variation of peak frequencies or widths with development. For f_2=8 kHz, for example, there was a prominent peak near $0.50f_2$ at all ages. Neither the center frequency nor the peak width changed consistently with age. Similar results hold for f_2=16 kHz, which had a very sharp, stable peak at $0.56f_2$ at all ages, and for f_2=32 kHz, which consistently had two peaks near $0.50f_2$ and $0.70f_2$. The observed peaks could be quite sharp, with half widths at the 10-dB level as small as $0.04f_2$.

C. Relationship of active and passive emissions

Passive emissions are operationally defined as those emissions measured at the time of maximum furosemide effect, when the endocochlear potential has been substantially reduced (Mills *et al.*, 1994; Mills and Rubel, 1994). These passive emissions are almost always weaker than (or equal to) the normal emissions (e.g., see Fig. 4) and can only be measured during a brief time. The measurement of passive emissions is therefore more difficult than for the active emissions. In particular, in this experiment, for the younger animals the passive emissions were generally so weak that the programmable filter had to be employed, to reduce the amplitude of the much stronger primaries prior to A/D conversion (see Sec. I). This resulted in the loss of accurate phase information for the passive emissions from the younger animals, so that the passive signal delays could not generally be calculated. Further, even with the additional filtering often only the passive emissions near the maximum amplitude (as a function of stimulus ratio) could be adequately measured. While there was enough information available on all animals to estimate cochlear amplifier gains and active–passive stimulus levels (as in Fig. 4), there was only sufficient information to adequately characterize the passive emission amplitude distribution and the signal delay for several individual, mostly older animals.

The question naturally arose, in seeing the sharp peaks in the active emissions (Fig. 8), whether the same peaks occurred in the passive, post-injection response at similar stimulus levels. This might already be concluded, since we have already shown: (1) that the emissions at high stimulus levels are typically little changed by furosemide intoxication (Fig. 4), i.e., the pre-injection and passive emissions have the same distribution at high stimulus levels; and (2) there is little change in the pre-injection emission amplitude distribution comparing high and low stimulus levels (Fig. 7). One might also like to compare the two distributions directly, by comparing the pre-injection emissions at low stimulus levels with the ''passive,'' post-injection emissions at the same stimulus levels, i.e., 20–30 dB below L_x. This is not strictly possible, because the passive emissions are simply not measurable at these stimulus levels (e.g., Fig. 4). As a compromise, in Fig. 11 we have compared the active, pre-injection emissions at low stimulus levels for several individual animals to the post-injection, passive emissions measured in each individual *at a stimulus level 30 dB higher*.

Using the established slope of 2:1 for the passive emissions, the passive amplitude distribution was then shifted downward by 60 dB, so that it would represent the amplitude expected at the *same* stimulus level as the active emissions (Mills *et al.*, 1994; Mills and Rubel, 1996). The amplitude scales of the active and passive emissions are given on the left and right sides of Fig. 11, respectively. The scales have been set so that the active and passive emissions would overlap in these panels *if* the difference between the active and passive emission amplitudes at the same stimulus level was 40 dB. For the stimulus levels we are considering here, this would generally occur if the shift ΔC was 40 dB (Fig. 4).

It is clear from Fig. 11 that the major peaks observed in the active emissions (as a function of stimulus frequency)

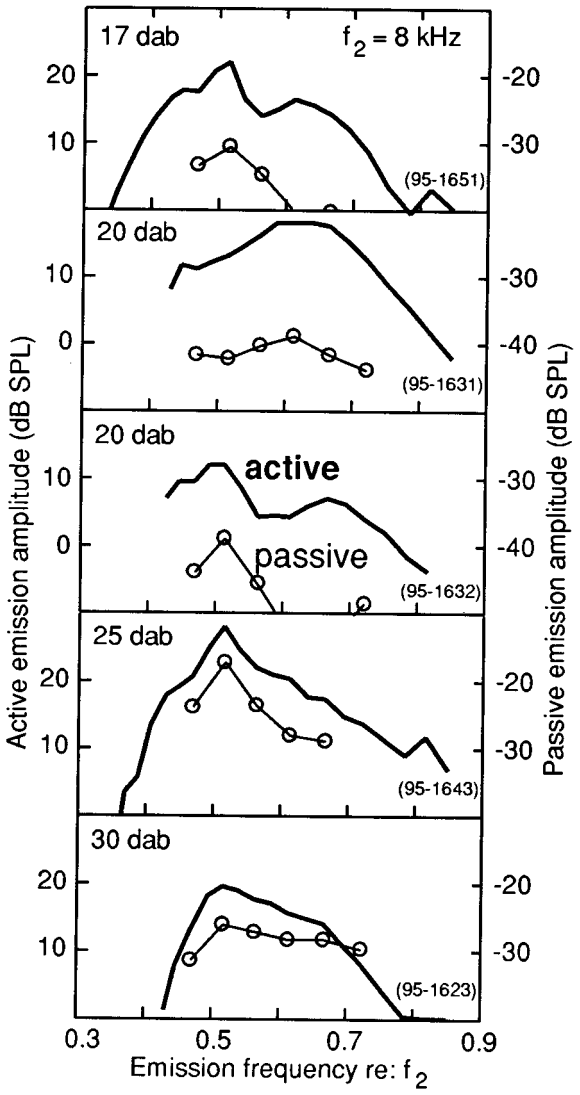


FIG. 11. Comparison of active and passive ''filter'' responses for individual animals at f_2=8 kHz. For the active responses at each age, the stimulus levels were the same as that in Figs. 8 and 11, i.e., were $L_1 \times L_2$=60×50 dB SPL for the 17 dab animal, and 50×40 dB SPL for the others. The passive responses were obtained at the time of maximum furosemide effect, at the stimulus level 30 dB higher than that for the active response. The passive responses were then extrapolated down to the same stimulus level as the active, using the established 2:1 slope, i.e., 60 dB was subtracted from the observed amplitude. The scales for the active and passive amplitudes are different, and are on the left and right sides, respectively. The scales were set so that, if the data points overlapped, the active emission amplitude was 40 dB more than the passive amplitude (i.e., if the active and passive points overlapped, A_c=40 dB).

were typically found in the passive, post-injection emissions as well. There was sometimes a modest difference between active and passive emission distributions in the *relative* peak amplitudes noted in individual animals. For example, in the 17 dab animal, the peak at higher frequencies (at 0.62 f_2) was relatively larger for the active emission than for the passive. This is probably related to the similar shift in the relative amplitudes of the peaks of the normal emission as a function of stimulus level (Fig. 7).

Data such as that in Fig. 11 lead to the following conclusion: For low stimulus levels, when the cochlear amplifier

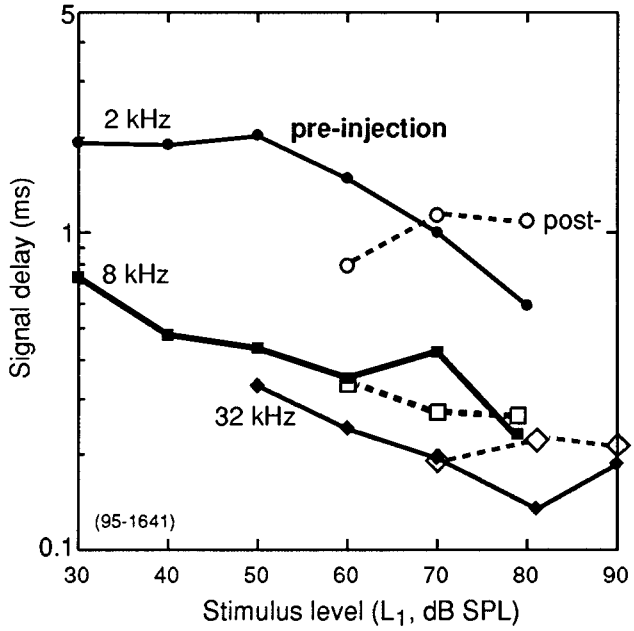


FIG. 12. Comparison of signal delay pre- and post-injection for an individual 30 dab animal. The parameter listed is the f_2 frequency. The pre-injection data are shown in filled symbols and solid lines, the post-injection in open symbols and light lines. Note that the symbols for f_2 frequencies of 8 and 32 kHz at a stimulus level of 80 dB SPL are displaced slightly in the horizontal direction so that they do not overlap.

was made inoperative by furosemide, the entire emission amplitude pattern shifted downward. The amount of shift in amplitude was approximately the same at all stimulus frequency ratios, and was equal to the shift, ΔC, at that f_2 frequency.

Finally, Fig. 12 displays the signal delay found as a function of frequency for a typical young adult animal. The pre-injection results are shown in solid lines and filled-in symbols, with the post-injection delays shown in shaded lines and open symbols.

As noted earlier, the pre-injection signal delays generally decreased as the amplitude increased. That is, the delays associated with the active process (the normal emissions at low stimulus levels) were larger than those found at higher stimulus levels. The signal delays for the pre-injection high level emissions, on the other hand, were very similar to the passive emissions as defined here. The signal delays for the post-injection, passive emissions, on the other hand, did not appear to change consistently with a change in stimulus level, at least over the relatively limited range over which the passive emissions could actually be measured.

III. SUMMARY OF RESULTS

The response of the cochlea as a function of stimulus frequency ratio (f_1/f_2) is complex; it is complex at a given developmental age and the response changes with age in a complex, often counterintuitive manner. The complexity applies both to changes in phase angle, including the derived signal delay for a narrow range of stimulus frequency ratios, and to changes in amplitude, including the observed multiple peak structure. Because of this complexity, it seems useful

here to summarize the data obtained in this study. The goal here is to provide a *tentative* summary of the most salient features of the results to date. Obviously, many of these trends require confirmation.

(1) The active and passive emissions have essentially the same signal delay and amplitude response as a function of f_1/f_2 (Figs. 11 and 12). That is, the normal signal delay measured at high stimulus levels is about the same as that measured at the same high levels when the cochlear amplifier is temporarily nonfunctional. Further, the passive amplitude/filter response is similar in form to the active filter response at *all* stimulus levels. In other words, when furosemide is injected, the emission amplitudes for low stimulus levels shift downward by approximately the same amount, relatively independent of stimulus frequency ratio. However, the signal delay associated with the active response (i.e., the normal delay at low stimulus levels) is higher than that associated with the passive or active responses (at high stimulus levels). The passive signal delay at low stimulus levels is not measurable.

(2) At a given age, the signal delay of active emissions typically decreases with increasing stimulus level, except for a complex response at high stimulus levels (Figs. 3 and 12).

(3) This complex response is associated with the active–passive transition (Figs. 4 and 5). It seems useful to compare responses at different frequencies and ages not on the basis of the absolute stimulus levels, but on the basis of the stimulus levels *relative* to the active–passive transition level, L_x. The comparison seems valid for both phase and amplitude responses. To study active emission characteristics, responses for stimuli well below the active–passive transition should be compared.

(4) At all frequencies, the derived signal delay typically *decreases* only slightly during maturation. However, there is an interesting relative minimum in signal delay which occurs, first at lower frequencies and subsequently at higher frequencies [Fig. 6(A)].

(5) The changes in signal delay with development are relatively small and complex. Depending on the observed frequencies and age groups, increases or decreases could be measured [Fig. 6(B)]. In terms of its signal delay characteristics, the cochlea in gerbils seems essentially mature near hearing onset.

(6) The difference in signal delay between low and high frequencies appears to be dominated by intrinsic changes in the larger, low-frequency delay [Fig. 6(A)].

(7) The structure in the amplitude responses, as a function of stimulus frequency, changes little with stimulus level at any age (Fig. 7). This includes the peak center frequencies, which are essentially unchanged with stimulus level. The peaks can be quite sharp, with mean half widths at 10 dB below the peak as small as $0.04f_2$.

(8) At all ages, the responses of the CDT at $2f_1-f_2$ and the fifth order term at $3f_1-2f_2$ are very similar when plotted as a function of the *emission* frequency (Fig. 8). The most prominent peak in adult animals is usually at an emission frequency between $0.5f_2$ and $0.6f_2$.

(9) For $f_2=2$ kHz, the center frequency of the most prominent peak decreases with age. At higher frequencies, the center frequencies of individual peaks are quite stable, but there may be a modest tendency for the relative amplitudes of the peaks to change with age. That is, the most prominent peaks in the younger animals tend to be those at higher frequencies (closer f_1 and f_2 frequencies), while those in adult animals tend to be those at the mid-frequencies (emission frequencies equal to $0.50f_2$ to $0.60f_2$). There are no consistent developmental trends in the peak center frequencies or peak widths themselves (Fig. 10).

(10) Overall, except for the changes noted at 2 kHz, the cochlea appears essentially mature near hearing onset in the characteristics which account for the variation of amplitude and phase with stimulus frequency ratio.

IV. DISCUSSION

A. Decrease in signal delay with stimulus intensity

From simple arguments, the largest component of the observed signal delay for distortion product emissions in the mammalian cochlea is likely to be phase buildup in the ingoing traveling wave associated with the stimulus (Eggermont, 1979; Mahoney and Kemp, 1995). Our observation that the relative signal delays change very little with development (Fig. 6) implies that the cochlear amplifier is mature in this respect even at early ages. The phase changes due to place code shifts are presumed to be comparatively small, and so not detected in these measurements.

It is well known that the signal delay, calculated by Eq. (1) from the observed phase changes, decreases with increasing stimulus intensity (Brown and Kemp, 1985). The observed behavior apparently follows from two simple assumptions:

(1) The cochlear amplifier reaches nonlinear levels even for generally small stimulus levels. For typical mammalian ears, the saturation of the cochlear amplifier begins to produce signal compression at such low signal levels that the existence of a linear regime cannot easily be established (Brown, 1993; Goldstein, 1967).

(2) The dominant odd-order emissions are generated in areas where the cochlear amplifier operation has become nonlinear. This assumption rests on the idea that the nonlinear mechanism which produces the emission is intrinsically related to, or is the same as, the nonlinear mechanism which causes the saturation. This idea has support in measurements which show that the vulnerable odd-order emissions are intrinsically tied to the mechanisms which produce amplification and sharp tuning at low stimulus levels (Rübsamen *et al.*, 1995). This assumption is further based on the simple idea that the dominant distortion will be produced where the response amplitude is sufficient to involve nonlinear mechanics.

The traveling wave phase buildup is very rapid as the peak of emission is reached (Robles *et al.*, 1986; Zweig, 1991). As the stimulus level increases, the cochlear amplifier saturation begins to occur more and more basally (Johnstone *et al.*, 1986). If the centroid of the emission generation area moves even slightly basally with increasing stimulus levels, this effect could easily cause a reduction in the apparent signal delay with increasing stimulus intensity.

B. Interpretation of development of signal delay characteristics

For these measurements, the frequency f_2 was fixed and the frequency f_1 varied. The argument for this approach is that the location of the site of emission generation in the cochlea, at least for weak stimuli, is believed to be near the f_2 place (Brown and Kemp, 1984; Kummer *et al.*, 1995; Martin *et al.*, 1987). That is, the emission originates near the location on the BM where a single weak tone of frequency f_2 would have a maximum response. There is then some hope that varying f_1 while fixing f_2 would leave the physical site of generation approximately unchanged, allowing for a simpler interpretation of the phase changes (Mahoney and Kemp, 1995).

Even for constant stimulus levels and f_2 frequency, the derived signal delay depends somewhat on the frequency ratio. For practical and theoretical reasons, it has been usual to determine the signal delay for a small range of stimulus frequency ratios which span the ratio where the emission amplitude is a maximum. For this report, we chose a single range in stimulus frequency ratio which was a compromise, approximately spanning the CDT emission maximum as a function of f_2. Other possibilities exist. We could, for example, have changed the frequency ratio for each f_2, and even for each age. This seems like a needless complication, which could potentially confuse the interpretation. In any case, the center frequencies of the peaks turned out to be insensitive to age.

Using either the phase delay at constant stimulus levels in the ear canal [Fig. 6(C)] or at a constant level below the active–passive transition level [Fig. 7(A) and (B)], we found a modest decrease in signal delay with age. Such a decrease, found in human infants, was used to suggest that there was no shift of the place code after this age (Brown *et al.*, 1994). However, this conclusion does not appear to be justified by these results, for the following reasons. In gerbil, the shift in the place code for mid- and high frequencies appears to be primarily due to a shift in the passive cochlear response, associated with a shift in the passive base cutoff frequency (Mills *et al.*, 1994). The place code shift merely moves the place on the BM where the cochlear amplifier begins to respond at a given frequency. The phase buildup associated with the passive, linear response of the incoming stimulus, *up to the point where the cochlear amplifier starts to respond*, is estimated to be relatively small compared to the phase buildup associated with the cochlear amplifier (Eggermont, 1979; Zweig, 1991). Therefore, it is expected that developmental changes in the cochlear amplifier itself would dominate signal delay changes, masking any effects due to place code shifts.

This expectation appears to be supported by the data found for the gerbil. After normalizing the stimulus levels to the active–passive transition, there was very little change in signal delay found with age. During the same period, it is known that there are significant changes in the place code in the base of the gerbil cochlea (Arjmand *et al.*, 1988; Harris and Dallos, 1984; Mills *et al.*, 1994; Mills and Rubel, 1996). Unless there were fortuitous, and unlikely, compensating changes, the known place code changes cannot have affected the signal delay very much. We conclude that, in gerbils at least, the place code shift cannot be determined by measurement of the signal delay characteristics. In addition, we conclude that from the onset of hearing to adulthood in gerbils, there are relatively small changes in the cochlear properties which are responsible for the signal delay characteristics. In contrast, there are other characteristics of the cochlea and cochlear amplifier which change considerably during the same period (Sanes *et al.*, 1989; Rubel, 1978; Rubel and Ryals, 1983; Walsh and Romand, 1992; Woolf and Ryan, 1984; Woolf *et al.*, 1986; Yancey and Dallos, 1985).

C. Emission amplitude response as a function of stimulus frequency ratio

We have described an amplitude response in gerbils with multiple maxima at emission frequencies of $0.50f_2$ to $0.75f_2$, at f_2 frequencies of 8 kHz and higher (Figs. 8 and 12). The relative amplitudes of the peaks seem to change with stimulus level and with development. However, the widths and center frequencies of the peaks seem quite stable, both with stimulus level (Fig. 7) and with age (Figs. 8 and 10). In contrast, Brown and colleagues have proposed that there is a single important amplitude maximum, which occurs when the emission frequency is a half-octave below f_2, i.e., when the emission frequency is about $0.7f_2$ (Brown, 1987; Brown and Gaskill, 1990a, b; Brown *et al.*, 1992, 1993; Brown and Kemp, 1985; A. M. Brown *et al.*, 1994; Brown and Williams, 1993; Gaskill and Brown, 1990). The stimulus levels usually employed in their recent measurements are $L_1 \times L_2 = 55 \times 40$ dB SPL, and they note that the half-octave relationship is ''lost'' if higher L_2 levels are allowed, but only for $2f_1 - f_2$ (Brown and Williams, 1993). While some of our gerbil data would fit with the half octave proposal, clearly most of our observed responses are more complicated. For example, peaks at an emission frequency a full octave below f_2 are found frequently and prominently, as well as peaks near $0.60f_2$. These differences appear real, and seem relatively unchanged as a function of stimulus amplitude. We certainly find these peaks for $L_1 \times L_2 = 50 \times 40$ dB SPL, for example.

Some of the differences between our results and those of Brown's group may be species differences, in that the majority of the data supporting their conclusions are drawn from humans and guinea pigs. In contrast, their gerbil observations (Brown and Kemp, 1985, Fig. 1) do show for $f_2 = 4$ kHz a prominent peak located at about $0.62\ f_2$, which persists at high stimulus levels and is joined by a lower-frequency peak. Further, a close examination of their human data shows that there often appear to be multiple peaks, including peaks at frequencies lower than a half octave below f_2 (although not as low as a full octave below). While it is a useful first approximation, therefore, the observed emission amplitude responses appear to be more complicated than the ''half octave'' rule can encompass, at least in gerbils.

D. The interpretation of the amplitude response

The structure in the amplitude response has been interpreted as evidence in favor of a ''second filter'' in the co-

chlea (Allen and Fahey, 1993a, b; Brown and Williams, 1993). The general idea is that distortion forces are generated in the outer hair cells (OHCs). In this model, the force generation is presumed to be relatively insensitive to stimulus frequency ratio, that is, insensitive on the scale of the amplitude structure, which involves changes in f_1/f_2 of 0.10 or less. At low stimulus levels, the maximum of distortion generation is assumed to occur in the OHCs located near the f_2 place. To be observed in the ear canal, these distortion forces must be translated into BM motion. It is in this translation that the emissions are assumed filtered by the micromechanical structure of the cochlea. One specific model has suggested that the filtering occurs in the tectorial membrane–outer hair cell stereocila system, which causes a ''short-circuit'' for distortion frequencies a half octave below f_2 (Allen and Fahey, 1993a).

This specific model now seems unlikely to be adequate, given recent results that show similar filtering in barn owls, and in certain alligator lizards who do not even have a tectorial membrane in the relevant frequency range (Tashenberger *et al.*, 1995). We may further note that the reported responses in these two very different species are remarkably similar to each other, and fairly similar to those in the gerbil reported here. The shape of the amplitude response is insensitive to stimulus level for both barn owl and lizard, over the limited range reported. There are apparently multiple peaks in both, for the barn owl at emission frequencies of $0.50f_2$ and $0.66f_2$ for f_2=7.9 kHz, and for the lizard at $0.43f_2$ and $0.70f_2$ for f_2=4 kHz. The peak frequencies noted above for the gerbil span these values.

Hearing in the gerbil does not begin until 13–14 days after birth (dab), and the endocochlear potential does not reach adult levels until after 20 dab. There are still structural changes involving the tectorial membrane through the early period, as well as possible BM changes associated with the place code shift (Schweitzer *et al.*, 1996). Further, there are significant changes in cochlear amplifier operation and base cutoff frequency through this period (Mills *et al.*, 1994; Mills and Rubel, 1996). In contrast, we have found in this study that the center frequencies and widths of the peaks in the emissions seem to be remarkably stable through most of this developmental period. Further, both amplitude and phase variations are generally the *same* for the active processes and the passive responses. This implies that, whatever the cause of the ''filtering'' of the emissions seen here, it is insensitive to specific details of the developing micromechanical BM structure. In short, the filtering itself does not appear to be associated with the operation of the amplifier operation. The generally increased overall amplitude of the emissions at low stimulus levels in mammals, of course, is directly attributable to the cochlear amplifier operation.

These results all argue that the structure in the emission amplitude response is unlikely to be a useful indicator of the tuning of the BM. Indeed, studies which have attempted to directly relate structure to tuning have so far found little or only moderately significant correlations (Brown *et al.*, 1993). The origin of the emission amplitude structure remains unknown.

It also should be noted that the structure discussed here is different from the ''threshold microstructure'' which is well known in humans (e.g., Long and Tubis, 1988a, b), and the fine structure in distortion product emissions recently demonstrated in humans (He and Schmiedt, 1993). Compared to the structure noted here, the threshold microstructure and emission fine structure both have a generally finer scale, are much more variable between individuals, and generally disappear with increasing stimulus level.

In contrast to the stable location in relative frequency of the amplitude peaks, mammals seem to have some variation in the ''envelope'' of the amplitude response as a function of age and frequency. That is, there is a variation in relative peak amplitudes with age. This variation may be related to the fact that an upper stimulus level limit for measurement of the ''half octave'' response in mammals has been reported for $2f_1-f_2$ (Brown and Gaskill, 1990a; Brown and Williams, 1993), e.g., this effect could be due to another component becoming dominant as the stimulus level increases. Here we suggest a modest change in relative peak amplitude in gerbils with age, stimulus level, and frequency. However, the envelope response seems of little practical use, since the emission amplitude response is so strongly dominated by the multiple peak structure.

We have also found significant developmental changes in the amplitude response at 2 kHz with age. This is in agreement with earlier studies in gerbils, showing that the cochlear amplifier in the apex continued to mature over these ages (Mills *et al.*, 1994; Mills and Rubel, 1996).

V. CONCLUSIONS

(1) Caution is indicated in interpreting signal delays from emission measurements. Because the variation of observed delay time depends strongly on stimulus level, a difference in the intrinsic passive threshold can cause an apparent change in signal delay. Conceptually, it makes more sense to base measurements on stimulus levels at a constant level relative to cochlear function, than relative to the amplitude in the ear canal. The procedure used here, to normalize the stimulus levels to the active–passive crossover level, appears promising but its validity remains to be independently established.

(2) Even with this correction, the round trip signal delay measured using distortion product emissions is not an adequate method of estimating physical distances in the cochlea, e.g., the distance to the place where the traveling wave peaks. Instead, the signal delay appears to depend primarily on the phase buildup of the traveling wave to the region in the cochlea where the emissions are generated. The considerable variation of the measured delay with stimulus amplitude (probably due to saturation effects in the cochlear amplifier) makes it impossible to accurately determine the contribution of any other delays.

(3) The presence of multiple peaks in the active emission is remarkably stable with age in both the emission frequencies at the amplitude maxima and the peak half widths, and is similar to that reported in birds and lizards. In contrast to these nonmammalian species, there appears to be a modest change in the relative magnitudes of the peaks in the $2f_1-f_2$ emission with stimulus intensity. Peaks at higher

emission frequencies (i.e., closer to f_2) appear to be relatively more prominent in younger animals.

(4) The cochlear amplifier at the base of the cochlea appears to be quite mature from the onset of hearing, in the aspects responsible for signal delay. This is in agreement with earlier studies (Mills *et al.*, 1994; Mills and Rubel, 1996), which found that the limitation on auditory function in the base of the cochlea stemmed from immaturity in the passive response. In contrast, significant development of cochlear amplifier function itself was noted for f_2 frequencies of 1–2 kHz. This delayed apical development is supported by the larger changes observed here with age in the amplitude structure at 2 kHz.

(5) There is an interesting minimum in the derived signal velocity which occurs first at low frequencies, then later at high frequencies, typically a few days to a week after the emission first becomes measurable at that frequency. The significance of this change is unknown.

(6) The pre- and post-injection emissions show the same amplitude and phase responses at similar stimulus levels. The simplest hypothesis to account for this observation is that they are produced by the same mechanism, emitted at the same place in the cochlea, and if ''filtered,'' filtered by the same process after production. Remarkably, the active emissions, those odd-order emissions at low stimulus levels which are interrupted by furosemide injection, also have the same basic structure as the high-level pre- and post-injection emissions. The active signal delays are generally larger than that for the passive emissions, which agrees with simple ideas about the effect of the cochlear amplifier on the signal delay. There is a rapid variation in the derived signal delay at the crossover between active and passive levels.

ACKNOWLEDGMENTS

We thank Brandon Warren for assistance with software development on this project, and two anonymous reviewers for constructive comments on an earlier version of the manuscript. Support was provided by Grant No. NIH DC 00395 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

Allen, J. B., and Fahey, P. F. (**1993a**). ''Evidence for a second cochlear map,'' in *Proceedings of the International Symposium on Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, River Edge, NJ), pp. 296–303.

Allen, J. B., and Fahey, P. F. (**1993b**). ''A second cochlear-frequency map that correlates distortion product and neural tuning measurements,'' J. Acoust. Soc. Am. **94,** 809–816.

Arjmand, E., Harris, D., and Dallos, P. (**1988**). ''Developmental changes in frequency mapping of the gerbil cochlea: Comparison of two cochlear locations,'' Hear. Res. **32**, 93–96.

Brown, A. M. (**1987**). ''Acoustic distortion from rodent ears: A comparison of responses from rats, guinea pigs, and gerbils,'' Hear. Res. **31**, 25–38.

Brown, A. M. (**1993**). ''Distortion in the cochlea: Acoustic $f_2 - f_1$ at low stimulus levels,'' Hear. Res. **70**, 160–166.

Brown, A. M., and Gaskill, S. A. (**1990a**). ''Can basilar membrane tuning be inferred from distortion measurement?'' in *The Mechanics and Biophysics of Hearing: Proceedings of a Conference held at the University of Wisconsin, Madison, WI, June 25–29, 1990*, edited by P. Dallos, C. D. Geisler, J. W. Mathews, M. A. Ruggero, and C. R. Steele (Springer-Verlag, New York), pp. 164–169.

Brown, A. M., and Gaskill, S. A. (**1990b**). ''Measurement of acoustic distortion reveals underlying similarities between human and rodent mechanical responses,'' J. Acoust. Soc. Am. **88,** 840–849.

Brown, A. M., Gaskill, S. A., Carlyon, R. P., and Williams, D. M. (**1993**). ''Acoustic distortion as a measure of frequency selectivity: Relation to psychophysical equivalent rectangular bandwidth,'' J. Acoust. Soc. Am. **93,** 3291–3297.

Brown, A. M., and Kemp, D. T. (**1984**). ''Suppressibility of the $2f_1 - f_2$ stimulated acoustic emissions in gerbil and man,'' Hear. Res. **13**, 29–37.

Brown, A. M., and Kemp, D. T. (**1985**). ''Intermodulation distortion in the cochlea: Could basal vibration be the major cause of round window CM distortion?'' Hear. Res. **19**, 191–198.

Brown, A. M., Gaskill, S. A., and Williams, D. M. (**1992**). ''Mechanical filtering of sound in the inner ear,'' Proc. R. Soc. London Biol. **250**, 29–34.

Brown, A. M., and Williams, D. M. (**1993**). ''A second filter in the cochlea,'' in *Proceedings of the International Symposium on Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, River Edge, NJ), pp. 72–77.

Brown, A. M., Sheppard, S. L., and Russell, P. T. (**1994**). ''Differences between neonate and adult cochlear mechanical responses,'' Aud. Neurosci. **1**, 169–181.

Brown, D., Kimberley, B., and Eggermont, J. (**1994**). ''Cochlear traveling-wave delays estimated by distortion-product emissions in normal hearing adults and term-born neonates,'' J. Otolaryngol. **23**, 234–238.

Cohen, Y. E., Doan, D. E., Rubin, D. M., and Saunders, J. C. (**1993**). ''Middle-ear development V: Development of umbo sensitivity in the gerbil,'' Am. J. Otolarygol. **14**, 191–198.

Eggermont, J. J. (**1979**). ''Narrow-band AP latencies in normal and recruiting human ears,'' J. Acoust. Soc. Am. **65,** 463–470.

Fahey, P. F., and Allen, J. B. (**1986**). ''Characterization of the cubic intermodulation distortion products in the cat external auditory meatus,'' in *Peripheral Auditory Mechanisms: Proceedings of a Conference Held at Boston University, Boston, MA, August 13–16, 1985*, edited by J. B. Allen, J. L. Hall, A. Hubbard, S. T. Neely, and A. Tubis (Springer-Verlag, New York), pp. 314–321.

Gaskill, S. A., and Brown, A. M. (**1990**). ''The behavior of the acoustic distortion product, $2f_1 - f_2$, from the human ear and its relation to auditory sensitivity,'' J. Acoust. Soc. Am. **88,** 821–839.

Goldstein, J. L. (**1967**). ''Auditory nonlinearity,'' J. Acoust. Soc. Am. **41,** 676–689.

Harris, D., and Dallos, P. (**1984**). ''Ontogenic changes in frequency mapping in a mammalian ear,'' Science **225,** 741–743.

He, N.-j., and Schmiedt, R. A. (**1993**). ''Fine structure of the $2f_1 - f_2$ acoustic distortion product: Changes with primary level,'' J. Acoust. Soc. Am. **94,** 2659–2669.

Johnstone, B. M., Patuzzi, R., and Yates, G. K. (**1986**). ''Basilar membrane measurements and the traveling wave,'' Hear. Res. **22**, 147–153.

Keefe, D. H., Bulen, J. C., Arehart, K. H., and Burns, E. M. (**1993**). ''Ear-canal impedance and reflection coefficient in human infants and adults,'' J. Acoust. Soc. Am. **94,** 2617–2638.

Kimberley, B. P., Brown, D. K., and Eggermont, J. J. (**1993**). ''Measuring human cochlear traveling wave delay using distortion product emission phase responses,'' J. Acoust. Soc. Am. **94,** 1343–1350.

Kummer, P., Janssen, T., and Arnold, W. (**1995**). ''Suppression turning characteristics of the $2f_1 - f_2$ distortion-product otoacoustic emission in humans,'' J. Acoust. Soc. Am. **98,** 197–210.

Long, G. R., and Tubis, A. (**1988a**). ''Investigations into the nature of the association between threshold microstructure and otoacoustic emissions,'' Hear. Res. **36**, 125–138.

Long, G. R., and Tubis, A. (**1988b**). ''Modification of spontaneous and evoked otoacoustic emissions and associated psychoacoustic microstructure by aspirin consumption,'' J. Acoust. Soc. Am. **84,** 1343–1353.

Mahoney, C. F. O., and Kemp, D. T. (**1995**). ''Distortion product otoacoustic emission delay measurement in human ears,'' J. Acoust. Soc. Am. **97,** 3721–3735.

Martin, G. K., Lonsbury-Martin, B. L., Probst, R., Scheinin, S. A., and Coats, A. C. (**1987**). ''Acoustic distortion products in rabbit ear canal. II. Sites of origin revealed by suppression contours and pure-tone exposures,'' Hear. Res. **28**, 191–208.

Norton, S. J., Bargones, J. Y., and Rubel, E. W (**1991**). ''Development of otoacoustic emissions in gerbil: Evidence for micromechanical changes underlying development of the place code,'' Hear. Res. **51**, 73–92.

Mills, D. M., Norton, S. J., and Rubel, E. W (**1993**). ''Vulnerability and

adaptation of distortion product otoacoustic emissions to endocochlear potential variation,'' J. Acoust. Soc. Am. **94,** 2108–2122.
Mills, D. M., Norton, S. J., and Rubel, E. W (**1994**). ''Development of active and passive mechanics in the mammalian cochlea,'' Auditory Neurosci. **1**, 77–99.
Mills, D. M., and Rubel, E. W (**1994**). ''Variation of distortion product otoacoustic emissions with furosemide injection,'' Hear. Res. **77**, 183–199.
Mills, D. M., and Rubel, E. W (**1996**). ''Development of the cochlear amplifier,'' J. Acoust. Soc. Am. **100,** 428–441.
Norton, S. J., and Rubel, E. W (**1990**). ''Active and passive ADP components in mammalian and avian ears,'' in *The Mechanics and Biophysics of Hearing: Proceedings of a Conference Held at the University of Wisconsin, Madison, WI, June 25–29, 1990*, edited by P. Dallos, C. D. Geisler, J. W. Mathews, M. A. Ruggero, and C. R. Steele (Springer-Verlag, New York), pp. 219–227.
Plonsey, R., and Collin, R. E. (**1961**). *Principles and Applications of Electromagnetic Fields* (McGraw-Hill, New York).
Robles, L., Ruggero, M. A., and Rich, N. C. (**1986**). ''Basilar membrane mechanics at the base of the chincilla cochea. I. Input-output functions, tuning curves, and response phases,'' J. Acoust. Soc. Am. **80,** 1364–1374.
Rubel, E. W (**1978**). ''Ontogeny of structure and function in the vertebrate auditory system,'' in *Handbook of Sensory Physiology*, edited by M. Jacobson (Springer-Verlag, New York), pp. 135–237.
Rubel, E. W, and Ryals, B. M. (**1983**). ''Development of the place principle: Acoustic trauma,'' Science **219,** 512–514.
Rübsamen, R., Mills, D. M., and Rubel, E. W (**1995**). ''Effects of furosemide on distortion product otoacoustic emissions and on neuronal responses in the anteroventral cochlear nucleus,'' J. Neurophys. **74**, 1628–1638.
Sanes, D. H., Merickel, M., and Rubel, E. W (**1989**). ''Evidence for an alteration of the tonotopic map in the gerbil cochlea during development,'' J. Comp. Neurol. **279,** 436–444.
Schweitzer, L., Lutz, C., Hobbs, M., and Weaver, S. P. (**1996**). ''Anatomical correlates of the passive properties underlying the developmental shift in the frequency map of the mammalian cochlea,'' Hear. Res. **97**, 84–94.
Tashenberger, G., Gallo, L., and Manley, G. A. (**1995**). ''Filtering of distortion-product otoacoustic emissions in the inner ear of birds and lizards,'' Hear. Res. **91**, 87–92.
Walsh, E. J., and Romand, R. (**1992**). ''Functional development of the cochlea and the cochlear nerve,'' in *Development of Auditory and Vestibular Systems II*, edited by R. Romand (Elsevier, New York), pp. 161–219.
Whitehead, M. L., McCoy, M. J., Lonsbury-Martin, B. L., and Martin, G. K. (**1995a**). ''Dependence of distortion-product otoacoustic emissions on primary levels in normal and impaired ears. I. Effects of decreasing L_2 below L_1,'' J. Acoust. Soc. Am. **97,** 2346–2358.
Whitehead, M. L., Stagner, B. B., McCoy, M. J., Lonsbury-Martin, B. L., and Martin, G. K. (**1995b**). ''Dependence of distortion-product otoacoustic emissions on primary levels in normal and impaired ears. II. Asymmetry in L_1,L_2 space,'' J. Acoust. Soc. Am. **97,** 2359–2377.
Woolf, N. K., and Ryan, A. F. (**1984**). ''The development of auditory function in the cochlea of the mongolian gerbil,'' Hear. Res. **13**, 277–283.
Woolf, N. K., Ryan, A. F., and Harris, J. P. (**1986**). ''Development of mammalian endocochlear potential: Normal ontogeny and effects of anoxia,'' Am. J. Physiol. **250,** R493–R498.
Yancey, C., and Dallos, P. (**1985**). ''Ontogenic changes in cochlear characteristic frequency at a basal turn location as reflected in the summating potential,'' Hear. Res. **18**, 189–195.
Zweig, G. (**1991**). ''Finding the impedance of the organ of Cŏrti,'' J. Acoust. Soc. Am. **89,** 1229–1254.

A time-domain, level-dependent auditory filter: The gammachirp

Toshio Irino[a)]
NTT Basic Research Laboratories, 3-1 Morinosato Wakamiya, Atsugi-shi, Kanagawa, 243-01 Japan

Roy D. Patterson[b)]
MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, United Kingdom

(Received 28 June 1996; accepted for publication 4 September 1996)

A frequency-modulation term has been added to the gammatone auditory filter to produce a filter with an asymmetric amplitude spectrum. When the degree of asymmetry in this ''gammachirp'' auditory filter is associated with stimulus level, the gammachirp is found to provide an excellent fit to 12 sets of notched-noise masking data from three different studies. The gammachirp has a well-defined impulse response, unlike the conventional roex auditory filter, and so it is an excellent candidate for an asymmetric, level-dependent auditory filterbank in time-domain models of auditory processing. © *1997 Acoustical Society of America.* [S0001-4966(97)02701-X]

PACS numbers: 43.66.Ba, 43.66.Dc [WJ]

INTRODUCTION

In time-domain auditory models, the spectral analysis performed by the basilar membrane is often simulated by a bank of gammatone auditory filters (see, for example, Patterson *et al.*, 1995). The impulse response of the gammatone is

$$g_t(t)=at^{n-1}\exp(-2\pi b\ \mathrm{ERB}(f_c)t)\cos(2\pi f_c t+\phi) \quad (t>0), \quad (1)$$

where a, b, n, f_c, and ϕ are parameters. $\mathrm{ERB}(f_c)$ is the equivalent rectangular bandwidth of the filter, and at moderate levels $\mathrm{ERB}(f_c)=24.7+0.108f_c$ in Hz (Glasberg and Moore, 1990). The filter gets its name from the fact that the envelope formed by the power function and the exponential is a gamma distribution function, and the cosine carrier is a tone when it is in the auditory range. The amplitude spectrum of the gammatone filter is essentially symmetric on a linear frequency scale.

The gammatone function was introduced by Johannesma (1972) to characterize impulse-response data gathered physiologically from primary auditory fibers in the cat (see Carney and Yin, 1988, for an overview). The gammatone has also been used to characterize spectral analysis in humans at moderate levels where the amplitude characteristic of the auditory filter is nearly symmetric on a linear frequency scale (see Patterson 1994, for an overview).

The use of the gammatone filter is limited, however, by the repeated demonstration that, below its center frequency, the skirt of the auditory filter broadens substantially with increasing stimulus level, and above its center frequency the skirt sharpens a little with increasing level (Lutfi and Patterson, 1984; Patterson and Moore, 1986; Moore and Glasberg, 1987). The level dependence of the auditory filter has been modeled using the ''roex'' function (Patterson *et al.*, 1982; Glasberg and Moore, 1990; Rosen and Baker, 1994). But the roex auditory filter does not have a well-defined impulse response which largely precludes its use in auditory filterbanks. More physiological models of cochlear mechanics (for example, Giguère and Woodland, 1994) do not provide good fits to human masking data; nor do they have sufficiently simple impulse responses for the traditional filterbank architecture.

Irino (1995, 1996) recently demonstrated that an analytic relative of the gammatone function, referred to as the ''gammachirp'' function, is a theoretically optimum auditory filter, in the sense that it leads to minimal uncertainty in a joint time and scale representation of auditory signal analysis. The derivation of the gammachirp function is based on operator methods (Gabor, 1946; Cohen, 1991, 1993) involving the Mellin transform (Titchmarsh, 1948); it is summarized in Appendix A. The gammachirp auditory filter is the real part of the analytic gammachirp function, Eq. (A20). It has an asymmetric amplitude characteristic, and in the following we show that, when the asymmetry is associated with stimulus level, the gammachirp filter provides an excellent fit to human masking data. The gammachirp has a well-defined impulse response and, with only one parameter more than the gammatone, it would appear to be an excellent candidate for an asymmetric, level-dependent auditory filterbank.

I. METHOD

A. The power spectrum model with a gammachirp filter

The impulse response of the gammachirp auditory filter is

$$g_c(t)=at^{n-1}\exp(-2\pi b\ \mathrm{ERB}(f_r)t)\times\cos(2\pi f_r t+c\ \ln t+\phi) \quad (t>0). \quad (2)$$

The only difference between it and the impulse response of the gammatone [Eq. (1)] is the term $c\ \ln t$; c is an additional parameter, and ln is the natural logarithmic operator. The filter has a monotonically frequency-modulated carrier (a chirp) with an envelope that is a gamma distribution func-

[a)]Electronic mail: irino@nttlab.brl.ntt.co.jp; WWW: http://www.brl.ntt.co.jp/people/irino/index.html
[b)]Electronic mail: roy.patterson@mrc-apu.cam.ac.uk

tion, and hence the name ''gammachirp.'' [1] We use f_r instead of f_c for the frequency parameter because the peak frequency of the amplitude spectrum varies with c, and to a lesser extent, b and n. The equivalent rectangular bandwidth of the filter varies with stimulus level, but for convenience, we associate the parameter b with stimulus level so that the basic formula for filter width, $\mathrm{ERB}(f_r)=24.7+0.108f_r$, is the same as in Eq. (1).

The auditory filter shape is derived using the power spectrum model of masking (Fletcher, 1940; Patterson, 1976). In the experiment, the listener is required to detect a brief sinusoidal signal, referred to as a ''probe'' tone, in the presence of a masker which is a noise with a spectral notch in the frequency region of the probe tone. This ''notched noise'' has a constant spectrum level N_0 in a band below the tone between $f_{l_{\min}}$ and $f_{l_{\max}}$ and in a band above the tone between $f_{u_{\min}}$ and $f_{u_{\max}}$. The level of the probe tone is varied to determine the power required to make it just audible (probe ''threshold''), as a function of the width of the notch in the noise. The details of the experiment and the criterion for threshold are described in Patterson (1976). If the ''shape'' of the auditory filter (that is, its power spectrum) is represented by the weighting function, $W(f)$, then the power spectrum model is

$$P_s=K+N_0+10\log_{10}\left(\int_{f_{l_{\min}}}^{f_{l_{\max}}}W(f)\,df+\int_{f_{u_{\min}}}^{f_{u_{\max}}}W(f)\,df\right), \quad (3)$$

where P_s is the power of the probe tone at threshold in dB, and K is a constant which is related to the efficiency of the detection mechanism following the auditory filter. Following Patterson *et al.* (1982), a parameter r is introduced to limit the dynamic range of the filter. The weighting function is associated with the power spectrum of the gammachirp, $|G_C(f)|^2$, as follows:

$$W(f)=(1-r)\cdot W_{om}(f)\cdot|G_C(f)|^2+r. \quad (4)$$

Here, $W_{om}(f)$ is the ''ELC'' correction recommended by Glasberg and Moore (1990) to simulate the effects of the outer and middle ears. The maximum absolute magnitude of $W(f)$ is normalized to unity (See Appendix B for the analytic form of the amplitude spectrum of the gammachirp.)

B. Parameters and fitting procedure

We characterize the level dependence of the auditory filter shape in terms of the level dependence of the five parameters of the gammachirp: n, b, c, K, and r. The auditory filter becomes broader on the low side and sharper on the high side as stimulus level increases (Moore and Glasberg, 1987). Changes in the parameters n and b have little effect on the asymmetry of the amplitude spectrum, and K and r do not affect asymmetry since they are not filter parameters. Thus, the degree of asymmetry is primarily determined by c. Rosen and Baker (1994) showed, in an analogous fit with the roex auditory filter, that the level dependence can be summarized with a linear function. In our initial fit, then, we provide two coefficients for c and one coefficient for each of n, b, and K.

The parameters n and b affect bandwidth reciprocally; the bandwidth of the filter decreases, either when n increases or when b decreases, and vice versa. We are mainly concerned with the filter shape around the center frequency and in this case we can fix either n or b and let the other vary to match the auditory filter shape. Preliminary simulations applied to several data sets and previous work with the gammatone suggested that we begin by fixing n at the value 4.

The fitting procedure is broadly similar to the PolyFit procedure of Rosen and Baker (1994). Thresholds were calculated for a range of filters with center frequencies around the probe frequency. The value of the filter giving the highest signal-to-noise ratio was chosen as the threshold estimate P_s (Patterson and Nimmo-Smith, 1980). We used the Levenberg–Marquardt method (Press *et al.*, 1988) to minimize the squared error between the data and P_s; this is a standard procedure for a nonlinear least-mean-square problem. This fitting procedure is referred to as the ''gammachirp fit'' in the following.

C. Data sets

We began by applying the gammachirp fit to the notched-noise masking data of Rosen and Baker (1994) because the results can be compared directly to their results with the roex filter and the PolyFit procedure. We will specify the data source in the following by the initials of the subject and the probe frequency, for example, ''LM at 2000 Hz'' for this data set which contains 78 tone-in-noise thresholds. Rosen and Baker used total squared error in dB^2 to evaluate alternative fits. We will use rms (root-mean-squared) error in dB; it is a more intuitive measure and makes it easier to compare fits when the data sets have different numbers of thresholds.

The gammachirp was also fitted to subsets of the notched-noise data reported in Lutfi and Patterson (1984) and Moore *et al.* (1990). The data from Lutfi and Patterson are those of HM, RL, RM, and WW at 1000 and 4000 Hz. Each set contains 39 data points distributed over three different noise levels, except for RM at 4000 Hz, where there are 52 data points over four noise levels. The data of Moore *et al.* (1990) are those of CP at 200, 400, and 800 Hz. Each set contains 75 data points distributed over five noise levels.

II. RESULTS

A. Rosen and Baker (1994)

Rosen and Baker (1994) fitted a wide range of roex filter models to a set of masking data gathered with both probe-fixed and masker-fixed conditions to maximize the range of signal levels represented by the data. They discuss a subset of their fits for probe-dependent models with 24, 15, 10, 8, 7, 6, and 5 variable coefficients. The rms errors for the 78 thresholds are 1.15, 1.16, 1.19, 1.19, 1.19, 1.20, and 1.42 dB.[2] The rms errors for masker-dependent models are much greater than these values, and this is why they restricted their attention to probe-dependent models. The focus of their dis-

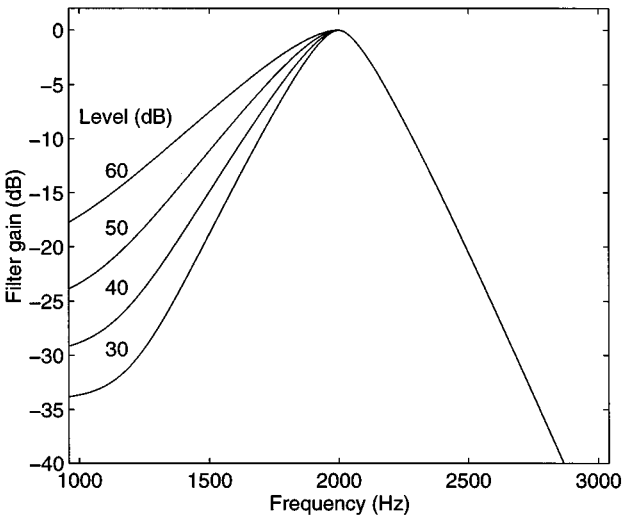


FIG. 1. The roex auditory filter shape as a function of probe level (30–60 dB SPL in 10-dB steps) with six variable coefficients [adapted from Rosen and Baker (1994)].

cussion is the trade-off between number of free parameters and goodness of fit, and they conclude that a roex(p,r) model with six coefficients is the most appropriate. Specifically, their fit employed one parameter for each of p_u and k, and two for each of p_l and r_l. The auditory filter shapes produced by this fit are shown in Fig. 1 as a function of probe level. The lower side of the filter becomes considerably broader as level increases; the upper side is invariant.

Column 7 of Table I shows rms error values (dB) obtained with the gammachirp fit using probe-dependent models with various numbers of coefficients. The integers in columns 2–6 show the number of coefficients used for the gammachirp parameter in that column. Following Rosen and Baker, we used the absolute threshold value (22.7 dB) to limit the minimum value of P_s. We also investigated masker-dependent models, but found, like Rosen and Baker, that the rms errors were always much greater than those with the probe-dependent models. Consequently, we only consider probe-dependent models in what follows. The fit without the parameter c, i.e., the gammatone fit, is shown in the last row of Table I. The rms error is about 30% greater than that of the gammachirp fits, indicating that the gammatone filter is not suitable as an asymmetric, level-dependent filter.

The rms errors in column 7 are the same for gammachirp models with between 4 and 7 variable coefficients. This indicates that the coefficients converge even with relatively few coefficients. The rms errors with the gammachirp are greater than those for roex models with six, seven, and eight variable coefficients (compare columns 7 and 8), and smaller for models with five variable coefficients. The gammachirp model with four variable coefficients, where n is fixed to 4, produces the same rms error as the model with five variable coefficients, where the estimated value of n is 3.89. Accordingly, the model with four variable coefficients seems sufficient to explain the masking data. Rosen and Baker do not report results with a four-coefficient model. The fixed-n model also has advantages when fitting smaller data sets and when comparing coefficients obtained with different data sets.

The coefficients for the four-coefficient model are listed in row ''LM 2000'' in Table II. The auditory filter shapes produced by this fit are shown in Fig. 2 as a function of probe level. In the fitting process, the peak frequency of the amplitude spectrum varies with level, as described previously; for clarity, however, the peak frequency is normalized to 2000 Hz in the figure by adjusting the value of f_r in Eq. (2). Below its peak frequency, the skirt of the gammachirp auditory filter broadens substantially with increasing stimulus level; above its peak frequency, the skirt sharpens a little with increasing level. These shapes are quite similar to the roex filter shapes in Fig. 1, although there are small differ-

TABLE I. Relationship between the number of filter coefficients and rms error. Columns 7 and 8 show rms (root-mean-squared) errors in dB obtained with the gammachirp filter and the roex filter when fitting the probe-dependent model with various numbers of coefficients to all 78 data points in Rosen and Baker (1994). The rms errors in column 8 are calculated from the total squared errors in Rosen and Baker (1994). The integers in the first column show the total number of variable coefficients for both the gammachirp and the roex. The integers in other columns show the number of coefficients used for the gammachirp parameter in that column: ''1'' indicates a filter parameter that is constant across signal level and ''2'' indicates a linear dependence of the parameter on signal level. The symbol ''-'' indicates an r value of −100 dB (practically zero in linear terms); ''*'' indicates an n value of 4; ''—'' indicates no model fitted at that value. The last row shows the results without parameter c, i.e., the gammatone fit.

Number of coefficients	Gammachirp						roex
	n	b	c	K	r	rms error	rms error
10	2	2	2	2	2	1.18	1.19
9	2	2	2	2	1	1.27	—
8	2	2	2	1	1	1.29	1.19
7	2	2	2	1	-	1.33	1.19
6	1	2	2	1	-	1.33	1.20
6	2	1	2	1	-	1.33	—
5	1	1	2	1	-	1.33	1.42
4	*	1	2	1	-	1.33	—
4	1	2	0	1	-	1.72	—

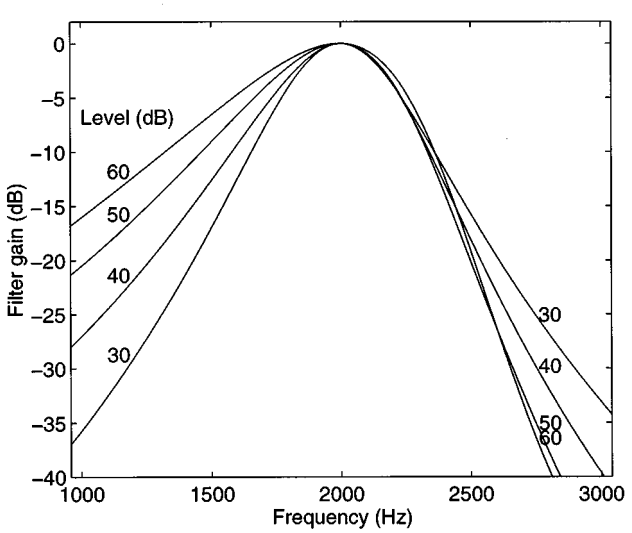


FIG. 2. The gammachirp auditory filter shape as a function of probe level (30–60 dB SPL in 10-dB steps) with four variable coefficients when applied to the masking data of Rosen and Baker (1994). The peak frequency is normalized to 2000 Hz.

ences in the flatness around the peak frequency and in the variability of the upper skirt. Unlike the roex, the upper skirt of the gammachirp has the ''backward S'' shape observed in the dense threshold functions in Patterson (1976). The gammachirp filter naturally introduces the physical constraints of realistic filters into the estimation of the auditory filter shape. The derived filter shapes are also in agreement with those reported in previous studies (Lutfi and Patterson, 1984; Patterson and Moore, 1986; Moore and Glasberg, 1987).

It also appears that we do not need the parameter r when fitting the data of LM; absolute threshold is sufficient limit to the dynamic range of the fitting process. This is another advantage of using the gammachirp fit.

B. Other data sets

Following the results in the previous subsection, we applied several probe-dependent models to the notched-noise masking data of Lutfi and Patterson (1984) and Moore *et al.* (1990). The models with four, five, and six variable coefficients were fitted to each set of data. As before, the model with four variable coefficients proved most appropriate and so we begin with it. Given their limited size, each set of masking data was fitted with seven different sets of initial values; the set that produced coefficients giving minimum mean squared error is listed with the rms error in Table II. It is clear that the value of b converges between 1 and 2 and that c is always negatively correlated with probe level P_s, as in the previous fits for LM at 2000 Hz. The filter shapes are similar to the shapes in Fig. 2 in terms of change in slope with level, except for four conditions: the filter shape is almost level independent for CP at 200 Hz and RM at 4000 Hz; the upper slope changes as much as the lower slope does for HM at 1000 and 4000 Hz. The last two rows in Table II show the means and standard deviations of the parameter values. Since the mean coefficients are close to those for LM, a ''typical'' auditory filter set resembles those shown in Fig. 2 when the peak frequencies are normalized to unity. Thus, the gammachirp with four variable coefficients provides a reasonable summary to the masking data in these data sets, although the rms errors are larger than those for the data set of LM.

For completeness, we also performed the gammachirp fit with five variable coefficients (1 n, 1 b, 2 c's, and 1 K) and six variable coefficients (1 b, 2 c's, 1 K, and 2 r's). Only two of the models with five variable coefficients reduced the rms error more than 5%, reductions that are negligible when compared with the variance in the data sets. Since absolute threshold values were not included in these fits, the model with level-dependent r was also applied to each set (i.e., six

TABLE II. Rms errors and coefficients obtained with the gammachirp auditory filter when fitting a probe-dependent model with four variable coefficients. The first column specifies the data source by the initials of the subject. CP represents data from Moore *et al.* (1990); HM, RL, RM, and WW represent data from Lutfi and Patterson (1984); LM represents data from Rosen and Baker (1994). The second column is probe frequency in Hz. The third column shows rms error in dB. The remaining columns show the best coefficients for b, c, and K with $n=4$ and $r=-100$ (dB). The last two rows show the means and standard deviations for b, c, and K.

Subject	Frequency	rms error	b	c		K
CP	200	4.72	1.19	−0.59	$-0.0097P_s$	2.82
CP	400	2.92	1.43	2.64	$-0.082P_s$	−1.15
CP	800	2.65	1.75	2.16	$-0.070P_s$	−4.35
HM	1000	4.04	1.17	8.43	$-0.180P_s$	−3.25
RL	1000	4.46	1.59	4.98	$-0.146P_s$	−11.70
RM	1000	2.93	1.21	5.27	$-0.148P_s$	−7.25
WW	1000	3.46	1.38	3.56	$-0.098P_s$	−6.64
LM	2000	1.33	1.68	3.38	$-0.107P_s$	−6.08
HM	4000	4.14	1.85	6.31	$-0.153P_s$	−7.67
RL	4000	5.10	1.75	5.17	$-0.182P_s$	−14.39
RM	4000	4.94	1.50	0.61	$-0.019P_s$	−3.63
WW	4000	2.75	1.79	4.18	$-0.110P_s$	−5.87
mean	—	—	1.51	3.88	$-0.109P_s$	−5.73
s.d.	—	—	0.24	2.46	0.057	4.51

variable coefficients). All of the derived values for the parameter r were negatively correlated with probe level P_s and were smaller than -40 dB, even when the signal level was 30 dB SPL. That is, the values are less than absolute threshold, and so, absolute threshold is a more suitable limit to the dynamic range when fitting these data. Moreover, the rms error was reduced more than 10% in only three cases. Thus, parameter r does not seem necessary to explain the general form of the masking data, and the model with four variable coefficients seems sufficient to explain these masking data, as well as those of Rosen and Baker.

III. SUMMARY

A ''gammachirp'' function derived as an optimum auditory filter (Irino, 1995, 1996) is shown to have an asymmetric amplitude characteristic in frequency. Using the power spectrum model of masking, and the assumption that the asymmetry is associated with stimulus level, the amplitude spectrum of the gammachirp was fitted to notched-noise masking data from 12 data sets reported in 3 different studies. A probe-dependent model with four variable coefficients is shown to provide an excellent fit to the masking data. The resultant gammachirp filter shape is similar to that obtained with a six-coefficient roex filter by Rosen and Baker (1994). The gammachirp has a well-defined impulse response unlike the roex auditory filter and, thus, it is an excellent candidate for an asymmetric, level-dependent auditory filterbank in time-domain models of auditory processing.[3]

ACKNOWLEDGMENTS

The authors wish to thank Brian C. J. Moore and Brian R. Glasberg of Cambridge University, and Stuart Rosen and Richard J. Baker of University College, London for providing their notched-noise masking data, as well as fruitful discussions. The first author wishes to thank Takesi Okadome of NTT BRL for the suggestion to study optimality of the auditory filter.

APPENDIX A: THE DERIVATION OF THE GAMMACHIRP FUNCTION

The gammachirp function arose from consideration of the contrast between the traditional representation of sound, the spectrogram, and the representation produced by auditory filterbanks designed to mimic the spectral processing of the cochlea. The contrast is set out in Sec. A of this Appendix. It has led to the hypothesis that the time-frequency representation of sound observed at the output of the cochlea is an intervening representation produced by the auditory system to support a subsequent ''scale transform,'' and that the function that minimizes uncertainty in the time-scale representation is the gammachirp. The scale transform and the gammachirp function are the subjects of Secs. B and C, respectively.

A. The spectrogram and auditory filtering

The spectrogram is a typical example of a joint time-frequency representation of sound. It is produced by converting successive segments of the sound wave into spectral frames with a Fourier transform. There is a trade-off between the resolution of time and the resolution of frequency in this representation. The trade-off is known as the uncertainty principle, and Gabor (1946) showed that the function which satisfies minimal uncertainty in the joint time-frequency representation is a complex sinusoidal carrier with a Gaussian envelope. This ''Gabor function'' is symmetric in time and symmetric in frequency; moreover, the frequency bands all have the same width in this time-frequency representation.

The spectral analysis produced by auditory filtering differs significantly from that produced by the Fourier transform: The impulse response of the auditory filter is asymmetric in time with a fast rise and a slow decay (de Boer and de Jongh, 1975; Carney and Yin, 1988); the amplitude spectrum of the auditory filter is definitely not Gaussian (Patterson, 1976), and at high sound levels, it is asymmetric with the lower skirt shallower than the upper skirt (Glasberg and Moore, 1990). The gammatone function [Eq. (1)] provides a much better fit to auditory filtering data than the Gabor function. It is clear, however, that to the extent that the gammatone differs from the Gabor function, it does not satisfy minimum uncertainty in a joint time-frequency representation of sound. Moreover, the bandwidth of the auditory filter increases with center frequency; in the region above about 500 Hz, it is essentially a ''constant-Q system,'' that is, bandwidth is proportional to center frequency (Greenwood, 1990; Glasberg and Moore, 1990).

It is possible that the auditory system is non-optimal because it has to satisfy some mechanical or physiological constraint that is not compatible with minimal uncertainty, and which restricts the bandwidth to be a proportion of the center frequency. On the other hand, it seemed reasonable, on encountering the discrepancy between optimality and auditory filtering, to explore the possibility that the auditory system is optimal, but optimal for a different representation of sound. It is this hypothesis that led to the ''scale transform'' and the derivation of the gammachirp function.

B. Scale analysis

Cohen (1991, 1993) has suggested that ''scale'' is a physical attribute of a signal just like time and frequency, and that a time-scale representation is more appropriate than Fourier analysis for ''scaled'' signals. A ''scaled'' signal is simply one that is compressed or extended in time relative to the original, as when a tape recording is replayed at a rate faster or slower than that at which it was recorded. Cohen (1991) introduced a ''scale transform'' in the form of an orthogonal Mellin transform (Titchmarsh, 1948) to produce the scale representation. It is described in Sec. C. The Mellin transform converts a scaled signal into (a) an invariant absolute distribution in the scale representation, and (b) a value specifying the scale value of the signal. When the speed of a recording fluctuates on playback, it has a pronounced effect on the pitch we hear, but the source of the sound is not perceived to change. This suggests that the ''scale'' value and the ''invariant distribution'' of the Mellin transform may be analogous to pitch and timbre in the auditory system. Thus, the time-scale representation of sound could have distinct advantages when analyzing systems where a vibrating

source with variable rate excites a complex resonator. ''Source-filter'' models of this sort are commonly used to explain the production of sound by the vocal tract (Fant, 1970) and musical instruments (Fletcher and Rossing, 1991).

The scale transform can be applied to a sound wave directly. It is clear, however, that in the auditory system, the scale transform would have to be applied after auditory filtering. The wavelet transform is similar to the auditory filterbank inasmuch as it is a ''constant-Q'' system; both the envelope and the carrier of the impulse response scale with center frequency in these systems. The Mellin transform converts the individual wavelets into an invariant distribution in the scale representation. Thus, with a wavelet filterbank, when a sound is scaled, its components shift to wavelet filters that have been scaled by the same amount. So, the outputs of the scaled filters are exactly the same as the scaled versions of the outputs of the original filters. Both the scaled and unscaled filter outputs are transformed into the same distribution in the scale representation. Thus, the wavelet filterbank is ''transparent'' to the scaling of sounds, and in this sense, the wavelet transform is optimal as a preprocessor for the Mellin transform. This implies that the auditory filterbank would be a near optimum preprocessor for the Mellin transform.

The optimal relationship between the Mellin transform and the wavelet transform does not uniquely determine the form of the wavelet that produces minimal uncertainty in a joint time-scale representation; Clearly, the Gabor function does not; After Cohen (1991, 1993) adapted Klauder's (1980) results on affine variables in quantum mechanics to produce the scale representation for time domain functions, he showed that the optimal function for minimal uncertainty in a time-scale representation has a gamma envelope and a monotonically frequency-modulated carrier. In Cohen's case, the instantaneous frequency of the carrier starts at infinity and converges on zero as time proceeds. This solution is not suitable for relatively narrow-band applications like auditory filtering. This led Irino (1995, 1996) to introduce a frequency shift term that makes it possible to model the bandwidth/center-frequency function of the auditory system, and produce a narrow-band filter centered on a specific frequency. This, in turn, led to the derivation of the gammachirp function through the optimality constraint. This, then, is the logic for the time-scale representation of sound and the gammachirp function.

C. Mathematical derivation

a. The Mellin transform

The Mellin transform (Titchmarsh, 1948) of a signal, $s(t)$ $(t>0)$, is defined as

$$S(p)=\int_0^\infty s(t)t^{p-1}\,dt, \tag{A1}$$

where p is a complex argument. One of the important properties is

$$\text{if } s(t)\Rightarrow S(p), \text{ then } s(at)\Rightarrow a^{-p}S(p), \tag{A2}$$

where the arrow ($\Rightarrow$) indicates ''is transformed into'' and a is a real dilation constant. That is, the distribution $S(p)$ is just multiplied with a constant a^{-p} when the function $s(t)$ is scaled in time. If p is denoted by p_r+jp_i,

$$a^{-p}=a^{-(p_r+jp_i)}=a^{-p_r}a^{-jp_i}=a^{-p_r}\exp(-j\ \ln p_i), \tag{A3}$$

where $j=\sqrt{-1}$, and exp and ln are the exponential and natural logarithmic operators. Since $|a^{-p}S(p)|=|a^{-p_r}|\cdot|S(p)|$, the absolute distribution $|S(p)|$ is not affected by a scaling of the signal, except for the constant that specifies the scale of the current signal; nor is it affected when the distribution is normalized.

b. Minimal uncertainty and operator methods

With the Mellin transform, questions concerning minimal uncertainty in a joint representation are assessed with operator methods. They were introduced into signal processing from quantum mechanics by Gabor (1946) because of the similarity in mathematical formalism. The following is a tutorial on operator methods based on the derivation of the Gabor function; it is adapted from Cohen (1991, 1993).

Time and frequency operators are defined as $\mathscr{T}=t$ and $\mathscr{W}=-j(d/dt)$ in the time domain. When the operator $\mathscr{W}$ is applied to the function $Ae^{j\omega t}$, the result is

$$\mathscr{W}\,Ae^{j\omega t}=\left(-j\,\frac{d}{dt}\right)Ae^{j\omega t}=\omega\,Ae^{j\omega t}. \tag{A4}$$

Thus, for a complex exponential, the operator $\mathscr{W}$ introduces the frequency term ω. This is the essence of operator methods. The commutator between these operators is again an operator; namely,

$$[\mathscr{T},\mathscr{W}]=\mathscr{T}\mathscr{W}-\mathscr{W}\mathscr{T}=t\left(-j\,\frac{d}{dt}\right)-\left(-j\,\frac{d}{dt}\right)t=j. \tag{A5}$$

It is easy to prove by applying this operator to the function $Ae^{j\omega t}$. Since the commutator is not zero, time and frequency do not commute. Thus, time and frequency cannot be measured independently and there is uncertainty between them, and in this case, it is

$$\Delta t\cdot\Delta\omega\geqslant\tfrac{1}{2}|\langle[\mathscr{T},\mathscr{W}]\rangle|=\tfrac{1}{2}|\langle j\rangle|=\tfrac{1}{2}, \tag{A6}$$

where $(\Delta.)$, $|.|$, and $\langle.\rangle$ denote the standard deviation, the absolute value, and the average, respectively. Functions which satisfy minimal uncertainty are solutions to the equation

$$(\mathscr{W}-\langle\mathscr{W}\rangle)s(t)=\lambda(\mathscr{T}-\langle\mathscr{T}\rangle)s(t), \tag{A7}$$

where

$$\lambda=\langle[\mathscr{T},\mathscr{W}]\rangle/2(\Delta\mathscr{T})^2=j/2(\Delta t)^2. \tag{A8}$$

Using $\mathscr{T}=t$, $\mathscr{W}=-j\ d/dt$, $\langle\mathscr{T}\rangle=\langle t\rangle$, and $\langle\mathscr{W}\rangle=\langle\omega\rangle$, Eq. (A7) is expanded as follows:

$$\left(-j\,\frac{d}{dt}-\langle\omega\rangle\right)s(t)=\lambda(t-\langle t\rangle)s(t);$$

$$j\,\frac{d}{dt}\,s(t)+\lambda ts(t)+(\langle\omega\rangle-\lambda\langle t\rangle)s(t)=0;$$

$$\frac{d}{dt}\,s(t)+\frac{1}{2(\Delta t)^2}\,ts(t)+\left(-j\langle\omega\rangle-\frac{\langle t\rangle}{2(\Delta t)^2}\right)s(t)=0. \tag{A9}$$

The nontrivial solution is

$$s(t)=a\,\exp\left(-\frac{1}{4(\Delta t)^2}\,t^2+\frac{\langle t\rangle}{2(\Delta t)^2}\,t+j\langle\omega\rangle t\right)$$

$$=a'\,\exp\left\{-\frac{1}{4(\Delta t)^2}\,(t-\langle t\rangle)^2\right\}\exp(j\langle\omega\rangle t), \tag{A10}$$

where a and a' are constants. This is the ''Gabor function,'' and the example shows how it was derived using the constraint that the required function satisfy minimal uncertainty in a joint time-frequency representation.

c. The Mellin operator

Cohen (1991, 1993) introduced the concept of a scale operator into signal processing in the form:

$$\mathscr{C}=\tfrac{1}{2}(\mathscr{T}\mathscr{W}+\mathscr{W}\mathscr{T})=\mathscr{T}\mathscr{W}-\tfrac{1}{2}j. \tag{A11}$$

Previously it had been known as the operator representing an affine variable in quantum mechanics (Klauder, 1980). The corresponding transform, that is, ''the scale transform'' (Cohen, 1993), is

$$D(c)=\frac{1}{\sqrt{2\pi}}\int_0^\infty f(t)t^{-jc-1/2}\,dt. \tag{A12}$$

The correspondence between the scale transform and the Mellin transform is revealed by setting $p=-jc+\frac{1}{2}$ in Eq. (A1). Thus, Cohen's scale transform is the Mellin transform with a specific argument. In Eq. (A12), the argument is restricted in range; we can, however, extend it to cover the entire complex plane by the introduction of two real constants c_0 and μ as follows:

$$p=-j(c-c_0)+(\mu+\tfrac{1}{2}). \tag{A13}$$

The corresponding Mellin operator is

$$\mathscr{C}_m=\mathscr{T}\mathscr{W}+\{c_0+j(\mu-\tfrac{1}{2})\}. \tag{A14}$$

Since we are concerned with signal processing by an auditory filterbank, we introduce a ''frequency-shift'' term ω_0 into the operator to specify the individual filters. The form of the operator becomes

$$\mathscr{C}_a=\mathscr{T}(\mathscr{W}-\omega_0)+\{c_0+j(\mu-\tfrac{1}{2})\}. \tag{A15}$$

The frequency-shift term can be removed later following consideration of the fluctuation of components at the output of the auditory filter (Irino, 1996). The commutator between time and this operator is

$$[\mathscr{T},\mathscr{C}_a]=[\mathscr{T},\mathscr{C}_m]=[\mathscr{T},\mathscr{C}]=j\mathscr{T}. \tag{A16}$$

The operators in Eqs. (A14) and (A15) are not Hermitian except when $\mu=0$; nevertheless, $(\mathscr{C}_a-\langle\mathscr{C}_a\rangle)$ is Hermitian and, thus, the eigenvalue is real. The function that satisfies minimal uncertainty between time and the quantity represented by the operator in Eq. (A14) is the solution to the equation

$$(\mathscr{C}_a-\langle\mathscr{C}_a\rangle)s(t)=\lambda(\mathscr{T}-\langle t\rangle)s(t), \tag{A17}$$

where

$$\lambda=\langle[\mathscr{T},\mathscr{C}_a]/2(\Delta\mathscr{T})^2\rangle=j\langle t\rangle/2(\Delta t)^2. \tag{A18}$$

Equation (A15) expands to

$$t\left(-j\,\frac{d}{dt}\right)s(t)-(\omega_0+j\alpha_1)ts(t)+(-c_1+j\alpha_2)s(t)=0, \tag{A19}$$

where $\alpha_1=\langle t\rangle/2(\Delta t)^2$, $\alpha_2=\mu-\frac{1}{2}-\mathrm{Im}\langle c_a\rangle+\langle t\rangle^2/2(\Delta t)^2$, and $c_1=\mathrm{Re}\langle c_a\rangle-c_0$, Re and Im indicate the real and imaginary parts. The solution is

$$s(t)=at^{\alpha_2+jc_1}\,\exp(-\alpha_1 t+j\omega_0 t)$$

$$=at^{\alpha_2}\,\exp(-\alpha_1 t)\exp(j\omega_0 t+jc_1\,\ln t), \tag{A20}$$

where a is a constant. The envelope $t^{\alpha_2}\exp(-\alpha_1 t)$is a gamma distribution function $\gamma(t)$. The instantaneous frequency is ω_0+c_1/t; that is, a fractional function of time. When played as a sound, the carrier would be a chirp, and hence the name ''gammachirp'' function. When $c_1=0$, Eq. (A20) becomes a gammatone function. Thus, the gammatone function is a first order approximation to the gammachirp function.

APPENDIX B: THE AMPLITUDE SPECTRUM OF THE GAMMACHIRP FUNCTION

The Fourier spectrum of the gammachirp function can be derived analytically. For convenience, we consider a simplified version of the complex form of the gammachirp filter in Eq. (2).

$$g_c(t)=at^{n-1}\,\exp(-b't)\exp(j\omega_r t+jc\,\ln t)\quad(t>0)$$

$$=at^{n-1+jc}\,\exp(-b't+j\omega_r t)\quad(t>0), \tag{B1}$$

where $b'=2\pi b\,\mathrm{ERB}(f_r)$, $\omega_r=2\pi f_r$, and the phase term ϕ is ignored. The Laplace transform of Eq. (B1) is

$$G_C(s)=a\,\frac{\Gamma(n+jc)}{\{s-(-b'+j\omega_r)\}^{n+jc}}$$

$$=a\,\frac{\Gamma(n+jc)}{|s-(-b'+j\omega_r)|^{n+jc}e^{j\theta\cdot(n+jc)}}$$

$$=a\,\frac{\Gamma(n+jc)}{|s-(-b'+j\omega_r)|^{n}e^{-c\theta}\cdot|s-(-b'+j\omega_r)|^{jc}e^{jn\theta}}, \tag{B2}$$

where $\theta=\arg\{s-(-b'+j\omega_r)\}$. Thus, the absolute value is

$$|G_C(s)|=\frac{|a\Gamma(n+jc)|}{|s-(-b'+j\omega_r)|^{n}e^{-c\theta}}. \tag{B3}$$

Substituting $s=j\omega=j2\pi f$ into Eq. (B3) to derive the amplitude of the Fourier spectrum of the gammachirp function,

$$|G_C(f\,)|=\frac{|a\Gamma(n+jc)|}{|b'+j2\pi(f-f_r)|^{n}}\cdot e^{c\theta}, \tag{B4}$$

where $\theta=\arg\{b'+j2\pi(f-f_r)\}$.

[1]Several auditory filters with gamma distribution envelopes and monotonically frequency-modulated (FM) carriers have appeared recently. First, Lyon (1996) has reported an ''all-pole gammatone filter (APGF)'' based on reduction of zeros from the Laplace transform of the gammatone filter in the s plane. In this case, the intent was to simulate basilar partition motion

'as a function of stimulus level. Although the impulse response of the APGF is not mathematically equivalent to the gammachirp function, it is similar in having a monotonic FM carrier and a gamma distribution envelope. Second, in an attempt to produce an asymmetric gammatone filter, Baker (1995) replaced the pure-tone carrier with a monotonic FM carrier. Again, the result is not strictly a gammachirp function, but the impulse response has an FM carrier and a gamma distribution envelope. Finally, Laine and Härmä (1996) have suggested similar filters for an auditory filterbank on the Bark scale. The gammachirp function can be viewed as providing the theoretical background to the larger family of auditory filters with gamma distribution envelopes and chirp carriers.

[2]In their paper Rosen and Baker (1994) presented total-squared-error values of 103.6, 105.0, 110.9, 110.9, 111.0, 111.9, and 158.1 dB^2, respectively, for these conditions.

[3]To construct such a filterbank, we would need to develop a mechanism to measure the output level of each filter on a moment to moment basis to specify the appropriate value of c, and thus the filter's asymmetry, at any given moment. A mechanism of this sort has been developed by Lyon (1982) for a nonlinear filterbank simulating cochlear mechanisms. Thus, it would not appear to be an insurmountable problem to develop one for a gammachirp filterbank. It is, however, beyond the scope of the present paper.

Baker, R. (**1995**). Personal communication.

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.

Cohen, L. (**1991**). ''A general approach for obtaining joint representations in signal analysis and an application to scale,'' in *Advanced Signal-Processing Algorithms, Architectures, and Implementation II*, edited by T. L. Franklin, Proc. SPIE, **1566**, 109–133.

Cohen, L. (**1993**). ''The scale representation,'' IEEE Trans. Signal Process. **41**, 3275–3292.

de Boer, E., and de Jongh, H. R. (**1978**). ''On cochlear encoding: Potentialities and limitations of the reverse-correlation technique,'' J. Acoust. Soc. Am. **63,** 115–135.

Fant, G. (**1970**). *Acoustic Theory of Speech Production* (Mouton, The Hague).

Fletcher, H. (**1940**). ''Auditory patterns,'' Rev. Mod. Phys. **12,** 47–65.

Fletcher, N. H., and Rossing, T. D. (**1991**). *The Physics of Musical Instruments* (Springer-Verlag, New York).

Gabor, D. (**1946**). ''Theory of communication,'' J. IEE (London) **93**, 429–457.

Giguère, C., and Woodland, P. C. (**1994**). ''A computational model of the auditory periphery for speech and hearing research. I. Ascending path,'' J. Acoust. Soc. Am. **95,** 331–342.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hear. Res. **47**, 103–138.

Greenwood, D. D. (**1990**). ''A cochlear frequency-position function for several species—29 years later,'' J. Acoust. Soc. Am. **87,** 2592–2605.

Irino, T. (**1995**). ''An optimal auditory filter,'' in *Proc. IEEE Signal Processing Society, 1995 Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY.

Irino, T. (**1996**). ''A 'gammachirp' function as an optimal auditory filter with the Mellin transform,'' in Proc. IEEE Int. Conf. Acoust., Speech Signal Processing (ICASSP-96), **2**, 981–984, Atlanta, GA.

Johannesma, P. I. M. (**1972**). ''The pre-response stimulus ensemble of neurons in the cochlear nucleus,'' in *Symposium on Hearing Theory* (IPO, Eindhoven, Holland), pp. 58–69.

Klauder, J. R. (**1980**). ''Path integrals for affine variables,'' in *Functional Integration: Theory and Applications*, edited by J. P. Antoine and E. Tirapgui (Plenum, New York).

Laine, U. K., and Härmä, A. (**1996**). ''On the design of Bark-FAMlet Filterbanks,'' in Proc. Nordic Acoustical Meeting, Helsinki, Finland.

Lutfi, R. A., and Patterson, R. D. (**1984**). ''On the growth of masking asymmetry with stimulus intensity,'' J. Acoust. Soc. Am. **76,** 739–745.

Lyon, R. F. (**1982**). ''A computational model of filtering, detection, and compression in the cochlea,'' in Proc. IEEE Int. Conf. Acoust., Speech Signal Processing (ICASSP-82), 1282–1285, Paris, France.

Lyon, R. F. (**1996**). ''The all-pole gammatone filter and auditory models,'' in Proc. Forum Acusticum '96, Antwerp, Belgium.

Moore, B. C. J., and Glasberg, B. R. (**1987**). ''Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns,'' Hear. Res. **28**, 209–225.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1990**). ''Auditory filter shapes at low center frequencies,'' J. Acoust. Soc. Am. **88**, 132–140.

Patterson, R. D. (**1976**). ''Auditory filter shapes derived with noise stimuli,'' J. Acoust. Soc. Am. **59,** 640–654.

Patterson, R. D. (**1994**). ''The sound of a sinusoid: Spectral models,'' J. Acoust. Soc. Am. **96,** 1409–1418.

Patterson, R. D., Allerhand, M., and Giguère, C. (**1995**). ''Time-domain modelling of peripheral auditory processing: a modular architecture and a software platform,'' J. Acoust. Soc. Am. **98,** 1890–1894.

Patterson, R. D., and Moore, B. C. J. (**1986**). ''Auditory filters and excitation patterns as representations of frequency resolution,'' in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London).

Patterson, R. D., and Nimmo-Smith, I. (**1980**). ''Off-frequency listening and auditory filter asymmetry,'' J. Acoust. Soc. Am. **67,** 229–245.

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.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (**1988**). *Numerical Recipes in C* (Cambridge U.P., Cambridge).

Rosen, S., and Baker, R. J. (**1994**). ''Characterising auditory filter nonlinearity,'' Hear. Res. **73**, 231–243.

Titchmarsh, E. C. (**1948**). *Introduction to the Theory of Fourier Integrals* (Oxford U.P., London), 2nd ed.

Detectability of simultaneously masked signals as a function of masker bandwidth and configuration for different signal delays[a)]

Beverly A. Wright[b)]
Keck Center, Box 0732, University of California at San Francisco, San Francisco, California 94143-0732

(Received 15 December 1995; revised 11 July 1996; accepted 16 September 1996)

Signal detectability was measured in three temporal conditions as a function of the bandwidth and configuration of simultaneous maskers that either did or did not spectrally overlap the signal. The 20-ms signal was 250 Hz wide and was centered at 2500 Hz (f_s). Although there were marked individual differences, performance was typically poorer when signal onset came 1 ms rather than 250 ms after the onset of a 420-ms masker, and poorest when signal onset came 1 ms after the onset of a 23-ms masker. The results support the idea that two separate across-channel processes contribute to temporal changes in signal detectability. One process contributes to the improvement observed as signal onset is delayed from masker onset, and its influence is reduced by the presence of masking components at f_s only when the masker extends exclusively below f_s. The other process is associated with the improvement observed as masker offset is delayed from signal offset, and its influence is reduced by the presence of masking components at f_s when the masker extends exclusively above, or both below and above f_s. Both of these processes are primarily activated by frequencies ranging from 0.6 to $0.8f_s$ and 1.2 to $1.4f_s$. The data also demonstrate that the measured critical bandwidth narrows as signal onset is delayed from masker onset. © *1997 Acoustical Society of America.* [S0001-4966(97)04901-1]

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Mk [WJ]

INTRODUCTION

The amount of signal energy required for the detection of a brief simultaneously masked signal often decreases as signal onset is delayed from masker onset. This phenomenon will be referred to as a ''temporal effect'' here, but has also been denoted by terms such as ''overshoot'' (e.g., Zwicker, 1965a, b) and ''enhancement'' (e.g., Viemeister, 1980). This paper reports how the magnitude of the temporal effect measured with maskers that do and do not spectrally overlap the signal is influenced by the masker bandwidth and configuration, the choice of reference condition, and individual differences. Experiments reported by Wright (1995a) in a companion paper employed the same subjects and addressed the same general themes but focused on manipulations of signal rather than masker bandwidth.

Temporal effects are particularly interesting because they are associated with temporal changes in frequency selectivity. Such dynamic changes in resolution may strongly influence the perception of real-world sounds such as speech. Furthermore, a deeper understanding of temporal effects in normal-hearing listeners may aid in the diagnosis and treatment of hearing disorders, because temporal effects are much smaller in hearing-impaired than normal-hearing listeners (e.g., Kimberley *et al.*, 1989).

A. Masker bandwidth and configuration

Thirty years ago, Elliott (1965, 1967, 1969) suggested that two distinct mechanisms each produce a temporal effect. Supporting her idea, there are now two lines of evidence that different processes are evoked depending upon whether the signal and masker pass primarily through the same (within-channel), or separate (across-channel), cochlear filters.

One line of evidence is that any given stimulus manipulation tends to have a similar influence on the temporal effects measured with narrow-band maskers centered at the signal frequency (on-frequency narrow-band maskers) and wideband maskers, but a different influence with maskers presented only at frequencies remote from the signal (off-frequency maskers). Thus the influence of a particular stimulus manipulation on the temporal effect depends upon whether or not there are masking components in the narrow *within-channel* frequency region surrounding the signal. For example: (1) As the masker level is increased, the magnitude of the temporal effect initially grows and then declines for on-frequency narrow-band (Carlyon and Moore, 1986) and wideband (e.g., Bacon, 1990) maskers, but steadily grows for off-frequency maskers (Viemeister, 1980); (2) To reset signal detectability to its normal value at masker onset requires that only a brief silent period be inserted before the signal in otherwise continuous on-frequency narrow-band (Elliott, 1969) and wideband (e.g., Zwicker, 1965a) maskers, but requires a much longer gap for off-frequency maskers (e.g., Elliott, 1969; Viemeister, 1980); (3) The magnitude of the temporal effect increases with signal frequency for on-frequency narrow-band (e.g., Carlyon and Moore, 1986) and wideband (e.g., Zwicker, 1965a; Elliott, 1967) maskers, but generally remains constant for off-frequency maskers (e.g., Bacon and Moore, 1986); (4) Asymptotic performance is reached when the signal is delayed from masker onset by 100–200 ms for on-frequency narrow-band (e.g., Bacon and

[a)]Portions of this research were presented at the 121st Meeting of the Acoustical Society of America [J. Acoust. Soc. Am. **89**, 1914(A) (1991)].
[b)]Current address: Audiology and Hearing Sciences Program, Northwestern University, 2299 North Campus Drive, Evanston, IL 60208-3550.

Viemeister, 1985a) and wideband maskers (e.g., Elliott, 1965; Zwicker, 1965a), but requires a delay greater than 250 ms for off-frequency maskers (e.g., Viemeister, 1980; McFadden and Wright, 1990).

The second line of evidence that both within- and across-channel mechanisms are involved in the temporal effect is that, regardless of the influence of any particular stimulus manipulation, the overall magnitude of the temporal effect is generally larger for wideband and off-frequency maskers than for on-frequency narrow-band maskers. Thus the overall magnitude of the temporal effect depends upon whether or not there are masking components in the *across-channel* frequency regions remote from the signal. It appears that these regions encompass a broad range of frequencies above about 1.2 times, and below about 0.8 times the signal frequency, because the largest temporal effects are obtained when there are masking components within those regions (for a review, see Wright, 1995b). Both the within- and across-channel processes are frequently proposed to involve physiological adaptation, but it is not yet clear whether it is excitation (e.g., Smith and Zwislocki, 1975) or inhibition (e.g., Viemeister and Bacon, 1982) that is adapting.

Several previous studies have shown that the temporal-effect magnitude increases with increases in the bandwidth of a masker centered on the signal frequency (e.g., Zwicker, 1965b; Bacon and Smith, 1991). One other examination indicated that the temporal-effect magnitude initially decreases and then remains fairly constant as the bandwidth of the lower and upper bands of a notched-noise masker is increased, holding the masker spectrum level constant (Hicks and Bacon, 1992). There have also been numerous investigations showing that temporal effects are larger when the masker frequency is higher, as opposed to lower, than the signal frequency. This dependence on the relative frequencies of the masker and signal holds for tonal maskers (e.g., Green, 1969; Leshowitz and Cudahy, 1975; Bacon and Viemeister, 1985b; Kimberley *et al.*, 1989) and for high-pass and low-pass noise maskers that overlap the signal frequency (Schmidt and Zwicker, 1991). There has, however, been no systematic investigation of the combined effect of the bandwidth and relative frequency of the masker. One purpose of this paper is to report data on the influence of masker bandwidth and configuration on the magnitude of the temporal effect measured with maskers that do and do not spectrally overlap the signal. The results support the idea that both within- and across-channel processes are involved in the effect, and provide information about the interaction of the two processes. The changes in signal detectability across variations in masker bandwidth also indicate that the critical bandwidth is wider at masker onset, consistent with the conclusion originally reached by Scholl (1962).

B. The reference condition

A signal presented at the onset of a notched-noise masker can be much easier to detect when the masker outlasts the signal (short-delay condition) than when it does not (burst condition; Wright 1991, 1995a; Kidd and Wright, 1994). These detectability differences are much larger with wide-notched (Wright, 1995a) than with wideband maskers (Elliott, 1965; Zwicker, 1965a; Wright, 1995a). They thus appear to indicate the existence of an across-channel process that is sensitive to masker offsets in frequency regions remote from the signal, especially when the masker does not also overlap the signal (Wright, 1995a). These detectability differences also show that the temporal-effect magnitude can depend strongly upon whether performance in the burst or short-delay condition is used as the reference for calculating the size of the effect. The second purpose of this paper is to present additional data showing performance differences between the burst and short-delay conditions.

C. Individual differences

Wright (1991, 1995a) reported marked and systematic individual differences in the magnitudes of the temporal effects measured with both the burst and short-delay references. Such variations in performance in temporal-effect tasks have been emphasized in a number of recent papers (McFadden and Wright, 1990, 1992; Wright *et al.*, 1993) and were again revealed in an examination of 40 naive listeners (Wright, 1996). The third purpose of this paper is to report that the individual differences observed with the present manipulations of masker bandwidth and configuration were quite similar to those observed with the same subjects when signal bandwidth was varied (Wright, 1995a), indicating that these differences are quite robust.

I. EXPERIMENT 1: MASKERS SPECTRALLY REMOTE FROM THE SIGNAL

A. Method

1. Stimuli

The signal was a noise band, 250 Hz wide, centered at 2500 Hz. The masker was a noise band of variable width that extended in frequency either below (low-pass), above (high-pass), or both below and above (notched) the signal frequency. The low-pass masker was 0 (a tone), 250, 500, or 1900 Hz wide, and its highest component was fixed at 2000 Hz. The high-pass masker was 0 (a tone), 250, 500, 1900, or 4900 Hz wide, and its lowest component was fixed at 3000 Hz. The notched masker was created by combining low-pass and high-pass maskers that had the same bandwidth, yielding, in each case, a masker with a 1000-Hz wide notch centered at the signal frequency. Each notched masker will be identified by the bandwidth of its constituent components.

All waveforms were gated (Wilsonics, model BSIT) using a cosine-squared rise-decay time of 10 ms. The total duration of the signal was always 20 ms and that of the masker was either 23 or 420 ms. In the short-delay and long-delay conditions, the signal was gated 1 or 250 ms, respectively, after the onset of the 420-ms masker. In the burst condition, the signal was gated 1 ms after the onset of the 23-ms masker.[1] The magnitude of the short-delay temporal effect was calculated by subtracting signal detectability in the long-delay from that in the short-delay condition, and the magnitude of the burst temporal effect was calculated by subtracting signal detectability in the long-delay from that in the burst condition.

In most experiments in which masker bandwidth has been manipulated, the spectrum level of the masker was held constant, and the overall level was allowed to change (e.g., Fletcher, 1940). In the present experiment, however, the overall level of the low-pass and high-pass maskers was held constant at about 58 dB, and the spectrum level was allowed to decrease from about 58 dB with the 0-Hz bandwidth to 21 dB with the 4900-Hz bandwidth. The overall level of the notched masker was fixed at 61 dB so that the spectrum level for each particular masker bandwidth would be the same across the asymmetrical and notched maskers. The decision to hold overall level constant was aimed at minimizing the confound between increases in masker bandwidth and increases in overall level, because the latter are known to increase the magnitudes of both the short-delay and burst temporal effects obtained with notched and high-pass maskers (e.g., Viemeister, 1980; Bacon and Viemeister, 1985b).[2]

The signals and maskers were digitally synthesized and delivered to separate 16-bit D/A converters (using a 20-kHz sampling rate) followed by 10-kHz low-pass filters. Individual samples of brief narrow-band sounds can have very different sound-pressure levels. Therefore, individual waveform samples of the signal and masker (for each masker width, duration and configuration) were screened until 28 were found that covered a range in overall level of less than 1.25 dB. The particular signal and masker samples to be presented on a given observation interval were chosen at random, with replacement, from the appropriate set of 28 waveforms. For details, see Wright (1995a).

2. Procedure

All listening was through the left ear, over TDH-39 headphones mounted in circumaural cushions. The psychophysical task was two-interval forced-choice with feedback. The signal level was adjusted adaptively, using the three-down/one-up rule of Levitt (1971) which estimates the 79% correct point on the psychometric function. The final estimate of signal detectability in each of the temporal conditions was based on the mean of a minimum of five 60-trial blocks per subject. The mean within-subject standard error calculated across subjects and masker configurations was greatest in the burst condition at 1.5 dB. Mean between-subject standard errors are plotted in the figures. For details of the procedure, see Wright (1995a).

3. Subjects

Seven normal-hearing subjects were paid for their participation. These were the same subjects who had served in the experiments described by Wright (1995a).

B. Results

1. Partitioning of subjects

Figure 1 shows the overall signal level required for detection with the low-pass (left column), high-pass (middle column), and notched (right column) maskers. Mean results are plotted separately for three subsets of the seven subjects (top three rows) and for all seven subjects together (bottom row). Membership in each subject group was determined by

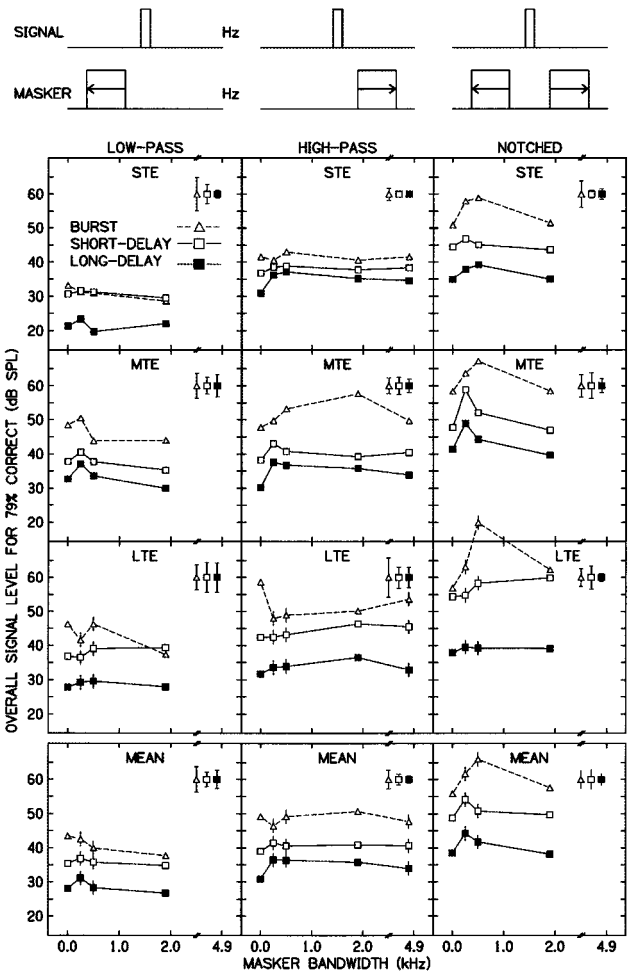


FIG. 1. Signal detectability in overall level plotted as a function of the bandwidth of the low-pass (left column), high-pass (middle column), and notched (right column) maskers. A schematic spectrum of the signal and masker is shown at the top of each column. The parameter is the temporal condition: burst (open triangles), short-delay (open squares), and long-delay (filled squares). The mean results are shown separately for the two Small-Temporal-Effect (STE; top row), the three Mixed-Temporal-Effect (MTE; second row), and the two Large-Temporal-Effect (LTE; third row) subjects, and for all seven subjects together (MEAN; bottom row). The error bars at the upper right of each panel represent the mean standard error between subjects for each of the conditions. The 20-ms signal was centered at 2500 Hz and was 250 Hz wide. The masking component nearest to the signal was fixed at 2000 Hz for the low-pass maskers and 3000 Hz for the high-pass maskers. The low- and high-pass maskers were combined to form the notched maskers. The *overall* level of the masker was held constant at about 58 dB SPL for the asymmetrical maskers, and about 61 dB for the notched maskers. The vertical lines marking some of the points in the bottom two rows indicate values based on the performance of only one of the two LTE subjects.

the magnitudes of the burst and short-delay temporal effects shown by each subject when the signal bandwidth was manipulated in the companion experiment (Wright, 1995a). In that experiment, two subjects consistently showed small burst and short-delay temporal effects. These will be called the Small-Temporal-Effect (STE) subjects (top row). Two other subjects consistently showed large burst and short-delay temporal effects. These will be called the Large-Temporal-Effect (LTE) subjects (third row). The remaining three subjects showed large burst temporal effects, but small short-delay temporal effects. These will be called the Mixed-

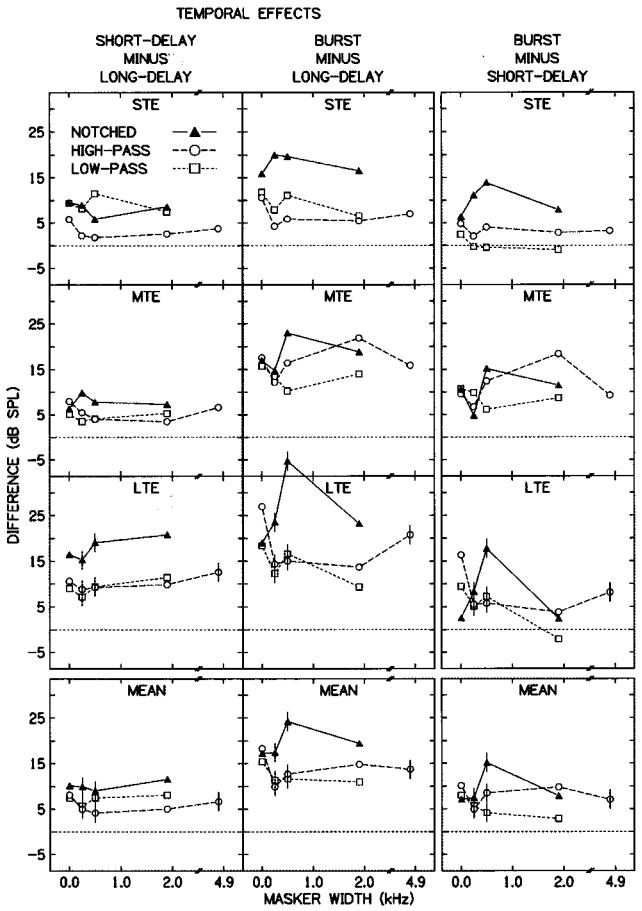


FIG. 2. The short-delay temporal effects (left column), burst temporal effects (middle column), and burst/short-delay differences (right column) plotted as a function of the masker bandwidth for the STE (top row), MTE (second row), and LTE (third row) subjects, and for all seven subjects together (MEAN; bottom row). The parameter is the masker configuration: low-pass (open squares), high-pass (open circles), and notched (filled triangles). The values were calculated from the data in Fig. 1 using the computations listed at the top of each column.

Temporal-Effect (MTE) subjects (second row). For clarity, the mean between-subject standard error for each condition is represented by a single error bar at the upper right of each panel. In Figs. 1 and 2, the values at intermediate masker bandwidths marked with a vertical line in the bottom two rows of panels reflect the performance of only one of the two LTE subjects. This is because time limitations prevented the complete retesting of the other LTE subject following her dramatic improvement in performance after months of practice (see Wright, 1991, pp. 253–259).

2. Signal levels

Three general patterns emerge from the measured signal levels plotted in Fig. 1. First, with few exceptions, detectability was poorest in the burst condition (open triangles), intermediate in the short-delay condition (open squares), and best in the long-delay condition (filled squares). Second, within each panel, the influence of increasing masker bandwidth on detectability tended to be similar for the short-delay and long-delay conditions, with the poorest performance typically occurring at the bandwidth of 250 Hz. For the burst condition, however, performance was poorest at masker bandwidths either narrower (low-pass maskers) or wider (high-pass and notched maskers) than 250 Hz. Third, the notched maskers generally produced the most masking, followed by the high-pass and then the low-pass maskers. The greater effectiveness of high-pass than low-pass maskers at low masker levels has been observed previously (e.g., Zwicker and Jaroszewski, 1982; Patterson and Moore, 1986).

3. Temporal effects

Figure 2 shows the group and overall mean magnitudes of the short-delay (left column) and burst (middle column) temporal effects calculated from the data in Fig. 1. The parameter is the masker configuration. The overall mean magnitude of the short-delay temporal effect (bottom left panel) remained relatively constant across masker bandwidths. Despite a 33-dB decrease in the masker spectrum level, as the masker bandwidth was increased from 0 to 1900 Hz the short-delay temporal effect only decreased by about 3 dB for the high-pass maskers (open circles), remained essentially unchanged for the low-pass maskers (open squares), and actually increased by about 2 dB for the notched maskers (filled triangles). In contrast, over the same range of masker bandwidths, the burst temporal effect (bottom middle panel) decreased by about 5 dB for both the high-pass and low-pass maskers and was clearly largest at a masker bandwidth of 500 Hz for the notched maskers. Common to both reference conditions, the overall mean magnitude of the temporal effect was nearly always largest for the notched maskers. The subject groups differed primarily in their relative performances across the different masker configurations and in the magnitudes of their effects. Just as reported by Wright (1995a), the short-delay temporal effect was largest for the LTE subjects (7–21 dB, 7 values $>$10 dB), and was smaller and similar for the STE (2–12 dB, 1 value $>$10 dB) and MTE (3–10 dB, 0 values $>$10 dB) subjects, and the burst temporal effect was typically larger and fairly similar for the MTE (10–23 dB, 13 values $>$10 dB) and LTE (9–37 dB, 12 values $>$10 dB) subjects, and was smaller for the STE subjects (4–20 dB, 7 values $>$10 dB).

4. Burst/short-delay differences

The differences in detectability between the burst and short-delay conditions calculated from the data in Fig. 1 are plotted in the right column of Fig. 2. These burst/short-delay differences had a similar pattern to that of the burst temporal effect, but their overall magnitudes were necessarily somewhat smaller than the burst temporal effect. The differences ranged from about -1 to 14 dB for the STE subjects (2 values $>$10 dB), 5 to 18 dB for the MTE subjects (6 values $>$10 dB), and -2 to 18 dB for the LTE subjects (2 values $>$10 dB). Consistent with their categorization as ''mixed,'' the MTE subjects generally showed the largest burst/short-delay differences, particularly for the wider masker bandwidths.

5. Individual differences

In the companion experiment (Wright, 1995a), the marked variations across subjects in the magnitudes of their

temporal effects arose primarily from performance differences in the two reference conditions; all subjects performed similarly in the long-delay condition. The same pattern occurred in the present experiment for the high-pass maskers. For the low-pass and notched maskers, however, the subjects differed in sensitivity in all three temporal conditions. Performance was typically best for the STE subjects, and poorer for the MTE and LTE subjects.

II. EXPERIMENT 2: MASKERS SPECTRALLY OVERLAPPING THE SIGNAL

A. Method

In experiment 2, both the bandwidth and configuration of a masker that spectrally overlapped the signal were manipulated. As in experiment 1, the signal was a noise band, 250 Hz wide, centered at 2500 Hz, and the masker bandwidth was either increased below (low-pass), above (high-pass), or both below and above (bandpass) the signal frequency. The low-pass masker was 630, 1130, or 2530 Hz wide, and its highest component was fixed at 2630 Hz (5 Hz above the upper edge of the 250-Hz-wide signal). The high-pass masker was 630, 1130, 2530, or 5530 Hz wide, and its lowest component was fixed at 2370 Hz (5 Hz below the lower edge of the signal). The bandpass masker was formed by combining low-pass and high-pass maskers having the same bandwidth, but without doubling the components at the signal frequency. The bandpass masker was 260, 1000, 2000, 4800, or 7900 Hz wide.

Here, the masker spectrum level was held constant at 25 dB, and the overall level was allowed to increase from about 49 to 64 dB as the masker bandwidth was increased from 260 to 7900 Hz. This was done to be consistent with other similar experiments (e.g., Zwicker, 1965b; Bacon and Smith, 1991), and because the magnitude of the temporal effect obtained with maskers that overlap the signal depends upon the masker spectrum level (e.g., Keiser, 1980; Carlyon and Moore, 1986; Bacon, 1990). In this experiment, only the signal and the 23-ms maskers were screened so that the individual waveform samples covered a range in overall level of less than 1.25 dB. Otherwise, the generation of stimuli, the procedure and the subjects were the same as for experiment 1. Just as in experiment 1, the mean within-subject standard error calculated across subjects and masker configurations was greatest in the burst condition at 1.2 dB.

B. Results

1. Signal levels

The data for the low-pass (left column), high-pass (middle column), and bandpass (right column) maskers are plotted in Fig. 3, in a manner parallel to Fig. 1. Note that the ranges on the ordinate and abscissa differ from Fig. 1. For comparison, the results for the bandpass masker that was 260 Hz wide are plotted in every panel at the narrowest masker bandwidth on each function. The signal was nearly always hardest to detect in the burst condition (open triangles). Signal detectability in the short-delay condition (open squares) was typically poorer than in the long-delay condition (filled squares) for the wider masker bandwidths, but that pattern

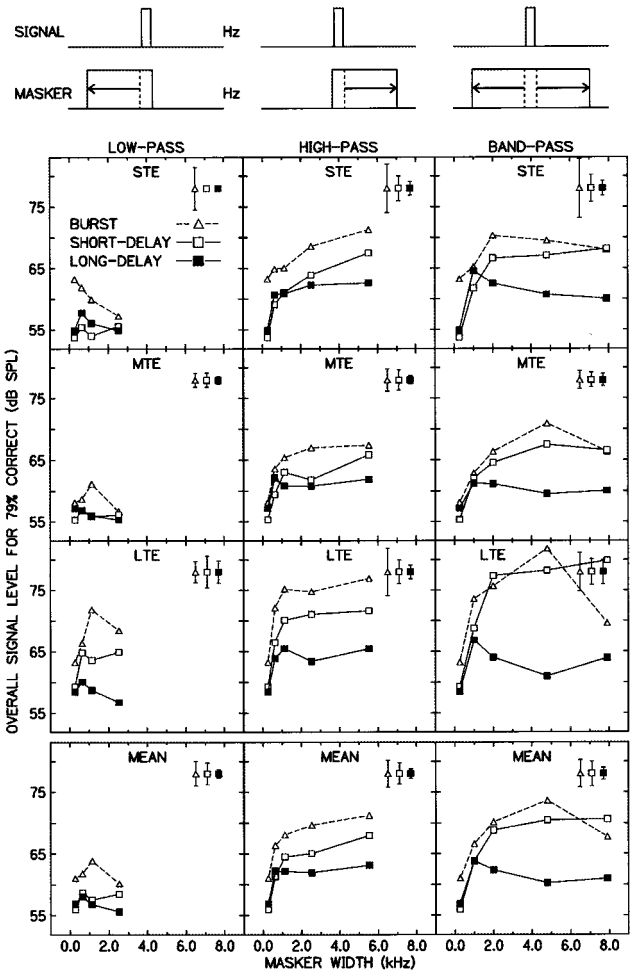


FIG. 3. Plotted as in Fig. 1, but for a masker that spectrally overlapped the signal. The upper edge of the masker was fixed at 2630 Hz for the low-pass masker and the lower edge of the masker was fixed at 2370 Hz for the high-pass masker. The low- and high-pass maskers were combined to form the bandpass maskers. The *spectrum* level of the maskers was held constant at 25 dB SPL.

reversed at the narrower masker bandwidths for both the STE and MTE subjects (also see Bacon and Smith, 1991). As masker bandwidth increased, detectability decreased rapidly up to a particular bandwidth, which varied across conditions, and typically remained relatively constant thereafter. As in experiment 1 in which the maskers were always spectrally remote from the signal, here sensitivity was generally poorest with the bandpass maskers, intermediate with the high-pass maskers, and best with the low-pass maskers for all three groups of subjects.

2. Temporal effects

Plotted in Fig. 4 are the magnitudes of the short-delay (left column) and burst (middle column) temporal effects calculated from the data in Fig. 3. The values are displayed above the actual bandwidths for the low-pass (open squares) and high-pass (open circles) maskers. The values for the bandpass masker (filled triangles) are plotted above the bandwidths for the asymmetrical maskers that had the same remote cutoff frequencies; for example, the results for the bandpass masker that was 2000 Hz wide are plotted with the

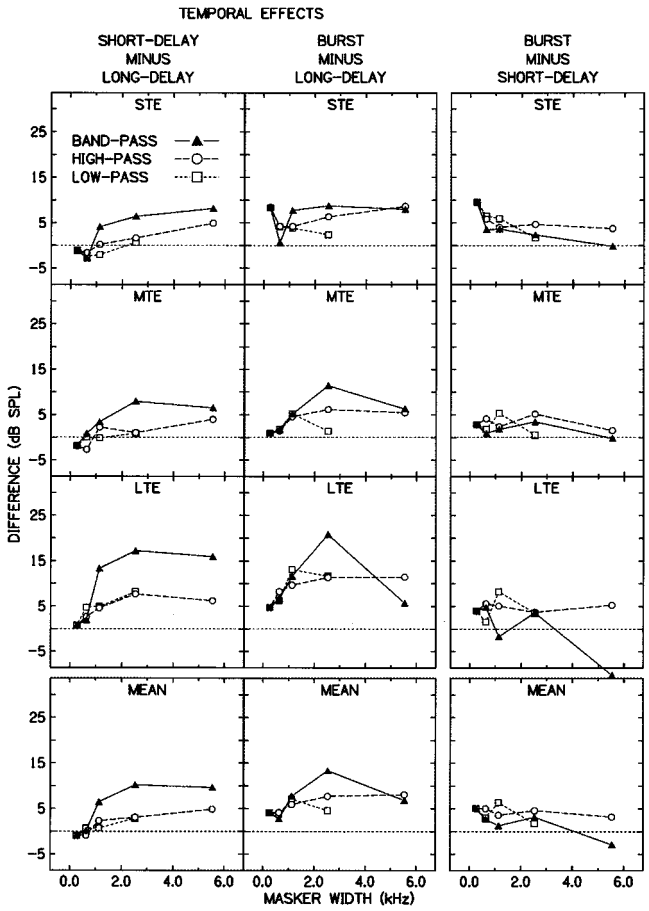


FIG. 4. Parallel to Fig. 2, but for a masker that spectrally overlapped the signal. The parameter is the masker configuration: low-pass (open squares), high-pass (open circles), and bandpass (filled triangles). The values were calculated from the data in Fig. 3.

data for the asymmetrical maskers that were 1130 Hz wide, because the remote cutoff frequencies of the bandpass and the asymmetrical maskers were the same (1500 and 3500 Hz).

The magnitudes of both the short-delay and burst temporal effects typically increased with increases in masker bandwidth for all three masker configurations. The short-delay temporal effect was actually negative at the narrowest bandwidths for the STE and MTE subjects (also see, e.g., Bacon and Smith, 1991), but the burst temporal effect was always positive. The temporal effects obtained with both reference conditions were generally similar across all three masker configurations at the narrower masker bandwidths. However, at the wider masker bandwidths, the effects were nearly always largest for the bandpass maskers, and were smaller, and similar, for the low-pass and high-pass maskers. Both the short-delay and burst temporal effects were largest for the LTE subjects (2–21 dB, of bandwidths wider than 260 Hz, 17 were >5 dB) and smaller, and similar, for the STE ($-3-+9$ dB, 7 were >5 dB) and MTE ($-3-+11$ dB, 8 were >5 dB) subjects. These results are thus in good accord with those reported by Wright (1995a) for the same subjects when the masker was a wideband noise and the signal bandwidth was manipulated.

3. Burst/short-delay differences

The right column of Fig. 4 shows the burst/short-delay differences calculated from the data in Fig. 3. These differences were relatively small—generally less than 6 dB—at all masker bandwidths in all three masker configurations, just as had been previously reported for wideband bandpass maskers (e.g., Elliott, 1965; Zwicker, 1965a; Wright, 1995a).

4. Individual differences

There were consistent individual differences in detectability in all three temporal conditions. In all cases, sensitivity was better, and similar, for the STE and MTE subjects, and was worse for the LTE subjects. The differences across the groups were smaller in the long-delay than in the other two conditions. This pattern of results resembles that obtained from the same subjects by Wright (1995a) using other maskers that spectrally overlapped the signal.

III. DISCUSSION

A. Evidence for within- and across-channel contributions to the temporal effect and burst/short-delay difference

One purpose of the present experiments was to determine whether the results obtained across variations in masker bandwidth and configuration formed a pattern consistent with the idea that both a within- and an across-channel process are involved in the temporal effect. The data support the two-mechanism hypothesis by showing that the magnitude of the temporal effect depended upon the masker spectrum. Maskers that presumably primarily stimulated the within-channel process, because they only had components at the signal frequency, consistently yielded small or even negative temporal effects (narrow maskers in Fig. 4). In contrast, maskers that presumably primarily stimulated the across-channel process, because they only had components remote from the signal frequency, produced temporal effects that were larger and always positive (Fig. 2).

The remainder of this section describes how the temporal effects and burst/short-delay differences measured when the across-channel process was activated were influenced by the simultaneous activation of the within-channel process. For convenience, in this discussion, all maskers that had components at the signal frequency (experiment 2) will be referred to as on-frequency maskers, and all maskers that did not will be referred to as off-frequency maskers.

Figure 5 shows the overall mean magnitudes of the short-delay (left column) and burst (middle column) temporal effects and of the burst/short-delay differences (right column) obtained with the on-frequency (filled symbols) and off-frequency (open symbols) maskers. The data are replotted from Figs. 2 and 4. The abscissa indicates the masker bandwidth either below 2000 Hz ($0.8f_s$) for the low-pass maskers (top row) or above 3000 Hz ($1.2\ f_s$) for the high-pass (middle row) and symmetrical (bottom row) maskers. The frequencies of 2000 and 3000 Hz were the inner-most masking components in experiment 1, and the remote edges of the masker bandwidths of 630 and 1000 Hz in experiment 2. Hence, as the masker bandwidth was increased below

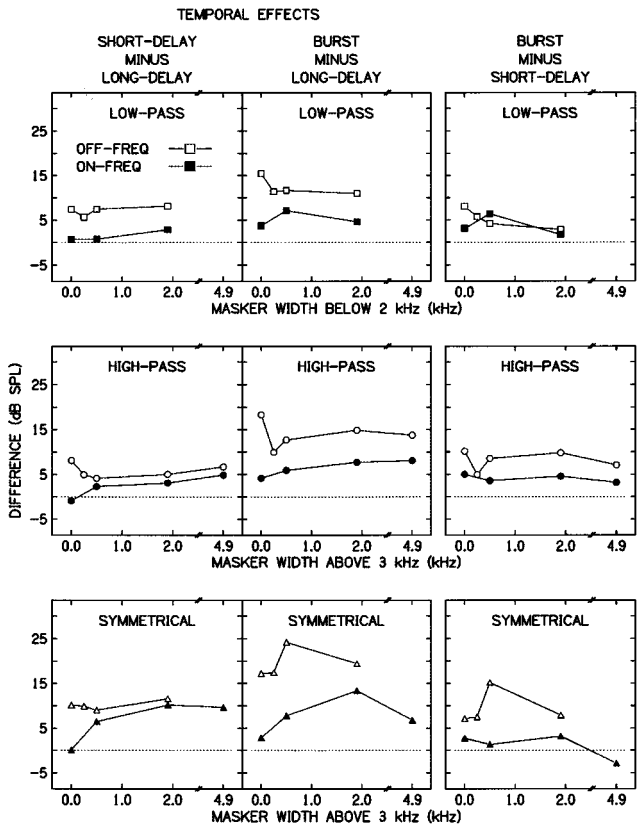


FIG. 5. A comparison of the effect of increasing the remote-masker bandwidth of the low-pass (top row), high-pass (middle row), and symmetrical (bottom row) maskers on the overall mean magnitude of the short-delay (left column) and burst (middle column) temporal effects and the burst/short-delay difference (right column). The parameter is the masker type: off-frequency (open symbols; maskers from experiment 1) or on-frequency (filled symbols; maskers from experiment 2). The data are replotted from Figs. 2 and 4.

2000 Hz and/or above 3000 Hz, the same frequency regions remote from the signal frequency were stimulated in both experiments. The only difference was that in experiment 2 there were also masking components at the signal frequency, while in experiment 1 there were not. The masker bandwidths listed on the abscissae of Fig. 5 therefore will be referred to as remote-masker bandwidths. The masker overall level was fixed in experiment 1 (off-frequency maskers) and the masker spectrum level was fixed in experiment 2 (on-frequency maskers). Nevertheless, the spectrum levels of the maskers in the two experiments were within 6 dB of each other for remote-masker bandwidths of 500 Hz and greater, and were the same for the 1900-Hz bandwidth.

For the low-pass maskers (top row),[3] the temporal effects measured with both reference conditions were always about 6 dB larger with the off-frequency than the on-frequency maskers at remote-masker bandwidths of 500 Hz and greater. These data demonstrate that there is an across-channel region below the signal frequency capable of producing temporal effects (also see Carlyon, 1989; Hicks and Bacon, 1992). They also show that the ability of this across-channel region to yield a temporal effect is considerably reduced when there is also masking energy in the within-channel region. Consistent with this conclusion, Schmidt and Zwicker (1991) observed only small burst temporal effects with their low-pass on-frequency masker. In contrast to the two temporal effects, the burst/short-delay difference obtained with the low-pass maskers was quite similar across the two masker types at the wider masker bandwidths. This outcome indicates that these burst/short-delay differences depended entirely on the masking components in the across-channel region, because adding masking energy to the within-channel region had little influence on the effect.

A different pattern of results emerged for the high-pass (middle row) and symmetrical (bottom row) maskers. For these two masker configurations, the magnitudes of all three difference scores were always larger with the off-frequency than the on-frequency maskers. However, for remote-masker bandwidths of 500 Hz and greater, the difference in detectability across the two masker types was only about 2 dB for the short-delay temporal effect, but was always greater than 7 dB for the burst temporal effect and burst/short-delay difference. These data suggest that the temporal effects and burst/short-delay differences obtained with both the high-pass and symmetrical maskers were produced by masking components in across-channel regions located above (high-pass maskers) or both above and below (symmetrical maskers) the signal frequency. They also show that the burst temporal effect and burst/short-delay difference are considerably reduced when the within- and across-channel regions are simultaneously activated. Wright (1995a, her Fig. 6) reports parallel results for signals of variable width presented in symmetrical on-frequency and off-frequency maskers.

One additional aspect of the results obtained with the symmetrical maskers deserves mention. Unlike for the other masker configurations, the burst temporal effects and burst/short-delay differences obtained in the symmetrical configuration differed in form as well as magnitude across the on-frequency and off-frequency maskers (also see Wright, 1995a). Thus the combination of the across-channel regions below and above the signal frequency changed both the pattern and size of the results. These differences support the suggestion of Wright (1991, 1995a) and Kidd and Wright (1994) that there may be a separate across-channel process that is sensitive to masker offsets especially when the masker does not also spectrally overlap the signal.

B. Masker bandwidth

Masker bandwidth influenced the magnitude of the temporal effect particularly for the high-pass and symmetrical configurations. One of the clearest examples of this was that both the short-delay and burst temporal effects grew by approximately 10 dB as the bandwidth of the bandpass on-frequency masker was increased to 2530 Hz (Fig. 4, filled triangles). This growth presumably represents the transition into activation of the across-channel process.

Increasing the masker bandwidth within the across-channel region generally lead first to a rapid, and then to a slower, growth in the temporal-effect magnitude. This pattern is most obvious for the on-frequency maskers for which the masker spectrum level was held constant. The largest growth in the size of the temporal effect usually occurred as the remote-masker bandwidth was increased from 0 to 500

Hz (Fig. 5, filled symbols). The lower and upper frequency edges of the remote maskers at these bandwidths correspond to the ratios of 0.6 and $0.8f_s$ and 1.2 and $1.4f_s$. These same frequency ranges were previously concluded to be the most influential within the across-channel region based upon a review of the literature (Wright, 1991) and the results of a more recent variation of the bandwidth of off-frequency maskers (Hicks and Bacon, 1992). As the remote-masker bandwidth was increased beyond 500 Hz, the temporal-effect magnitude typically continued to grow, but at a much slower rate (Fig. 5). A similar pattern can be seen in other data (e.g., Zwicker, 1965b; Bacon and Smith, 1991; Schmidt and Zwicker, 1991). These results indicate that although the most influential across-channel regions range from 0.6 to $0.8f_s$ and 1.2 to $1.4f_s$, there are measurable across-channel contributions from frequencies much more remote from the signal.

The most unusual result from the manipulation of masker bandwidth was that both the burst temporal effect and the burst/short-delay difference measured with symmetrical off-frequency maskers were clearly largest at a remote-masker bandwidth of 500 Hz (Fig. 5, bottom middle and bottom right panels). Therefore it appears that the frequency regions between 0.6 and $0.8f_s$ and 1.2 and $1.4f_s$ are also the most important for the across-channel mechanism that is sensitive to masker offsets, but that simultaneously stimulating frequencies below $0.6f_s$ and above $1.4f_s$ may actually inhibit the action of that across-channel process. Hicks and Bacon (1992) examined the burst temporal effect under somewhat similar circumstances. In their data, the effect at 1000 Hz initially decreased and then remained steady as the upper frequency edge of the masker increased from 1.3 to $2.0f_s$. The discrepancy between their results and the present ones may have resulted from any of a number of differences in the stimuli or subjects.

C. Masker configuration

The masker configuration only influenced the magnitudes of the short-delay and burst temporal effects when the masker bandwidth was wide enough to activate the across-channel regions. For these wider on-frequency and off-frequency maskers, the average temporal effects were largest with the symmetrical maskers, and smaller and similar with the low-pass and high-pass maskers (Figs. 2 and 4, bottom rows). These data thus differed from the more common reports of greater temporal effects for high-pass than for low-pass maskers (e.g., burst reference: Bacon and Viemeister, 1985b; short-delay reference: Kimberley *et al.*, 1989). Masker configuration had very little influence on the burst/short-delay differences measured with the on-frequency maskers (Fig. 4, bottom row), but was largest for the notched maskers at a bandwidth of 500 Hz for the off-frequency maskers (Fig. 2, bottom row).

D. Filtering

As in other experiments in which the bandwidth of a masker centered on the signal frequency was manipulated (Zwicker, 1965b; Bacon and Smith, 1991), in experiment 2, the measured critical bandwidth was wider when the signal was presented at masker onset rather than after a delay.[4] When two least-square error lines were fitted to the mean results for the symmetrical maskers (Fig. 3, bottom right panel) in each of the three conditions,[5] the best fitting lines intersected at a critical bandwidth of about 522 Hz in the long-delay condition, 880 Hz in the short-delay condition, and 1600 Hz in the burst condition. The absolute values of these estimates are surely influenced by the use of a 250-Hz-wide noise signal and only five masker bandwidths. Nevertheless, the tendency toward wider critical bandwidths for signals presented at masker onset is clear. The same subjects showed a similar trend when the critical bandwidth was measured by increasing the signal bandwidth within a wideband masker (Wright, 1995a). These data thus provide further evidence that frequency selectivity improves over time (e.g., Scholl, 1962; Kimberley *et al.*, 1989; Carlyon, 1989; Wright, 1992).

E. Individual differences

The individual differences in the magnitudes of the temporal effects and burst/short-delay differences shown by the seven subjects here paralleled those shown by the same subjects in the companion experiment (Wright, 1995a). The consistency of these results suggests that individual differences may provide insights into the mechanisms involved in these effects. Supporting this idea, Wright (1996) recently reported that the magnitude of the burst temporal effect in 40 naive subjects was negatively correlated with the amount of psychophysical two-tone suppression (e.g., Houtgast, 1974) and with a measure of frequency selectivity in forward masking. For these same naive subjects, a single factor appeared to underlie detectability in all of the listening conditions used to calculate the magnitudes of the temporal and suppression effects. Similar results are present in the data of six highly trained subjects (Wright *et al.*, 1993). It is likely that the single factor associated with signal detectability is related to the sharpness of frequency tuning, which, in turn, is often proposed to increase as the amount of suppression increases (e.g., Houtgast, 1974). Thus these individual-difference patterns indicate a link between temporal effects, suppression, and frequency selectivity.

IV. SUMMARY

(1) *Experimental conditions:* Signal detectability was measured in three temporal conditions as a function of the bandwidth and configuration of simultaneous maskers that either did or did not spectrally overlap the signal. In the short-delay and long-delay conditions, the 20-ms signal was gated 1 or 250 ms, respectively, after the onset of a 420-ms masker. In the burst condition, the signal was gated 1 ms after the onset of a 23-ms masker. The signal was 250 Hz wide and was centered at 2500 Hz (f_s).

(2) *Performance in the short-delay versus long-delay conditions:* When the masker included components below $0.8f_s$ and/or above $1.2\ f_s$, the signal was nearly always more difficult to detect in the short-delay than in the long-delay condition. The magnitude of this short-delay temporal effect was similar across maskers that did and did not spectrally

overlap the signal when the masker extended only above or both below and above f_s, but was consistently smaller for maskers that overlapped the signal when the masker extended only below f_s (Fig. 5, left column). These results were taken as evidence that (a) an across-channel process located in frequency regions below and above f_s mediates the improvement in signal detectability as signal onset is delayed from masker onset, and (b) the influence of this across-channel process is reduced by the presence of masking components at f_s only when the masker extends exclusively below f_s.

(3) *Performance in the burst versus short-delay conditions:* When the masker included components below $0.8f_s$ and/or above $1.2\ f_s$, the signal was hardest to detect in the burst condition. The pattern of the detectability difference between the burst and short-delay conditions was the inverse of that observed for the short-delay temporal effect. The magnitude of the burst/short-delay difference was similar for maskers that did and did not spectrally overlap the signal when the masker extended only below f_s, but was smaller for maskers that overlapped the signal when the masker extended only above or both below and above f_s (Fig. 5, right column). These data were taken as evidence that (a) a second across-channel process located in frequency regions below and above f_s mediates the improvement in signal detectability as masker offset is delayed from signal offset, and (b) the influence of this across-channel process is reduced by the presence of masking components at f_s when the masker extends exclusively above, or both below and above f_s.

(4) *Masker bandwidth:* The results from the manipulation of masker bandwidth showed that the frequency regions ranging from 0.6 to $0.8f_s$ and 1.2 to $1.4f_s$ were the most important for the production of the detectability differences between the three temporal conditions (e.g., Fig. 4, filled triangles).

(5) *Masker configuration:* When present, the detectability differences between the three temporal conditions tended to be largest for maskers that extended both below and above f_s, and smaller and similar for maskers that extended either only below or only above f_s (Figs. 2 and 4, bottom rows).

(6) *Narrowing critical bandwidth:* The critical bandwidth measured by increasing the bandwidth of a masker centered on the signal frequency narrowed as signal onset was delayed from masker onset (Fig. 3, bottom right panel).

(7) *Individual differences:* The seven subjects showed marked individual differences that were consistent with their performances in a companion experiment (Wright, 1995a).

ACKNOWLEDGMENTS

This research is a portion of a doctoral thesis submitted to the Graduate School of the University of Texas. The members of my doctoral committee, W. S. Geisler, C. A. Champlin, R. L. Diehl, W. Wilczynski, and particularly my advisor, D. McFadden, gave me valuable advice and consistent support. This manuscript was improved by the comments of A. Kohlrausch and one anonymous reviewer. E. G. Pasanen provided technical assistance. All of these data were collected, and much of the preparation of this manuscript completed, at the Department of Psychology at the University of Texas. Supported in part by research grants 1 R29 DC02997 (Wright) and 5 R01 DC00153 (McFadden) from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

[1]The frequencies of the masking components nearest to the signal and the rise-decay time were chosen, in part, to minimize the physical spread of the masker spectrum into the frequency region of the signal. Bacon and Viemeister (1985b) investigated the influence of masker splatter on signal detectability using stimuli roughly analogous to the tonal masker conditions employed here. They reported that the splatter from a 50-ms, 1250-Hz masker had little influence on the detectability of a 20-ms, 1000-Hz signal presented 15 ms after masker onset, as long as the rise-decay time on the masker was greater than 5 ms.

[2]It was reasoned that if the masker spectrum level were held constant, (1) an increase in the temporal-effect magnitude with increasing masker bandwidth could be due either to the increase in the masker bandwidth, or to the corresponding increase in the overall masker level, and (2) the temporal-effect magnitude presumably could not decrease with increases in the masker bandwidth, because then the overall masker level would correspondingly increase. On the other hand, by holding the overall masker level constant, (1) an increase in the temporal-effect magnitude with increasing masker bandwidth could only be attributable to the increase in the masker bandwidth, because the masker spectrum level would correspondingly decrease, and (2) a decrease in the temporal-effect magnitude could be due either to the increase in the masker bandwidth, or to the corresponding decrease in the masker spectrum level. Thus the choice to hold the overall level constant was admittedly an imperfect solution, but it was judged to be the more conservative approach.

[3]When the masker has a lower frequency than the signal, the magnitude of the burst temporal effect can be reduced, or even made negative, by the addition of a continuous low-pass noise designed to mask the frequency region of the cubic difference tone, or CDT (Leshowitz and Cudahy, 1975; Bacon and Viemeister, 1985a, b). Prompted by these reports, the mode of presentation of a CDT masker was manipulated in a pilot experiment. The results showed that the magnitudes of both the burst and short-delay temporal effects could be reduced, for some subjects, by the addition of a continuous low-pass CDT masker. However, it was also seen that the CDT masker produced a temporal effect itself. Attempts at gating the CDT masker to reduce its contribution to the final temporal effect frequently led to increases in the temporal-effect magnitude, and, when the CDT masker was gated with the signal, it even appeared to impair detectability when it did not spectrally overlap the combination product. Because of these complicated, and sometimes inconsistent results, no combination-product masker was employed in experiment 1. This means that the magnitude of the burst and short-delay temporal effects may have been somewhat exaggerated for the narrower, off-frequency, low-pass, and symmetrical maskers (also see Hicks and Bacon, 1992).

[4]The band-widening estimate of critical bandwidth was applied to the present data for two related reasons. First, the band-widening paradigm is accepted as providing an estimate of critical bandwidth for signals presented in long-delay conditions, so the same analysis technique should be acceptable for signals presented in other temporal conditions. Second, the critical bandwidth has been described as ''a purely empirical phenomenon [defined as] that bandwidth at which subjective responses rather abruptly change (Scharf, 1970, p. 159).'' By that definition, if an abrupt change occurs at different bandwidths in different conditions, the *empirical* critical bandwidth differs. Differences in the critical bandwidth across conditions may result from changes in both within- and across-channel contributions to masking.

[5]The two best fitting lines for each condition were determined using the technique of Bogartz (1968) for fitting two arbitrary straight lines to a set of points. Others, such as Spiegel (1979), have used this method to estimate critical bandwidth. With this method, the data were iteratively divided into two groups, corresponding to the narrower and wider masker bandwidths, and least-square error lines fitted to the points in each group. The two best fitting lines produced the least total squared error. This process indicated that the division between the data should occur between the 630- and 1130-Hz masker bandwidths for the short- and long-delay conditions, but between the 1130- and 2530-Hz masker bandwidths for the burst condition. Note that, because only five masker bandwidths were tested, for each of the conditions, one of the two best fitting lines encompassed only two points

and thus influenced the total squared error (and the subsequent intersection point of the two lines) by its perfect fit.

Bacon, S. P. (**1990**). ''Effect of masker level on overshoot,'' J. Acoust. Soc. Am. **88,** 698–702.

Bacon, S. P., and Moore, B. C. J. (**1986**). ''Temporal effects in simultaneous pure-tone masking: Effects of signal frequency, masker/signal frequency ratio, and masker level,'' Hear. Res. **23**, 257–266.

Bacon, S. P., and Smith, M. A. (**1991**). ''Spectral, intensive, and temporal factors influencing overshoot,'' Q. J. Exp. Psych. **43A**, 373–399.

Bacon, S. P., and Viemeister, N. F. (**1985a**). ''The temporal course of simultaneous tone-on-tone masking,'' J. Acoust. Soc. Am. **78,** 1231–1235.

Bacon, S. P., and Viemeister, N. F. (**1985b**). ''Simultaneous masking by gated and continuous sinusoidal maskers,'' J. Acoust. Soc. Am. **78,** 1220–1230.

Bogartz, R. S. (**1968**). ''A least squares method for fitting intercepting line segments to a set of data points,'' Psychol. Bull. **70,** 749–755.

Carlyon, R. P. (**1989**). ''Changes in the masked thresholds of brief tones produced by prior bursts of noise,'' Hear. Res. **41**, 223–236.

Carlyon, R. P., and Moore, B. C. J. (**1986**). ''Continuous versus gated pedestals and the 'severe departure' from Weber's law,'' J. Acoust. Soc. Am. **79,** 453–460.

Elliott, L. L. (**1965**). ''Changes in the simultaneous masked threshold of brief tones,'' J. Acoust. Soc. Am. **38,** 738–746.

Elliott, L. L. (**1967**). ''Development of auditory narrow-band frequency contours,'' J. Acoust. Soc. Am. **42,** 143–153.

Elliott, L. L. (**1969**). ''Masking of tones before, during, and after brief silent periods in noise,'' J. Acoust. Soc. Am. **45,** 1277–1279.

Fletcher, H. (**1940**). ''Auditory patterns,'' Rev. Mod. Phys. **12,** 47–65.

Green, D. M. (**1969**). ''Masking with continuous and pulsed sinusoids,'' J. Acoust. Soc. Am. **46,** 939–946.

Houtgast, T. (**1974**). ''Lateral suppression in hearing: a psychophysical study on the ear's capability to preserve and enhance spectral contrasts,'' Doctoral dissertation, Academische Pers B. V. Amsterdam, The Netherlands.

Hicks, M. L., and Bacon, S. P. (**1992**). ''Factors influencing temporal effects with notched-noise maskers,'' Hear. Res. **64**, 123–132.

Keiser, J. H. (**1980**). ''Time course of overshoot phenomenon in psychophysical masking,'' Doctoral dissertation, The University of Iowa, Iowa City, Iowa.

Kidd, G. Jr., and Wright, B. A. (**1994**). ''Improving the detectability of a brief tone in noise using forward and backward masker fringes: Monotic and dichotic presentations,'' J. Acoust. Soc. Am. **95,** 962–967.

Kimberley, B. P., Nelson, D. A., and Bacon, S. P. (**1989**). ''Temporal overshoot in simultaneous-masked psychophysical tuning curves from normal and hearing-impaired listeners,'' J. Acoust. Soc. Am. **85,** 1660–1665.

Leshowitz, B., and Cudahy, E. (**1975**). ''Masking patterns for continuous and gated sinusoids,'' J. Acoust. Soc. Am. **58,** 235–242.

Levitt, H. (**1971**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49,** 467–477.

McFadden, D., and Wright, B. A. (**1990**). ''Temporal decline of masking and comodulation detection differences,'' J. Acoust. Soc. Am. **88,** 711–724.

McFadden, D., and Wright, B. A. (**1992**). ''Temporal decline of masking and comodulation masking release,'' J. Acoust. Soc. Am. **91,** 3436–3442.

Patterson, R. D., and Moore, B. C. J. (**1986**). ''Auditory filters and excitation patterns as representations of frequency resolution,'' in *Frequency Selectivity in Hearing*, edited by Brian C. J. Moore (Academic, New York), pp. 123–177.

Scharf, B. (**1970**). ''Critical bands,'' in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), pp. 157–202.

Schmidt, S., and Zwicker, E. (**1991**). ''The effect of masker spectral asymmetry on overshoot in simultaneous masking,'' J. Acoust. Soc. Am. **89,** 1324–1330.

Scholl, H. (**1962**). ''Das dynamische Verhalten des Gehörs bei der Unterteilung des Schallspektrums in Frequenzgruppen,'' Acustica **12,** 101–107.

Smith, R. L., and Zwislocki, J. J. (**1975**). ''Short-term adaptation and incremental responses in single auditory-nerve fibers,'' Biol. Cybern. **17,** 169–182.

Spiegel, M. F. (**1979**). ''The range of spectral integration,'' J. Acoust. Soc. Am. **66,** 1356–1363.

Viemeister, N. F. (**1980**). ''Adaptation of masking,'' in *Psychophysical, Physiological and Behavioral Studies in Hearing*, edited by G. van den Brink and F. A. Bilsen (Delft U.P., Delft), pp. 190–198.

Viemeister, N. F., and Bacon, S. P. (**1982**). ''Forward masking by enhanced components in harmonic complexes,'' J. Acoust. Soc. Am. **71,** 1502–1507.

Wright, B. A. (**1991**). ''Contributions of stimulus bandwidth, the reference condition, and individual differences to the auditory temporal effect: Evidence for within- and across-channel mechanisms,'' Unpublished doctoral dissertation, University of Texas, pp. 277.

Wright, B. A. (**1992**). ''Masker duration affects frequency resolution in simultaneous and nonsimultaneous masking,'' J. Acoust. Soc. Am. **82,** 2364(A).

Wright, B. A. (**1995a**). ''Detectability of simultaneously masked signals as a function of signal bandwidth for different signal delays,'' J. Acoust. Soc. Am. **98,** 2493–2503.

Wright, B. A. (**1995b**). ''Across-channel contributions to auditory enhancement,'' *Advances in Hearing Research: Proceedings of the 10th International Symposium on Hearing*, edited by G. A. Manley, G. M. Klump, C. Koppl, H. Fastl, and H. Oeckinghaus (World Scientific, Singapore), pp. 576–588.

Wright, B. A. (**1996**). ''Correlated individual differences in conditions used to measure psychophysical suppression and signal enhancement,'' J. Acoust. Soc. Am. **100**, 3295–3303.

Wright, B. A., McFadden, D. M., and Champlin, C. A. (**1993**). ''Adaptation of suppression as an explanation of enhancement effects,'' J. Acoust. Soc. Am. **94,** 72–82.

Zwicker, E. (**1965a**). ''Temporal effects in simultaneous masking by white-noise bursts,'' J. Acoust. Soc. Am. **37,** 653–663.

Zwicker, E. (**1965b**). ''Temporal effects in simultaneous masking and loudness,'' J. Acoust. Soc. Am. **38,** 132–141.

Zwicker, E., and Jaroszewski, A. (**1982**). ''Inverse frequency dependence of simultaneous tone-on-tone masking patterns at low levels,'' J. Acoust. Soc. Am. **71,** 1508–1512.

A comparison of detection and discrimination of temporal asymmetry in amplitude modulation

Michael A. Akeroyd[a] and Roy D. Patterson
MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, United Kingdom

(Received 6 March 1996; revised 24 June 1996; accepted 11 September 1996)

Two compound experiments were performed to compare the detection of amplitude modulation with the discrimination of modulator shape when the modulators have strong temporal asymmetry. In experiment 1, an adaptive procedure was used to measure detection and discrimination as a function of modulation frequency from 4 to 400 Hz. In experiment 2, the method of constant stimuli was used to measure psychometric functions for detection and discrimination at one modulation frequency, 8 Hz. The asymmetric modulators were time-reversed pairs. Thus their envelope spectra are identical and models based on the envelope spectrum would predict no effect of asymmetry on detection or discrimination at any modulation depth. The detection results show, as predicted, that the direction of asymmetry does not affect the detectability of modulation in either experiment. In contrast, the discrimination results show that direction of asymmetry is readily discriminable for modulation frequencies less than about 50 Hz, indicating that envelope-spectrum models will require modification if they are to be extended to include discrimination of temporal asymmetry. [S0001-4966(97)04201-X]

PACS numbers: 43.66.Mk, 43.66.Ba [WJ]

INTRODUCTION

The ability to detect sinusoidal amplitude modulation applied to a noise carrier depends upon the frequency of the modulator. The sensitivity to the modulation decreases as the modulation frequency is raised above about 20 Hz (e.g., Viemeister, 1979). One popular model assumes that the modulator is smoothed in the auditory system by a leaky integrator and then detected using a specified decision statistic, such as the standard deviation of the envelope (Viemeister, 1979) or the ratio of the maximum of the envelope to its minimum (Forrest and Green, 1987).

An alternative decision statistic is based on the concept of a modulation filterbank. To detect modulation, it is assumed that a listener attends to a modulation filter centered on the modulation frequency. This model has been used to account for the phenomenon of modulation masking: when sinusoids are used to modulate noise, the presence of modulation at one modulation frequency can interfere with detection of modulation at a neighboring modulation frequency (e.g., Bacon and Grantham, 1989; Houtgast, 1989). The listener attends to the modulation filter centered on the signal frequency, but that filter passes some of the interfering modulator. The amount of interference is determined by the level of the interfering modulator passed by the modulation filter. One method for quantifying the model is to assume that listeners perform a Fourier transform on the envelope of the signal and base their decisions on its magnitude spectrum. For convenience, the magnitude spectrum of the envelope is referred to as the ''envelope spectrum.''

An important restriction of the envelope-spectrum model is that the Fourier transform is insensitive to the temporal direction of the modulators; the envelope spectrum does not change if the stimuli are played backward instead of forward. Three sets of experiments with temporally asymmetric modulators have been reported recently (Patterson, 1994a, b; Akeroyd and Patterson, 1995; Irino and Patterson, 1996), and they all demonstrate that listeners can discriminate the direction of temporal asymmetry when the modulation depth is 100%. Listeners were asked to discriminate between signals having temporally asymmetric, exponential modulators; exponentially decaying (*damped*) stimuli and exponentially rising (*ramped*) stimuli. The only difference between a damped stimulus and a ramped stimulus is that one is the other played backward, and so the envelope spectra of a pair of damped and ramped stimuli are identical. Nevertheless, pairs with the same half-life can be easily discriminated over a wide range of conditions.

None of the damped and ramped experiments, however, included data on the *detection* of temporally asymmetric modulation. In this paper, we compare detection of modulation for three modulator shapes and discrimination of the same three modulator shapes from each other. Two of the modulators are temporally asymmetric, whereas the third is temporally symmetric. The discrimination results are used to test the hypothesis that listeners use the envelope spectrum for modulation discrimination. The detection results are interpreted in terms of the leaky-integrator model and possible decision statistics.

In the discrimination experiments mentioned above, the asymmetric modulators had an exponential shape and the comparison was between rising and falling exponentials. Their envelope minima are essentially zero whenever the period of the modulator is more than about four times the half-life of the modulator. One of the decision statistics proposed for the leaky-integrator model, the maximum/minimum ratio, is unstable with exponential modulators. Accordingly, to

[a]Present address: MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, United Kingdom. Electronic mail: michael.akeroyd@ihr.mrc.ac.uk

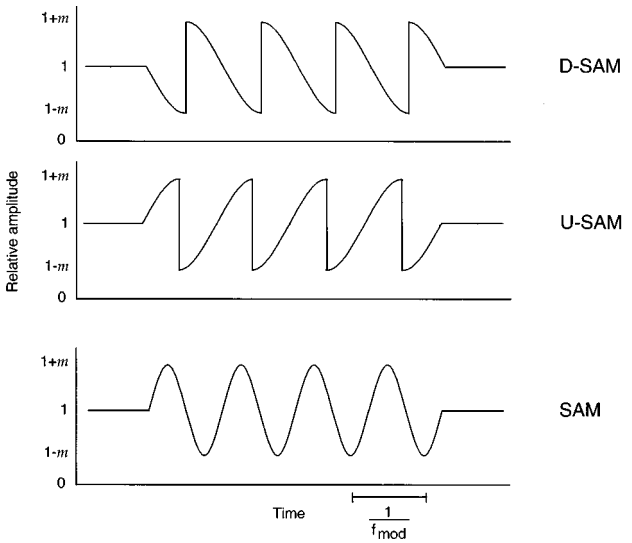


FIG. 1. Schematic illustration of the D-SAM, U-SAM, and SAM modulators.

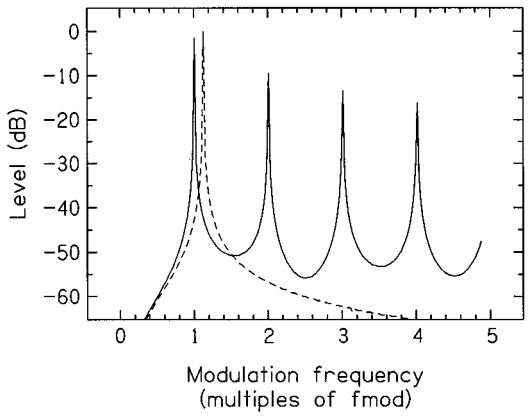


FIG. 2. Envelope spectra for D-SAM (solid line) and SAM (dashed line) modulators. The envelope spectrum for U-SAM is identical to that of D-SAM. Note that, for clarity, the SAM envelope spectrum has been displaced rightward by a small amount.

avoid near-zero minima, we developed a new group of temporally asymmetric sounds, based on sinusoidal amplitude modulation (SAM) as shown in Fig. 1. The damped exponential modulator was replaced by D-SAM (*downward*-sloping sinusoidal amplitude modulation), corresponding to the downward-sloping part of a full sinusoidal function, and the ramped exponential modulator was replaced by U-SAM (*upward*-sloping sinusoidal amplitude modulation), corresponding to the upward-sloping part of a full sinusoidal function. The only difference between D-SAM and U-SAM is that one is the other played backward. The D-SAM and U-SAM modulators are cut from a sinusoid of frequency $f_{\sin}$. The control modulator is a SAM function of frequency $2f_{\sin}$. The rate of the peaks in the modulators, f_{mod}, is equal to $2f_{\sin}$ and is the same in D-SAM, U-SAM, and SAM. The modulation depth is m, where m is the ratio of the amplitude of the modulating sinusoid to the continuous dc level ($0 \leqslant m \leqslant 1$).

The envelope-spectrum model makes specific predictions for both detection and discrimination. The magnitude of the fundamental component in the Fourier transform of a D-SAM or U-SAM modulator is 1.4 dB lower than the magnitude of the fundamental component in the Fourier transform of a SAM modulator (Stephenson, 1973). If it is assumed that listeners base their decisions on the envelope spectrum, then they will detect SAM at a modulation depth 1.4 dB below that at which they can detect D-SAM or U-SAM. It is thus predicted that detection of SAM will be easier than detection of either D-SAM or U-SAM and also that detection of D-SAM and U-SAM will be identical.

The envelope spectra of D-SAM and U-SAM contain higher harmonics which are not present in the envelope spectrum of SAM (Fig. 2). The second harmonic is the most intense, at a level 8.0 dB lower than the level of the D-SAM or U-SAM fundamental (Stephenson, 1973). If it is assumed that listeners base their decisions on the envelope spectrum, then they will be able to discriminate D-SAM from SAM when they can detect the higher harmonics of the D-SAM modulator, and they will be able to discriminate U-SAM from SAM when they can detect the higher harmonics of the U-SAM modulator. It is thus predicted that discrimination of D-SAM from SAM will occur at *exactly* the same modulation depth as discrimination of U-SAM from SAM. It is also predicted that listeners will be unable to discriminate D-SAM from U-SAM, because their envelope spectra are identical.

Experiment 1 reports measurements of threshold modulation depth for detection and discrimination, as a function of modulation frequency. Experiment 2 reports measurements of psychometric functions for detection and discrimination, as a function of modulation depth, for a constant modulation frequency.

I. EXPERIMENT 1: DETECTION AND DISCRIMINATION OF SAM, D-SAM AND U-SAM AS A FUNCTION OF MODULATION FREQUENCY

In the first part of the experiment, the minimum modulation depth at which D-SAM, U-SAM, and SAM could be detected was measured as a function of the modulation frequency. In the second part of the experiment, the minimum modulation depth at which the modulators could be discriminated was measured for the three comparisons, D-SAM versus U-SAM, D-SAM versus SAM, and U-SAM versus SAM.

A. Method

1. Stimuli

The stimuli were amplitude-modulated wideband noises. The modulating function was either the repeated full cycle of a sinusoid, or the repeated down-half of a sinusoid, or the repeated up-half of a sinusoid, referred to as SAM, D-SAM, and U-SAM, respectively. The starting phases were always 0°, 180°, and 0°, respectively, and the finishing phases were always 360°, 180°, and 360°, respectively. The duration of the modulators was 500 ms. The modulators were generated digitally at 20 000 samples per second, played through a 12-bit, low-distortion, digital-to-analog converter (Cambridge Electronic Design 1401) and an extremely sharp low-pass filter set to 8 kHz (Cambridge Electronic Design 1701). They

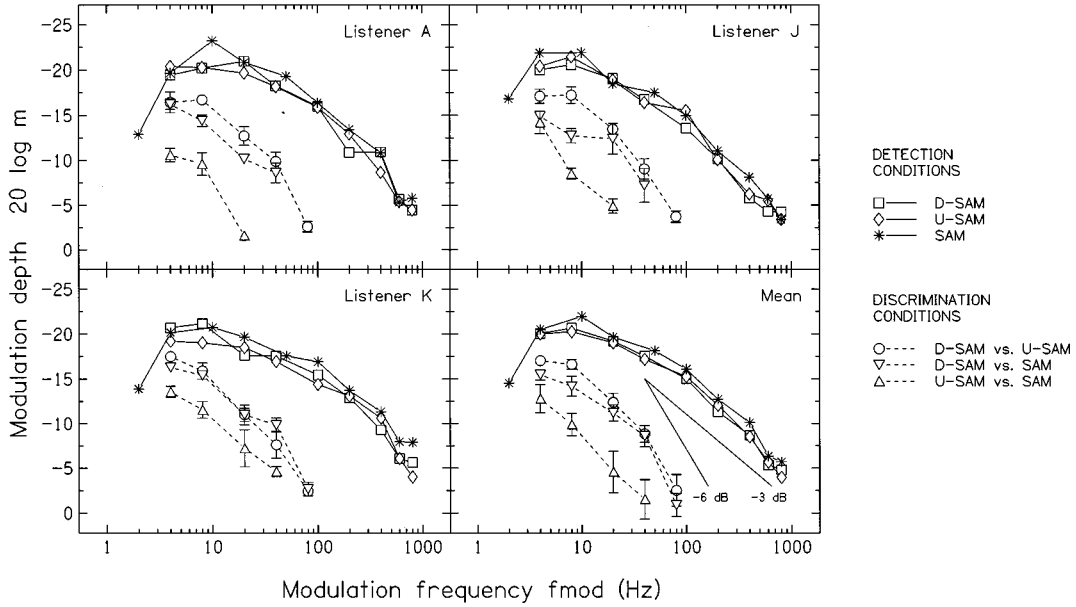


FIG. 3. Threshold functions for detection and discrimination of the three types of modulation. Each panel shows the results for an individual listener, with the mean results shown in the bottom-right panel. Decay rates of slope −3 dB per octave and −6 dB per octave are also shown in the bottom-right panel. Standard deviations (error bars) are indicated for the discrimination conditions only.

were added to a continuous dc level by an analog adder, and then multiplied with a wideband noise using an analog multiplier. The resulting signals were further attenuated and filtered (CED 1701), then high-pass filtered at 100 Hz (−3-dB cutoff frequency, 48-dB per octave slope; Kemo 21CF30/03), and low-pass filtered at 5 kHz (−3-dB cutoff frequency, 48-dB per octave slope; Kemo 21CF30/03). The signals were amplified by a Quad 303 power amplifier and played over one channel of a Sennheiser HD-414 headset. Listeners sat in double-walled IAC booths.

The spectrum level of the continuous noise carrier was 40 dB SPL, measured at 1 kHz using a Hewlett–Packard 3561A spectrum analyzer. In order to minimize the use of changes in overall level among the stimuli in any of the conditions, the spectrum level of the stimuli was roved randomly by up to −3 or +3 dB, in 1-dB steps using a CED 1701. The changes in modulation depth were made by attenuating the level of the modulator (using a CED 1701) before it was multiplied by the carrier noise.

2. Procedure

Thresholds were measured using a 3I-2AFC task and a three-down, one-up adaptive method, estimating the 79.4% point on the psychometric function (Levitt, 1971). This point corresponds to a d' of 1.16. The step size was initially 4 dB, changing to 2 dB after the first reversal. The run continued for 15 more reversals, and threshold was defined as the mean of the last six pairs of reversals. Each trial consisted of a 300-ms warning light, a 500-ms stimulus (interval one; the ''reminder stimulus''), a 500-ms gap, a 500-ms stimulus (interval two), a 500-ms gap, a third 500-ms stimulus (interval three), and then a ''respond now'' light. The intertrial interval varied slightly with load on the experimental computer. The stimulus intervals were marked by lights, and correct/incorrect feedback was provided. The stimulus timing and presentation, together with response collection and feedback, were controlled by a Dell Pentium microcomputer.

Pilot work showed that in the discrimination tests it was easy to forget the timbre of the ''correct'' modulator. The correct modulator was thus played at the beginning of each trial in interval one, in both the detection and discrimination conditions, with a modulation depth equal to that of the signal modulator. Listeners were asked to regard intervals two and three as a two-interval, two-alternative forced-choice (2I-2AFC) task, with the stimulus in interval one as a ''reminder'' stimulus, to be used in case they forgot what they were listening for. The reminder stimuli were D-SAM for the D-SAM versus U-SAM and D-SAM versus SAM conditions and U-SAM for the U-SAM versus SAM condition.

Three adaptive runs per cell per listener were measured. Any adaptive run with a within-run standard deviation greater than 3 dB was rejected as unreliable and an extra adaptive run was measured. This occurred in about 4% of the adaptive runs. Extra adaptive runs were also measured when the between-run standard deviation was greater than 3 dB, and in this case all the runs were included in the analyses. This occurred in about 3% of the conditions.

3. Listeners

Three listeners participated in the experiment, referred to as A, J, and K. All had normal audiometric thresholds. Listener A was the first author; listeners J and K were paid for their participation.

B. Results and discussion

1. Detection

The detection thresholds are shown in Fig. 3 as open symbols connected by solid lines. Each panel shows the results from an individual listener; the mean results are shown

in the bottom-right panel. The results are plotted as 20 log m with the lowest thresholds (best performance) toward the top of the ordinate. The parameter is modulator shape: D-SAM (squares), U-SAM (diamonds), and SAM (asterisks). The standard deviations are omitted for clarity, but had a mean value of 1.2 dB. Slopes of -3 dB per octave and -6 dB per octave are illustrated in the bottom-right panel. The remaining symbols show the discrimination thresholds, to be discussed later in this section.

The threshold functions for all three modulators show the expected low-pass shape, in that performance decreases as the modulation frequency is increased from about 10 Hz. The functions are similar to the functions observed by other authors, although the comparisons cannot be exact because of differences in stimulus level, bandwidth, and experimental method. Nevertheless, the decrease in performance at a modulation frequency of 2 Hz (SAM only; asterisks) has only occasionally been observed before with continuous carriers (e.g., by Bacon and Viemeister, 1985; Forrest and Green, 1987; Scott and Humes, 1990, but not by Viemeister, 1979; or by Grantham and Bacon, 1991), although it has been observed with gated carriers (e.g., Viemeister, 1979; Forrest and Green, 1987). The lowest thresholds were between -20 and -25 dB. These values are similar to those measured by some authors (e.g., Formby and Muir, 1988; Scott and Humes, 1990; Eddins, 1993) but are slightly higher than those measured by other authors (e.g., Viemeister, 1979; Bacon and Viemeister, 1985; Forrest and Green, 1987; Strickland and Viemeister, 1996). The thresholds increased by 3 dB from their minima at a modulation frequency of 30–50 Hz. These values compare well with those reported by other authors. Visually, the slopes of the threshold functions at high frequencies are closer to -3 dB per octave than to -6 dB per octave, in accordance with previous measurements.

The three modulators were not equally detectable. SAM was the easiest to detect; the asterisks are above both the squares and the diamonds in Fig. 3. D-SAM and U-SAM were approximately equally detectable. An ANOVA was used to estimate the statistical significance of the differences in thresholds, with factors of listener (three levels), modulator type (three levels), modulation frequency (4–800 Hz; seven levels), together with adaptive runs (3, 4, or 5 levels) forming the error estimates.[1] The analysis showed a significant main effect of modulator type [$F(2,130)=3.53$, $p<0.05$] and a significant main effect of modulation frequency [$F(6,130)=289$, $p<0.0001$]. A Tukey test showed that SAM performance was significantly better than both D-SAM and U-SAM performance, at the 0.05 level. The interaction of modulation type with modulation frequency was not significant [$F(12,130)=0.381$, $p>0.1$].

Figure 4 shows the threshold differences, together with the standard deviations across listeners (error bars). The differences are calculated as the D-SAM threshold minus the U-SAM threshold (asterisks, joined by dashed lines), the D-SAM threshold minus the SAM threshold (circles, joined by solid lines) and the U-SAM threshold minus the SAM threshold (double triangles, also joined by solid lines); if SAM is easier to detect than D-SAM and U-SAM then the second and third differences should be *positive*. It may be seen that the D-SAM minus U-SAM differences are clustered about 0 dB, as predicted by the envelope-spectrum model (the average difference was -0.14 dB). The D-SAM minus SAM differences and the U-SAM minus SAM differences are generally larger than 0 dB; the average differences were both $+0.96$ dB, respectively, close to the value $+1.4$ dB predicted by the envelope-spectrum model. Thus the envelope-spectrum model can account for the detection thresholds.

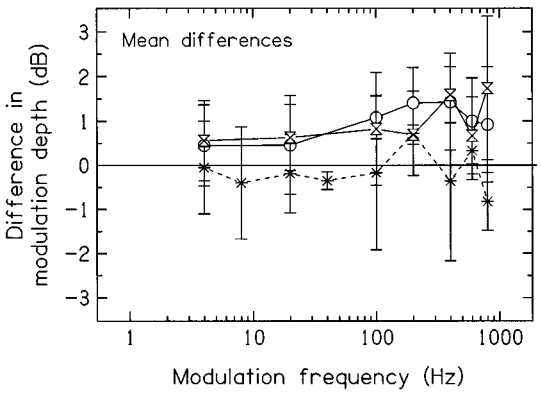


FIG. 4. The detection thresholds from Fig. 3, plotted as a function of the differences between D-SAM and U-SAM thresholds (asterisks), between D-SAM and SAM thresholds (circles), and between U-SAM and SAM thresholds (double triangles). The standard deviation across listeners are indicated by error bars.

2. Discrimination

The discrimination thresholds are shown in Fig. 3 as the symbols connected by dashed lines. The parameter is discrimination condition: D-SAM versus U-SAM (circles), D-SAM versus SAM (downward triangles) and U-SAM versus SAM (upward triangles). The standard deviations are indicated by the error bars; their mean value was 0.9 dB.

Discrimination of modulator shape was consistently harder than detection of modulation. The easiest discrimination was D-SAM versus U-SAM. Discrimination of D-SAM from SAM was easier than discrimination of U-SAM from SAM. The functions still have a low-pass shape. The modulation frequency at which D-SAM would become indistinguishable from U-SAM even at a modulation depth of 0 dB (i.e., 100% modulation) was, by extrapolation, about 100 Hz.

An ANOVA was used to estimate the statistical significance of the threshold differences, with factors of listener (three levels), discrimination type (three levels) and modulation frequency (4, 8, and 20 Hz), together with adaptive runs providing the error terms (3, 4, or 5 levels). The ANOVA was restricted to the frequencies at which all three discriminations could be made. The results showed a significant main effect of discrimination type [$F(2,54)=230$, $p<0.0001$] and a significant main effect of modulation frequency [$F(2,54)=193$, $p<0.0001$]. A Tukey test showed that performance on each of the three discriminations was significantly different from the other two at the 0.05 level. The interaction of modulation shape and modulation frequency was also significant [$F(4,54)=9.51$, $p<0.0001$].

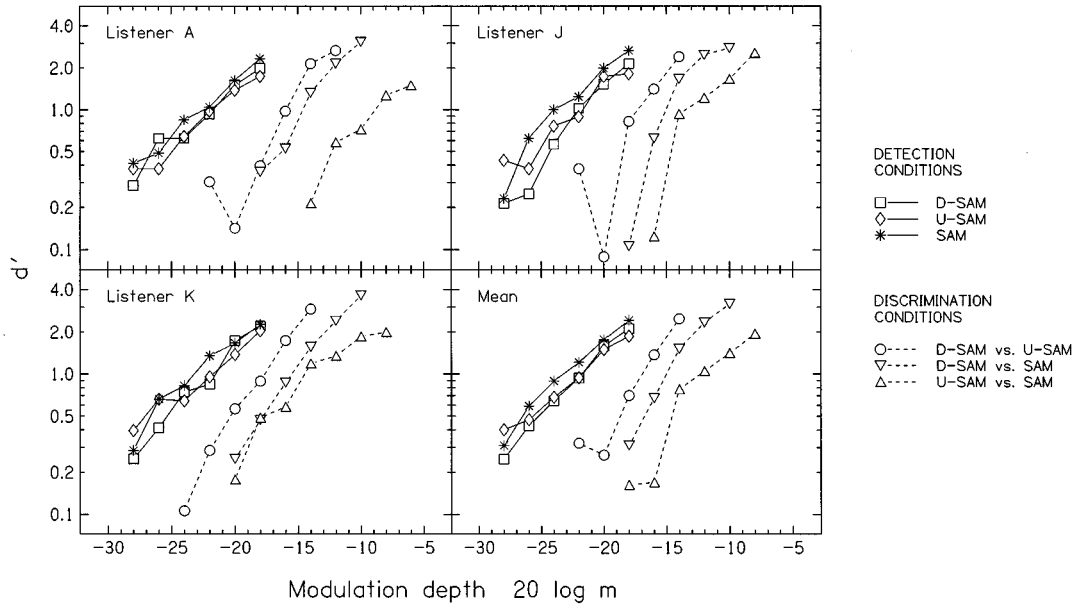


FIG. 5. Psychometric functions for detection and discrimination of the three types of modulation. Each panel shows the results for an individual listener, with the mean results shown in the bottom-right panel.

3. The envelope-spectrum model

The predictions of the envelope-spectrum model are supported by the detection data but not by the discrimination data. Listeners should have been as good at discriminating D-SAM from SAM as they were at discriminating U-SAM from SAM, and they should also have been unable to discriminate between D-SAM and U-SAM. Neither prediction was observed in the data, indicating that listeners did not use the envelope spectrum when making their discrimination decisions. Thus the envelope-spectrum model is only applicable near *detection* threshold.

An alternative hypothesis is that listeners use temporal processes for both detection of and discrimination of amplitude modulation. Bacon and Grantham (1989) noted in their original paper on modulation masking that the phenomenon may be due to temporal processes, and Strickland and Viemeister (1996) have used a temporal model to successfully predict modulation masking. We show below in Sec. III that an explicit calculation of the envelope spectrum is not required to predict a difference in detectability between SAM, D-SAM, and U-SAM; a temporal model can predict this difference.

II. EXPERIMENT 2: DETECTION AND DISCRIMINATION OF SAM, D-SAM, AND U-SAM NOISE AS A FUNCTION OF MODULATION DEPTH

Psychometric functions were measured for detection and discrimination of all three sinusoidal modulators. The purpose was to determine the relationship between detectability or discriminability, as measured by d', and modulation depth, m. The modulation frequency was fixed at 8 Hz.

A power relationship

$$d' = km^p \tag{1}$$

has been reported for detection of SAM sinusoids. The value of p was originally reported to be 2 (Moore and Sek, 1992, 1994a, b, 1995; Edwards and Viemeister, 1994), but more recently Moore and Sek (1996) reported a value of 1.6. For detection of SAM noises, Irwin (1989) suggested a value of 4, although a value of 2 gave fits almost as good, and Sheft and Yost (1990) reported a value of approximately 2. Eddins (1993) used a cumulative Gaussian function to relate percent correct and 20 log m, implying a linear relationship between d' and 20 log m.

A. Method

The stimuli were the same as in experiment 1, but the modulation frequency was fixed at 8 Hz. The apparatus was identical to that used previously. The method of constant stimuli was used, with a counterbalanced design. The structure of each trial was the same as earlier. Each block of trials consisted of 60 or 70 trials, made up of 6 or 7 modulation depths played ten times each per block. The conditions were fixed within a block. The ''signal'' modulators were the same as previously. The results are based on a total of 200 trials per cell per listener. The same three listeners participated; the experiment was conducted after they had completed experiment 1.

B. Results and discussion

1. Detection

The detection functions are shown in Fig. 5 by solid lines. Each panel shows the results from a different listener, with the mean results shown in the bottom-right panel. The parameter is the modulator. D-SAM (squares), U-SAM (diamonds) and SAM (asterisks). The dashed lines show the discrimination functions, to be discussed later.

The three detection functions were similar to each other. Listeners found SAM to be easier to detect than either

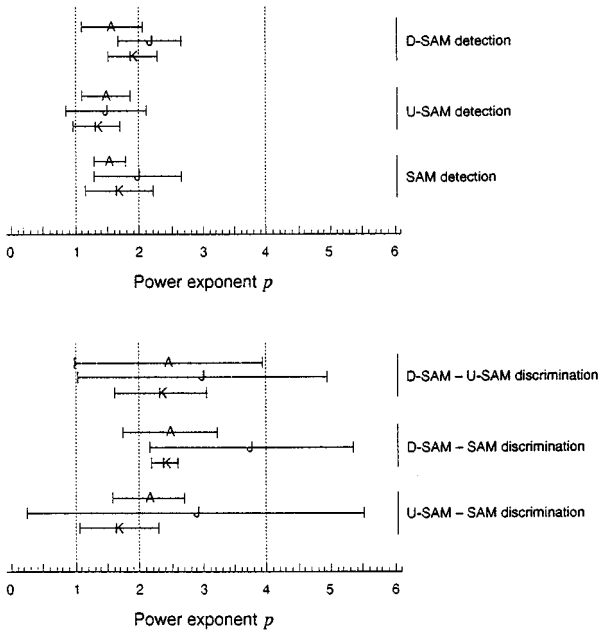


FIG. 6. (Top panel): The values of the power exponent p linking d' and modulation depth m for the detection psychometric functions. The error bars indicate ±95% confidence intervals. (Bottom panel): As the top panel, but for the discrimination psychometric functions.

D-SAM and U-SAM. Visually, the psychometric functions for D-SAM and U-SAM were almost indistinguishable. The logarithmic form of Eq. (6),

$$\log d' = p \log m + c, \tag{2}$$

was fitted to the data for each listener and condition, varying the parameters p and c to derive the best fit. The power exponents p were between 1.3 and 2.2, with a mean value of 1.7. The values of p are illustrated in Fig. 6 (top panel), together with 95% confidence intervals for p.[2] It may be seen that in only one condition (listener J, detection of D-SAM) was the value of p larger than 2. The values of c were between 1.5 and 2.3, with a mean value of 1.9. The values of c are, however, more informative when expressed as the ratio $20\ c/p$, giving the function

$$\log d' = \frac{1}{20}\, p\left(20 \log m + \frac{20c}{p}\right). \tag{3}$$

The ratio $20\ c/p$ gives the horizontal displacement, in dB, of the psychometric functions; the higher its value, the better performance is. The average ratios for the SAM functions were 1.0 dB larger than the average ratios for the U-SAM functions, and 0.9 dB larger than the average ratios for the D-SAM functions. Detection of SAM is thus easier than detection of either D-SAM or U-SAM by an amount equivalent to a 1-dB change in modulation depth. Detection of U-SAM is easier than detection of D-SAM by an amount equivalent to a 0.07-dB change in modulation depth. The squares of the correlation coefficients were between 0.91 and 0.99, with a mean of 0.96, indicating that a power function provides a good summary of the psychometric function in these experiments.

The results of experiment 2 support the results of experiment 1; the detectability of D-SAM and of U-SAM is almost identical. The mean value for the power exponent, p, across conditions was 1.7, which is close to Moore and Sek's (1996) value.

2. Discrimination

The discrimination functions are shown in Fig. 5 by dashed lines. The parameter is the condition: D-SAM versus U-SAM (circles), D-SAM versus SAM (downward triangles) and U-SAM versus SAM (upward triangles). In a small number of cases, performance was less than chance, giving a negative d'; these cases were excluded from Fig. 5 and from the data set used to estimate p and c.

Discrimination of modulator shape was consistently harder than detection of modulation; the discrimination psychometric functions are displaced rightward from the detection psychometric functions. The ordering of the discriminations is the same as in experiment 1; the easiest discrimination was D-SAM versus U-SAM (circles), and discrimination of D-SAM from SAM (downward triangles) was easier than discrimination of U-SAM from SAM (upward triangles).

The logarithmic form of Eq. (1) was fitted to the data for each listener and condition, varying the parameters p and c. The values of the power exponent p were between 1.7 and 3.8, with a mean of 2.6. The values and 95% confidence intervals are illustrated in Fig. 6 (bottom panel); they were consistently larger than those measured for detection, indicating, respectively, that the discrimination slopes were larger than the detection slopes but that the fits were less good (the squares of the correlation coefficients were 0.82–0.99, with a mean of 0.91). This reduction occurs because a subset of the psychometric functions have a downward curvature (e.g., the U-SAM versus SAM function). The values of the constant c were between 0.92 and 2.6, with a mean of 1.8. The values of the ratio $20\ c/p$ showed that the D-SAM versus U-SAM psychometric function was displaced rightward by 5.5 dB from the average position of the three detection psychometric functions. The D-SAM versus SAM psychometric function and the U-SAM versus SAM psychometric function were displaced rightward by 7.7 and 11 dB, respectively, from the average psychometric function for detection.

The results are similar to those of experiment 1; it is harder to discriminate U-SAM from SAM than it is to discriminate D-SAM from SAM. The mean value of p across conditions was 2.6. This is greater than the mean value for the detection conditions, indicating that the discrimination psychometric functions are steeper than the detection psychometric functions. It is possible that psychometric functions for discrimination are, in general, steeper than psychometric functions for detection; Macmillan (1971, 1973) observed steeper slopes for discrimination of amplitude increments and decrements in tones and noises, when compared to detection of the same increments and decrements.

III. GENERAL DISCUSSION

The data indicate that listeners did not use the envelope spectrum when making discrimination decisions. In this section, we consider the hypothesis that they used temporal pro-

cesses for *both* detection of and discrimination of amplitude modulation. By ''temporal processes'' we mean the traditional leaky integrator model. The attraction of the model is the suggestion that the modulation threshold data can be used to estimate the temporal resolution of the auditory system. The roll-off of the threshold function at high-modulation frequencies is assumed to provide a direct measure of the transfer function of a simple low-pass filtering process in the auditory system. In this case, the cutoff frequency of the modulation transfer function provides a single, numerical, value of temporal resolution. The cutoff frequency can be measured in two ways. The first method, following Jeffress (1967), is to construct an electrical or computational model of the auditory system and the experimental procedure (e.g., Viemeister, 1979; Forrest and Green, 1987; Strickland and Viemeister, 1996). The time constant (or cutoff frequency) of the low-pass filter is varied to fit the modulation transfer function to the experimental threshold data. The second method is to fit low-pass filter functions to the modulation threshold data directly. We report calculations using both methods.

A. A leaky-integrator model

Two variations on a single-channel temporal model of SAM detection are commonly discussed in the literature (Viemeister, 1979; Forrest and Green, 1987; Strickland and Viemeister, 1996). Both models consist of a predetection filter, a half-wave rectifier, and a low-pass filter, connected in series, followed by a decision device. The low-pass filter acts as a leaky integrator, and its output is an approximation to the envelope of the stimulus. The two models differ in their choice of decision statistic; Viemeister (1979) used the standard deviation of the envelope, whereas Forrest and Green (1987) used the maximum-to-minimum ratio of the envelope. Both models predict a threshold function which is similar to that observed, although Forrest and Green (1987) showed that the standard-deviation model predicts a high-frequency slope of −6 dB per octave, whereas the maximum-minimum model predicts a high-frequency slope of −3–4 dB per octave.

We implemented computational versions of both models, using the published parameters of Viemeister and of Forrest and Green. Detection thresholds for the models were measured using an average of five up–down adaptive runs, and are shown in Fig. 7 as symbols joined by solid lines (the key is the same as in Figs. 3 and 5). We found a small advantage for the detection of SAM when compared to the detection of D-SAM and U-SAM; the average SAM threshold was better than the average of the D-SAM and U-SAM thresholds, by 0.80 dB (standard-deviation model; Fig. 7, top panel) and by 0.79 dB (maximum/minimum-ratio model; Fig. 7, bottom panel). With the standard-deviation model, we also found a small, but statistically significant, advantage for the detection of U-SAM when compared to the detection of D-SAM: a difference in average threshold of 1.0 dB [$F(1,56)=11.1$, $p<0.01$]. With the maximum/minimum-ratio model, a smaller difference in the same direction was observed (0.30 dB), but it was not statistically significant [$F(1,56)=0.708$, $p>0.05$]. Thus both temporal models reproduced the detection advantage of SAM, without recourse to the envelope spectrum of the modulator. The two temporal models differ slightly in their predictions of the relative ease of detecting U-SAM. The maximum/minimum-ratio model's prediction is closer to the experimental data.

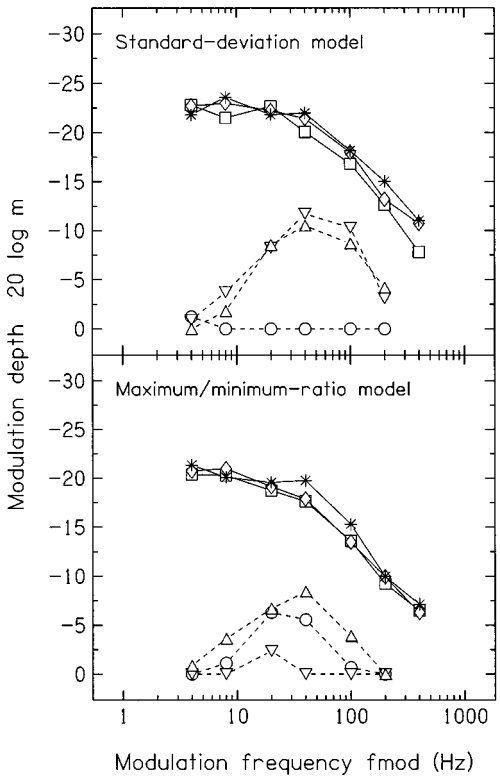


FIG. 7. Predicted threshold functions for detection and discrimination of the three types of modulation, for two variations on a single-channel temporal model, using the standard-deviation (top panel) or the maximum/minimum-ratio (bottom panel) as the decision statistic. The key is the same as in Figs. 3 and 5. Note that standard deviations are *not* shown, in contrast to Fig. 3.

Discrimination thresholds for the two models are shown in Fig. 7 as symbols joined by dashed lines. The functions show a band-pass rather than a low-pass shape. Furthermore, neither of the models predicted that the D-SAM versus U-SAM discrimination (circles) would be the easiest of the three discriminations. Since Akeroyd and Patterson (1995) showed that the two decision statistics also fail to predict discrimination of damped and ramped noises, we conclude that a new decision statistic or a new model is required in order to account for discrimination data.

Figure 8 shows envelopes from the standard-deviation model for a set of modulated noises, illustrating the discrimination problem. The modulation frequency is 50 Hz, and the modulation depth is −6 dB (top row) and −12 dB (bottom row). Each envelope is based on an average of 50 independent noises. One possible decision model is to assume that the auditory system uses an envelope-shape recognizer to decide which envelope is which. The experimental results show that discrimination of U-SAM from SAM (middle column versus right column) is, in general, more difficult than either discrimination of D-SAM from SAM (left column versus right column) or discrimination of D-SAM from U-SAM (left column versus middle column). Thus the envelope-

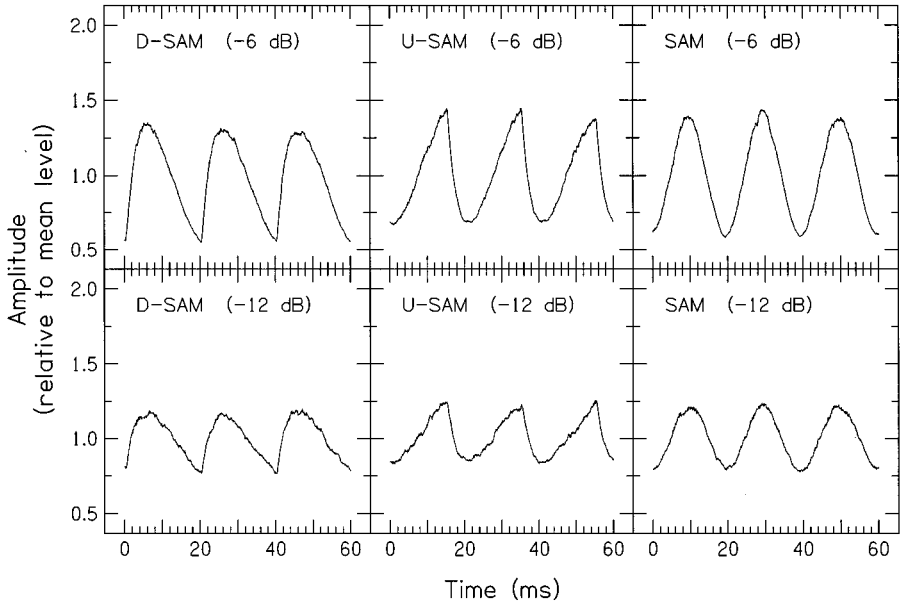


FIG. 8. The output of a single-channel, leaky-integrator model for D-SAM noise, U-SAM noise and SAM noise. The modulation frequency was 50 Hz, and the modulation was either −6 or −12 dB. Each panel shows the averaged output of 50 independent noises.

shape recognizer would be required to decide that a U-SAM shape is more similar to a SAM shape than a D-SAM shape is. The model does not seem promising, however, because the envelopes shown in Fig. 8 suggest that the D-SAM shape is more similar to the SAM shape.

Psychometric functions were also calculated for the detection conditions, based on 200 trials per value of 20 log m. The predicted psychometric functions were similar to the experimental psychometric functions, and thus the logarithmic form of Eq. (1) was fitted to the predicted functions. Using the standard-deviation model, the values of p were 1.8, 1.7, and 1.7, for D-SAM, U-SAM, and SAM, respectively. Using the maximum-minimum model, the values of p were 1.9, 1.2, and 1.4, respectively. The squares of the correlation coefficient were between 0.91 and 0.98. Thus both models give estimates of the power exponent p consistent with the experimental data.

Both the standard-deviation model and the maximum/minimum model are based on a single-channel model of temporal processing. In order to account for the overall efficiency of the listeners, the bandwidth of the single channel has to be set to a value far larger than the width of any individual auditory filter. This is because the amount of modulation passed by a filter is dependent upon the bandwidth of the filter; the wider the bandwidth is, the greater the modulation depth that is passed. We assume here, following previous authors (e.g., Viemeister, 1979), that the auditory system can integrate across the outputs of a set of auditory filters, preserving the fine time relationships as it does so, in order to approximate a single filter of 2- to 4-kHz bandwidth.[3]

B. The shape of the threshold function

Two numerical functions were fitted to the detection data (Fig. 9). One function (dashed line) assumes that the threshold function is determined by the transfer function of a first-order low-pass filter (Schroeder, 1981; Scott and Humes, 1990; Eddins, 1993); that is,

$$20 \log m = -1 \times 10 \log\left(\frac{1}{1+(2\pi\tau f_{\mathrm{mod}})^2}\right) + K \tag{4}$$

$$= 10 \log\left(1 + \frac{f_{\mathrm{mod}}^2}{f_c^2}\right) + K, \tag{5}$$

where f_{mod} is the modulation frequency, τ is the time constant of the simple low-pass filter, f_c is the −3-dB cutoff frequency of the first-order low-pass filter, and K is an additive constant. The −1 multiplier in Eq. (4) is required to convert the attenuation function of the filter to a threshold function. The high-frequency slope is −6 dB per octave. The other function (solid line) assumes that the threshold func-

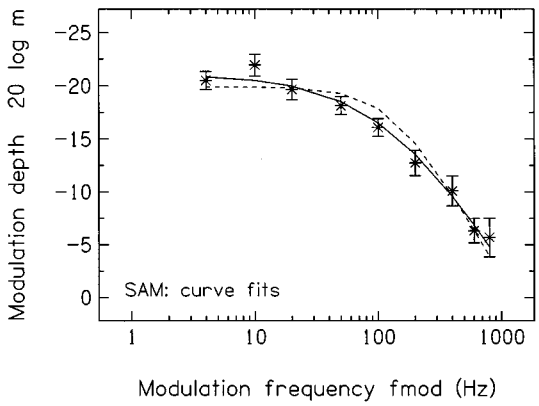


FIG. 9. The SAM detection thresholds from Fig. 3 (asterisks), together with numerical functions fitted to the thresholds. The numerical functions are a first-order low-pass equation (dashed line) and Formby and Muir's (1988) equation (solid line). The standard deviations across listeners are indicated by error bars.

tion is determined by an equation introduced by Formby and Muir (1988)

$$20 \log m = -1 \times 20 \log\left(\frac{1}{c + f_{\mathrm{mod}}}\right) + K, \tag{6}$$

where c is proportional to f_c

$$f_c = c(\sqrt{2} - 1) = \frac{1}{2\pi\tau}. \tag{7}$$

It can be expressed in a form similar to Eq. (5) as follows:

$$20 \log m = 10 \log\left(\frac{1}{(\sqrt{2}-1)^2} + \frac{2}{(\sqrt{2}-1)}\frac{f_{\mathrm{mod}}}{f_c} + \frac{f_{\mathrm{mod}}^2}{f_c^2}\right) + K'. \tag{8}$$

Both functions [Eqs. (5) and (8)] include a constant term and a f_{mod}^2/f_c^2 term, but the Formby–Muir function also includes a f_{mod}/f_c term. The high-frequency slope is again -6 dB per octave when f_{mod}^2/f_c^2 is much larger than f_{mod}/f_c, but for frequencies near f_c the slope is closer to -3 dB per octave. The free parameters in both functions are the -3-dB frequency f_c and the constant K, corresponding to the vertical displacement of the transfer functions.

The best-fitting values of the parameters were calculated for each combination of condition, listener and function separately (the thresholds at a modulation of 2 Hz were excluded from the analysis). The values of f_c for the first-order low-pass function were, 120, 130, and 130, for D-SAM, U-SAM, and SAM, respectively (rms of 1.6, 1.3, and 1.4 dB, respectively). The values correspond to time constants of 1.3, 1.2, and 1.2 ms, respectively. Using the Formby–Muir function, the values of f_c were 58, 59, and 63 Hz, respectively (rms errors of 1.0, 0.70, and 0.92 dB, respectively). The values correspond to time constants of 2.7, 2.8, and 2.5 ms, respectively.[4] The constants K were always smaller for the SAM conditions than for the D-SAM and U-SAM conditions, by, on average, 0.85 dB (low-pass) and 1.3 dB (Formby–Muir), in line with the experimental results. In each combination of listener and condition, the Formby–Muir function gave better fits than the first-order low-pass function, as the rms errors were lower for the Formby–Muir function than for the first-order low-pass function. As may be seen from Fig. 9, the Formby–Muir function (solid line) provides a better fit to the threshold data near f_c; it has a softer ''knee'' than the first-order low-pass function (dashed line).[5]

Hall and Grose (1994) compared the Formby–Muir function with the first-order low-pass function, and found acceptable fits to their threshold functions with both low-pass functions, although the least-square error from the Formby–Muir function was slightly greater than that for the first-order low-pass function. Scott and Humes (1990) used low-pass functions with asymptotic slopes of -6 dB per octave (first order) and of -12 dB per octave (second order), and found that the first-order function fitted better than the second-order function.

IV. SUMMARY AND CONCLUSIONS

The detection of modulation was measured for symmetric and asymmetric amplitude modulators, and compared to the discrimination of modulation shape for the same stimuli. The asymmetric modulators were time-reversed pairs made of the downward-sloping part of a sinusoid (D-SAM) and the upward-sloping part of a sinusoid (U-SAM). The symmetric modulator was the full cycle of a sinusoid (SAM).

(1) Detection of amplitude modulation does not depend upon the temporal direction of the modulation. Discrimination of amplitude modulation does depend upon the temporal direction of the modulation, and is generally more difficult than detection of amplitude modulation.

(2) Discrimination of D-SAM noise from U-SAM noise is easier than discrimination of either noise from SAM noise. Discrimination of D-SAM noise from SAM noise is easier than discrimination of U-SAM noise from SAM noise. This indicates that listeners do not use the envelope spectrum in making their decisions.

(3) Detection of SAM is slightly easier than detection of either D-SAM or U-SAM. A single-channel temporal model, using either the standard deviation of the envelope, or its maximum/minimum ratio, as the decision statistic can explain this difference.

(4) A low-pass filter function proposed by Formby and Muir [1988, Eq. (1)] provides a better fit to our detection thresholds than a first-order low-pass filter.

(5) The psychometric functions for detection of amplitude modulation can be summarized by the relationship $d' \propto m^{1.7}$. The psychometric functions for discrimination of amplitude modulation can be summarized by the relationship $d' \propto m^{2.6}$. The psychometric functions for discrimination are steeper than those for detection.

ACKNOWLEDGMENTS

We wish to thank Quentin Summerfield and two anonymous reviewers for their comments on a previous draft of this manuscript and the Institute of Hearing Research for use of their computational resources. This research was carried out while the first author was supported by a MRC Research Studentship and by a DRA grant (FRNIC/U/759).

[1]Due to a programming error, the 8- and 40-Hz SAM stimuli were actually 10 and 50 Hz, respectively. These two stimuli were not included in either the ANOVA or the measurement of average threshold differences.

[2]The analyses were conducted with the STATVIEW software package, and the 95% confidence intervals are equal to ± 2.76 times the standard error.

[3]The maximum/minimum-ratio model (Forrest and Green, 1987) used a bandwidth of 4 kHz whereas the standard-deviation model (Viemeister, 1979) used a bandwidth of 2 kHz. This difference is due to two effects. First, the maximum/minimum-ratio statistic is less efficient at detecting modulation in the presence of random noise than the standard-deviation statistic. Second, the experimental thresholds upon which the models are based differ; the lowest thresholds measured by Forrest and Green (1987, Fig. 11) were slightly lower than those measured by Viemeister (1979, Fig. 7). We did carry out some informal tests of a multichannel filterbank. We used three auditory filters centred at frequencies of 1.77, 2.00, and 2.25 kHz (the center frequencies are 1 ERB apart), but instead of using the remaining stages of the leaky-integrator model we used the auditory-image model and integrated its output over time interval (Akeroyd and Patterson, 1995). Unfortunately, the model was unable to predict the performance levels observed experimentally. At this juncture, we do not know if this failure

was due to the temporal integration processes in the auditory-image model or to the narrow bandwidths of the auditory filters.

[4]Strictly, a ''time constant'' can only be measured for a low-pass filter with an exponential impulse response, as it is defined as the time at which the response has decayed to $1/e$ ($=0.3679$) of its original value. A first-order low-pass filter has an exponential impulse response, and for this filter the time constant is equal to the reciprocal of the product of 2π and the -3-dB cutoff frequency [cf. Eq. (7)]. The impulse response of the Formby–Muir low-pass filter is not exponential and so the definition of time constant does not apply. It is, however, *approximately* exponential because of the f_{mod}^2/f_c^2 term [Eq. (8)] and as a result the reciprocal of the product of 2π and the -3 dB cutoff frequency is a useful summary measure.

[5]It will be noted that the values for the cutoff frequency of the first-order low-pass filter—an average of 130 Hz—are greater than the -3-dB points reported for the threshold functions—30–50 Hz (Sec. II B 1). The difference is due to the different methods used to measure the cutoff frequencies. The value of 130 Hz is based on fitting a low-pass function to *all* the data points (excepting those at a modulation frequency of 2 Hz, which give a bandpass function), even though as shown in Fig. 9, the low-pass function (dashed line) is a poor fit to the data points. The value of 50 Hz is based on finding the minimum threshold on the threshold function (Fig. 3), and then deriving graphically the frequency that gives a threshold higher by 3 dB. The cutoff frequencies of the leaky integrators used in Viemeister's (1979) and Forrest and Green's (1987) models were based on electrical or computational simulations. The values are different from values of the -3-dB points derived using a graphical method; Viemeister (1979) derived a -3-dB point of 50 Hz but used a cutoff frequency of 65 Hz in his model, and Forrest and Green (1987) measured a -3-dB point of about 20 Hz but used a cutoff frequency of 53 Hz in their model. It is of interest that the cutoff frequencies of the Formby–Muir function, approximately 60 Hz, are near the cutoff frequencies of the leaky integrator used in those models.

Akeroyd, M. A., and Patterson, R. D. (**1995**). ''Discrimination of wideband noises modulated by a temporally asymmetric function,'' J. Acoust. Soc. Am. **98,** 2466–2474.

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 Viemeister, N. F. (**1985**). ''Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners,'' Audiology **24**, 117–134.

Eddins, D. A. (**1993**). ''Amplitude modulation detection of narrow-band noise: Effects of absolute bandwidth and frequency region,'' J. Acoust. Soc. Am. **93,** 470–479.

Edwards, B. W., and Viemeister, N. F. (**1994**). ''Modulation detection and discrimination with three-component signals,'' J. Acoust. Soc. Am. **95,** 2202–2212.

Formby, C., and Muir, K. (**1988**). ''Modulation and gap detection for broadband and filtered noise signals,'' J. Acoust. Soc. Am. **84,** 545–550.

Forrest, T. G., and Green, D. M. (**1987**). ''Detection of partially filled gaps in noise and the temporal modulation transfer function,'' J. Acoust. Soc. Am. **82,** 1933–1943.

Grantham, D. W., and Bacon, S. P. (**1991**). ''Binaural modulation masking,'' J. Acoust. Soc. Am. **89,** 1340–1349.

Hall, J. W., and Grose, J. H. (**1994**). ''Development of temporal resolution in children as measured by the temporal modulation transfer function,'' J. Acoust. Soc. Am. **96,** 150–154.

Houtgast, T. (**1989**). ''Frequency selectivity in amplitude-modulation detection,'' J. Acoust. Soc. Am. **85,** 1676–1680.

Irino, T., and Patterson, R. D. (**1996**). ''Temporal asymmetry in the auditory system,'' J. Acoust. Soc. Am. **99,** 2316–2331.

Irwin, R. J. (**1989**). ''Psychometric functions for the discrimination of differences in intensity of Gaussian noise,'' Q. J. Exp. Psychol. **41A** 655–674.

Jeffress, L. A. (**1967**). ''Stimulus-orientated approach to detection reexamined,'' J. Acoust. Soc. Am. **41,** 480–488.

Levitt, H. L. (**1971**). ''Transformed up–down methods in psychophysics,'' J. Acoust. Soc. Am. **49,** 467–477.

Macmillan, N. A. (**1971**). ''Detection and recognition of increments and decrements in auditory intensity,'' Percept. Psychophys. **10,** 233–238.

Macmillan, N. A. (**1973**). ''Detection and recognition of intensity changes in tone and noise: The detection-recognition disparity,'' Percept. Psychophys. **13,** 65–75.

Moore, B. C. J., and Sek, A. (**1992**). ''Detection of combined frequency and amplitude modulation,'' J. Acoust. Soc. Am. **92,** 3119–3131.

Moore, B. C. J., and Sek, A. (**1994a**). ''Discrimination of modulation type (amplitude modulation or frequency modulation) with and without background noise,'' J. Acoust. Soc. Am. **96,** 726–732.

Moore, B. C. J., and Sek, A. (**1994b**). ''Effects of carrier frequency and background noise on the detection of mixed modulation,'' J. Acoust. Soc. Am. **96,** 741–751.

Moore, B. C. J., and 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., and Sek, A. (**1996**). ''Detection of frequency modulation at low rates: Evidence for a mechanism based on phase locking,'' J. Acoust. Soc. Am **100**, 2320–2331.

Patterson, R. D. (**1994a**). ''The sound of a sinusoid: Spectral models,'' J. Acoust. Soc. Am. **96,** 1409–1418.

Patterson, R. D. (**1994b**). ''The sound of a sinusoid: Time-interval models,'' J. Acoust. Soc. Am. **96,** 1419–1428.

Schroeder, M. R. (**1981**). ''Modulation transfer functions: Definition and measurement,'' Acustica **49**, 179–182.

Scott, D. M., and Humes, L. E. (**1990**). ''Modulation transfer functions: A comparison of the results of three methods,'' J. Speech Hear. Res. **33**, 390–397.

Sheft, S., and Yost, W. A. (**1990**). ''Temporal integration in amplitude modulation detection,'' J. Acoust. Soc. Am. **88,** 796–805.

Stephenson, G. (**1973**). *Mathematical Methods for Science Students* (Longman Scientific and Technical, Harlow, Essex, UK), 2nd ed.

Strickland, E. A., and Viemeister, N. F. (**1996**). ''Cues for discrimination of envelopes,'' J. Acoust. Soc. Am. **99,** 3638–3646.

Viemeister, N. F. (**1979**). ''Temporal modulation transfer functions based upon modulation thresholds,'' J. Acoust. Soc. Am. **66,** 1364–1380.

Comparison of electrode discrimination, pitch ranking, and pitch scaling data in postlingually deafened adult cochlear implant subjects

Leslie M. Collins[a)]
Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina 27708-0291 and Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, Michigan 48109-2122

Teresa A. Zwolan
Department of Otolaryngology, University of Michigan, Ann Arbor, Michigan 48109-0312

Gregory H. Wakefield
Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, Michigan 48109-2122 and Department of Otolaryngology, University of Michigan, Ann Arbor, Michigan 48109-0312

(Received 5 January 1996; revised 24 May 1996; accepted 14 June 1996)

The goal of this study was to investigate the relationship between variation in electrode site of stimulation and the perceptual dimensions along which such stimuli vary. This information may allow more effective use of electrode place when encoding speech information. To achieve this goal, two procedures which measure pitch in subjects implanted with the Nucleus/Cochlear Corporation multichannel device were performed. Estimates of electrode discriminability that can be obtained from these procedures were compared to a more direct measure of electrode discriminability that was obtained in a previous study [Collins *et al.*, Assoc. Res. Otolaryng. Abstracts, No. 642 (1994)]. In the first task, subjects performed a pitch ranking procedure similar to that used in previous studies [Townshend *et al.*, J. Acoust. Soc. Am. **82**, 106–115 (1987); Nelson *et al.*, J. Acoust. Soc. Am. **98**, 1987–1999 (1995)]. Estimates of the pitch percept elicited by stimulation of each electrode as well as the discriminability of the electrodes were generated from the data using two different statistical analyses. In the second task, subjects performed a pitch scaling procedure similar to one used in a previous study [Busby *et al.*, J. Acoust. Soc. Am. **95**, 2658–2669 (1994)]. Again, two different statistical analyses were performed to generate estimates of the pitch percept corresponding to stimulation of each electrode and to generate estimates of electrode discriminability. In general, the estimates of the relationships between the pitch percepts obtained from the two procedures were not identical. In addition, the estimates of electrode discriminability were not equivalent to the electrode discrimination measures obtained from the same subjects during the previous study. Signal detection theory has been used to model the decision processes required by each of the procedures described above [e.g., Jesteadt and Bilger, J. Acoust. Soc. Am. **55**, 1266–1276 (1974)]. However, these models do not predict the differences that were observed between the data sets obtained during this study. An alternate model is proposed which may explain the data obtained from these subjects. This model is based on the assumption that the percept that is elicited by electrical stimulation of an electrode is multidimensional, as opposed to unidimensional in nature. Therefore, the perceived signal is more appropriately modeled using a multidimensional random vector, where each element of the vector represents the perceived value of one of the dimensions of the signal. [S0001-4966(97)05912-2]

PACS numbers: 43.66.Ts, 43.66.Fe, 43.66.Ba [JWH]

INTRODUCTION

The ability to discriminate between different speech tokens varies substantially across subjects implanted with a cochlear prosthesis, and the sources of this variability are not well understood. Two of the most common assumptions for speech coding under electrical stimulation are that each stimulation site provides an independent channel through which information can be transmitted, and that the pitch associated with the percept elicited by stimulation at each site follows the tonotopic organization of a normal cochlea. It is possible that differences in the validity of these two assumptions across subjects is one of the sources of the variability in speech recognition.

In a previous study (Collins *et al.*, 1994), we hypothesized that electrode sites were not equally discriminable for subjects implanted with the Nucleus/Cochlear Corporation device, and that poor electrode discriminability results in a degradation in the information that can be transmitted to the subjects by the speech processing system of the prosthetic

[a)]Electronic mail: lcollins@ee.duke.edu

device. The results of the study indicated that there was substantial variability in the discriminability of electrodes across subjects, and that data from the electrode discrimination tasks could be used to improve speech recognition in some subjects. The goal of the current study was to determine whether data from the more traditional procedures that are based on pitch discrimination could be used to obtain the same estimates of electrode discriminability, and whether speech recognition could also be enhanced if subjects' speech processors were reconfigured based on these pitch measures.

Previous studies that have investigated the relationship between variation in site of stimulation and the percepts elicited by such variation have focused on the perceptual attribute of pitch. Such studies have used pitch scaling, labeling, or ranking techniques to study the relationship between pitch and variation in site of stimulation (Busby *et al.*, 1994; Townshend *et al.*, 1987; Tong *et al.*, 1982, 1980; Clark *et al.*, 1988; Eddington *et al.*, 1978; Tong and Clark, 1985; Shannon, 1983; Crosby *et al.*, 1984; Pauka, 1989; Nelson *et al.*, 1995; Dorman *et al.*, 1990). These studies have reported that the pitch percepts associated with electrical stimulation vary in an orderly fashion along the electrode array for most subjects, and that the perceived order matches the presumed tonotopic organization of an unimpaired cochlea. As opposed to the results obtained in the Collins *et al.* study, measures of discrimination obtained from the pitch-based data have tended to support the assumption that each electrode is discriminable from every other electrode, and therefore that an electrode site is equivalent to an independent channel.

Two studies have examined the relationship between the pitch percept and the site of stimulation in subjects implanted with the Nucleus/Cochlear Corporation device in detail. Nelson *et al.* (1995) studied the distinctiveness of the pitch percept associated with stimulation of different electrodes in 14 subjects using a pitch ranking procedure. Large differences in performance between subjects were observed, but an orderly relationship between pitch and electrode position was reported for the majority of subjects. The authors also reported a correlation between consonant identification and performance on the pitch ranking task. However, the slope of the regression line relating these two measures was close to zero, indicating that one measure cannot be predicted from the other. Busby *et al.* (1994) also evaluated the relationship between the pitch percept and site of stimulation in subjects implanted with the Nucleus/Cochlear Corporation device. In this study, a scaling procedure was performed by the subjects. The authors reported that the pitch percepts followed a general tonotopic structure for most subjects. Instances in which there were local reversals of the pitch order of the electrodes were observed for several subjects. Neither study specifically analyzed the discriminability indices for individual electrodes.

There have been no studies that have rigorously investigated whether reconfiguring subjects' speech processor to correct for deviations from a tonotopic pitch structure can improve speech recognition. Several authors have suggested that electrodes should be ordered based on a corrected pitch structure (Crosby *et al.*, 1984; Pauka, 1989), however, no results have been presented comparing speech recognition data using both the natural order (electrodes ordered sequentially from base to apex) and an electrode ordering based on measured pitch. Although only anecdotal speech results were provided, Pauka (1989) reported that subjects preferred a pitch-ordered MAP[1] to their clinical MAP, which was based on a natural electrode order, and that the preference was noted immediately upon reprogramming.

In the present study, two of the techniques that have previously been used to measure the pitch percept that is associated with site of stimulation were duplicated in the same set of subjects for which electrode discrimination had previously been measured. The first procedure used a paired comparison task to measure pitch rank, while the second procedure used scaling to obtain a numerical estimate of pitch. The discriminability estimates that were obtained from the estimates of electrode pitch were compared to the discriminability measures that were obtained using a more direct technique (Collins *et al.*, 1994). In addition, two alternative methods of analyzing the pitch ranking and scaling data were explored. These methods were derived from the underlying statistical nature of each of the tasks.

The clinical impact of the data obtained in this experiment was also investigated, since it is likely that improvements in speech recognition performance can be achieved when the processor is better matched to the percepts elicited by electrical stimulation for each individual listener (Tyler, 1991; Collins *et al.*, 1994). The relationship between the pitch estimates and position along the electrode array provided an indication of any deviations from the expected tonotopic pitch structure presupposed by the natural ordering of the electrodes. Based on the measured pitch structure, the electrodes were reordered in an experimental MAP. The reordering attempted to create an orderly progression of pitch coincident with frequency band assignments, thus tailoring the speech processor to each individual subject. Subjects' speech recognition performance using this modified MAP was measured and compared to their scores obtained using their standard, or clinical, MAP.

I. METHODS

A. Subjects

Eleven postlingually deafened adults[2] served as subjects for this study. All subjects were implanted with the Nucleus/Cochlear Corporation multichannel device. Demographic information regarding the subjects is listed in Table I. All subjects had used their device a minimum of 6 months.

Only electrodes used in the subjects' clinically programmed MAPs were stimulated in the experimental tasks; the number of active electrodes ranged from 10 to 20.[3] The mode of stimulation used to test each subject was the same as that used in their clinically programmed MAP. Stimuli were presented in a bipolar plus one (BP+1) mode of stimulation for nine of the eleven subjects and in a bipolar plus 2 (BP+2) mode for two subjects. Subjects participated in eight to ten test sessions lasting 2 to 4 h each.

TABLE I. Biographical information for the implanted subjects. Dynamic range was measured in Cochlear device amplitude steps.

Subject	Sex	Age (years)	Age at onset of profound deafness (years)	Age at implantation (years)	Average dynamic range across all active electrodes	Mode of stimulation
S1	M	49	33	45	60.1	BP+1
S2	F	75	40	71	15.6	BP+1
S3	F	70	51	66	29.3	BP+1
S4	M	33	7	31	43.6	BP+1
S5	M	67	16	62	43.7	BP+1
S6	M	66	25	62	40.1	BP+1
S7	F	51	42	47	38.8	BP+1
S8	F	40	8	39	34.7	BP+1
S9	M	60	20	57	105.9	BP+2
S10	M	61	12	57	71.5	BP+2
S11	F	73	70	72	41.7	BP+1

B. Stimuli

1. Pitch ranking and pitch scaling tasks

The stimuli for each of the tasks were 200-ms pulse trains consisting of biphasic rectangular pulses presented at a rate of 250 pps. Pulse duration for all stimuli was held constant at 204 μs per phase. In the ranking task, the interstimulus interval was 500 ms. A loudness balancing procedure (Collins *et al.*, 1994; Jesteadt, 1980) had been performed previously to obtain stimulus amplitudes for the electrode discrimination task. These amplitude values (measured in cochlear device amplitude steps) were also used in this study to maintain consistency between the stimuli presented to the subjects across experiments. Prior to performing the ranking and scaling tasks, each electrode was stimulated at these amplitude values to verify that subjects' hearing had not changed to a point where the stimuli had become either inaudible or too loud. In all cases, subjects reported the stimuli to be comfortably loud, and none of the subjects reported differences in the perceived loudness across the stimuli.

Stimuli were presented to subjects through a Cochlear Corporation Mini Speech Processor connected to a Cochlear Corporation Dual Processor Interface. A Compaq 386 computer controlled the psychophysical procedure and the interface system. A color monitor was used to display stimulus cueing to the subjects.

2. Speech recognition task

Five different speech recognition tests were administered using taped stimuli: a 9-choice closed-set medial vowel recognition task (heed, had, heard, hud, who'd, hawed, hid, head, and hood) in which five tokens were presented per item, a 14-choice closed-set medial consonant recognition task (apa, aba, ama, ana, ata, ada, ava, afa, aga, aka, asa, aza, asha, and aja) in which five tokens were presented per item, the NU6 Monosyllabic Words Test (Tillman and Carhart, 1966) scored for both words and phonemes correct, and the CID Everyday Sentences test (Davis and Silverman, 1978). Male talkers were used in all cases.

C. Experiments

1. Pitch ranking task

This task was performed using a paired comparison procedure, where on each trial subjects judged the higher of the two pitches associated with stimulation of two different electrodes. A block of trials consisted of five presentations of each possible pair of stimuli, and the order of presentation across all possible pairs of electrodes was randomized. The order of presentation within a pair was also randomized, and no explicit balancing of order was imposed across the five presentations of each pair of stimuli in a block of trials. At least three blocks were collected for each subject following a single training block.

2. Pitch scaling task

In the pitch scaling task, subjects assigned a numerical value in the range of 1–100 to a stimulus delivered to one of the possible electrode sites. Presentation order was randomized within a block of trials. Subjects were instructed to score the pitch of the stimulus, with small and large numbers indicating low and high pitches, respectively. A training block where all possible electrodes were stimulated twice was completed prior to data collection. Final scores were determined by calculating the trimmed mean of the final eight of ten values recorded for stimulation of each electrode. (The trimmed mean was calculated by discarding the highest and lowest of the final eight scores assigned to each electrode, and then calculating the mean of the remaining six scores.)

3. Speech recognition task

The speech recognition tests were administered to subjects while they utilized an experimental MAP in which frequency bands were assigned to specific electrodes based on the pitch ordering results obtained from the nonparametric analysis of the pitch ranking data (described in the next section). These scores were compared to the scores measured when subjects utilized their clinical MAP, in which elec-

trodes were assigned to frequency bands in natural order. Subjects had no previous experience using the experimental MAPs prior to speech recognition testing.

The experimental MAP was created by reordering selected electrodes contained in the clinical MAP. The MAP creation software, provided by Nucleus/Cochlear Corporation (Cochlear Corporation, 1993), reassigned the frequency bands used for speech coding to the reordered electrodes based on the standard linear:logarithmic spacing. This type of frequency spacing split the total number of electrodes available for use into a ratio of approximately 1:2 (linear:logarithmic), so that about one third of the electrodes were assigned to the $F1$ frequency range while the remaining two thirds of the array were assigned to the $F2$ frequency range. The electrodes assigned to the $F1$ range were located in the apical section of the array and covered the input frequency range of 280–1000 Hz. The frequency bands assigned to these electrodes were linearly spaced. The remaining electrodes in the mid and basal section of the electrode array were assigned to the $F2$ input frequency range of 1000–4000 Hz and were divided into logarithmically equal frequency bands. Additional MAP parameters were identical for the two MAPs. All MAPs employed the MPEAK encoding strategy.

Subjects were tested in a sound-treated booth using their own speech processors set to a comfortable listening level. Taped stimuli were presented to the subjects at 70 dB SPL, and subjects were instructed to listen to the stimuli and to write down the word, sentence, or token following its presentation. Stimuli were presented a single time only and feedback was not provided. The vowel and consonant recognition tests consisted of five presentations of each stimulus in random order. The NU6 words test contained a randomized list of 50 words, and the CID sentences test consisted of 20 sentences containing a total of 100 key words. All scores are reported as percent correct.

II. RESULTS

The data from the pitch ranking task and the pitch scaling task were analyzed to obtain estimates of both electrode pitch and electrode discriminability. First, measures of the pitch of each electrode or the relative pitches between electrodes were determined for each task. Using these measures of pitch or relative pitch, the electrodes were reordered in an experimental MAP so that an orderly progression of pitch was achieved, and speech recognition was measured. Theoretically, the electrode ordering based on the experimental data should have provided a more accurate tonotopic structure to the speech coding algorithms. Next, measures which provided estimates of electrode discriminability were calculated for each subject, and the subset of discriminable electrodes was determined for each task and analysis method. The methods for obtaining the pitch and relative pitch estimates, electrode order, the various measures of electrode discriminability, and additional analyses are discussed in the following subsections.

TABLE II. Cumulative stimulus response matrix for subject S10. Entries are the number of times the row electrode is judged higher in pitch than the column electrode. The total number of trials for each comparison was 40 for this subject.

Electrode	11	12	13	14	15	16	17	18	19
10	24	37	34	38	38	37	38	39	39
11	…	27	34	33	37	37	39	40	38
12	…	…	28	27	38	38	37	35	38
13	…	…	…	23	29	37	36	39	30
14	…	…	…	…	28	30	32	30	33
15	…	…	…	…	…	24	23	24	28
16	…	…	…	…	…	…	16	22	27
17	…	…	…	…	…	…	…	24	25
18	…	…	…	…	…	…	…	…	20

A. Pitch ranking task

Stimulus response matrices were calculated based on subjects' responses to all of the combinations of stimuli presented during each block of trials. For each matrix, the (i,j) entry corresponds to the number of times electrode i was reported by the subject to be higher in pitch than electrode j in N trials. In a Nucleus/Cochlear Corporation device that is programmed for a bipolar plus one mode of stimulation, the electrodes that are available for stimulation are numbered 1–20 in a basal to apical direction. The tonotopic arrangement of the intact cochlea suggests that electrodes would be ordered in pitch from highest pitch at electrode 1 to lowest pitch at electrode 20. If the pitch percepts associated with the electrode array exactly followed the tonotopic organization of an intact cochlea, and if the subjects behaved as ideal observers, the resulting stimulus response matrix would be upper triangular, with all entries above the main diagonal having value N.

The individual stimulus response matrices from each block of trials were combined for each subject to form a cumulative stimulus response matrix. An example of a cumulative stimulus response matrix for subject S10 is shown in Table II. Only the data above the main diagonal of the matrix is reported, since the (j,i) entry can be determined from the (i,j) entry as $(j,i)=N_{\text{total}}-(i,j)$. None of the subjects in this study appeared to have pitch percepts that followed a tonotopic organization (all entries above the main diagonal of magnitude N_{total}).

Two methods were used to obtain measures of the relative pitches between electrodes and to generate measures of electrode discriminability from the cumulative stimulus response matrices. The first method, referred to as d' analysis, is the same as that used by Townshend *et al.* (1987) and Nelson *et al.* (1995). This method is based on the derivations presented by Green and Swets (1974), Durlach and Braida (1969) and Torgerson (1958). The second method, referred to as row sum analysis, is based on a statistical theory developed specifically for analyzing the results of paired comparison experiments (David, 1988). One benefit of the row sum method is that, unlike the d' method, it does not require the assumptions that the underlying decision process is based on a unidimensional percept, or that the distributions of the perceptual responses evoked by each stimulus have equal variance. An additional benefit of the row sum method is that the

results may be used to estimate whether the evoked percept is unidimensional, since a multidimensional percept may result in a complex or invalid pitch ordering (David, 1988).

1. d' analysis

In this method, the sensitivity index, or d', was calculated from the ''percent correct'' scores obtained from the cumulative stimulus response matrices. As described by Nelson *et al.* (1995), the value of d' was set to 3.3 when the percent correct score was 100% (d' for a perfect score is indeterminate, this value corresponds to 99 to 100 correct responses). In addition, when the percent correct score was 50%, the value of d' was set to 0.0. For an ideal observer whose pitch percepts follow a tonotopic organization, the percent correct score between two electrodes would always be 100%; therefore the d' scores would always be equal to 3.3.

Once each of the entries in the cumulative stimulus response matrices was converted to an entry in a d' matrix, a summary statistic that is a cumulative measure of the discriminability of the pitch percept between adjacent electrodes across the electrode array was calculated (Nelson *et al.*, 1995). The summation was initialized to zero at the apical end of the array (highest numbered electrodes) so that high values of the statistic were indicative of a percept with a higher pitch. The cumulative d' curves for a subject with a pitch percept that is tonotopic would linearly decrease with increasing electrode number at a rate of -3.3 cumulative d'/electrode. The cumulative d' results obtained for each of the subjects using this analysis procedure are shown by the curves marked with open-triangle symbols in Fig. 1. Electrode number is plotted along the abscissa, and the cumulative d' score is plotted along the ordinate. The cumulative d' scale is labeled on the left axis by the second of the two numbers at each tick mark.

The pattern of results illustrated by the cumulative d' scores is similar to the results previously presented by Nelson *et al.* (1995), although more variability in the level of performance was observed in the subjects who participated in this study. None of the subjects' performance was indicative of a tonotopic pitch structure. In fact, not all of the cumulative d' curves were monotonically decreasing. An increase in the value of the cumulative d' curve between an electrode and its adjacent electrode in the apical direction occurred when there was a pitch reversal, which is represented in the data as a negative d' value (''percent correct'' less than 50%). A cumulative d' curve that is not monotonically decreasing indicates that the natural electrode ordering did not provide an orderly progression of pitch as site of stimulation was varied along the electrode array.

Since the d' matrices provide a measure of discriminability between each electrode and the other electrodes in the array, the d' values calculated from the stimulus response matrices can also be used to select a subset of discriminable electrodes from the set of electrodes utilized by each subject. To maintain compatibility with the data obtained in the electrode discrimination experiment, a d' value of 1.0 was used to indicate discriminability. The d' matrix for subject S10 is shown in Table III. Using this data, the subset of discriminable electrodes can be obtained as follows. The first row of the d' matrix indicates that although electrode 11 was indiscriminable from electrode 10, electrode 12 was discriminable from electrode 10. The third row indicates that electrodes 13 and 14 were indiscriminable from electrode 12, but that electrode 15 was discriminable from electrode 12. Continuing in this fashion, the set of discriminable electrodes in this subject's array was determined to be {10, 12, 15}. Alternatively, electrode 10 can be eliminated since it is indiscriminable from electrode 11, and the same analysis can be performed beginning with electrode 11. This analysis results in the set of discriminable electrodes {11, 13, 15}. In a similar fashion, the optimal subset of discriminable electrodes, defined as that set consisting of the largest number of electrodes and having the largest cumulative d', was determined for each subject.

The results of this analysis for each subject's d' matrix are listed in Table IV, along with the results from the electrode discrimination experiment and the other analyses that will be described later in this section. Note that in Table IV, the results are listed as the subset of indiscriminable electrodes, or electrodes that would potentially be removed from a subject's MAP. In all but one case (subject S7), the number of indiscriminable electrodes that was estimated using this method exceeds the number of indiscriminable electrodes measured during the electrode discrimination experiment. However, the subset of indiscriminable electrodes that were established in the electrode discrimination experiment was usually-included in the subset of indiscriminable electrodes established by this analysis procedure.

The cumulative d' analysis provides insight into the structure of the relationships between the pitch percepts associated with stimulation of adjacent electrodes across the electrode array. However, it does not provide a methodology to generate an electrode ordering that results in an orderly progression of the pitch percept. A method was developed that used the d' data to generate an optimal, or tonotopic, electrode ordering. The order was selected to maximize the number of positive entries above the main diagonal of the d' matrix, and to minimize the magnitudes of the negative entries. This method generated cumulative d' curves that were as close to monotonic as possible. This approach is similar to the method proposed by Townshend *et al.* (1987). The electrode orders for each subject that were generated from this analysis are listed in Table V. Only one subject (S3) had a pitch-based ordering of electrodes that matched the natural ordering of electrodes. Two subjects had one set of adjacent electrodes in reverse order in the tonotopic ordering (S2, S10), and the other subjects had increasing numbers of electrodes whose ordering was inconsistent with the natural ordering.

2. Row sum analysis

This method of analyzing data obtained in a paired comparison experiment is based on a model of the probabilities associated with the decision process. This method is described in detail by David (1988). This method also uses the stimulus response matrices. A score a_i was determined for the ith electrode as the sum of the scores in the ith row of the

cumulative stimulus response matrix. This score corresponds to the number of times the ith electrode was perceived to be higher in pitch than other electrodes in the array. The scores that would be achieved by an ideal observer with a tonotopic pitch structure that matches the natural electrode ordering are given by $a_i=(t-i)N$, $i=1,\ldots,t$, where t is the total number of active electrodes for the subject.

The data obtained in the pitch ranking task were analyzed using this procedure for each subject. The cumulative row scores were normalized by converting them to a ''percent wins'' statistic based on the total number of trials for each comparison. (This is equivalent to comparing the row sum scores to the scores expected under a tonotopic distribution of pitch.) The results of this analysis may be used to generate a pitch order directly by ordering the electrodes in decreasing order of percent wins. The results of the row sum formulation of the data are shown by the curves marked with open-circle symbols in Fig. 1. The stimulated electrode is plotted along the ordinate and percent wins is plotted along the abscissa. The percent wins scale is labeled on the left axis by the first of the two numbers at each tick mark. Again, significant variability in the estimated pitch structure was

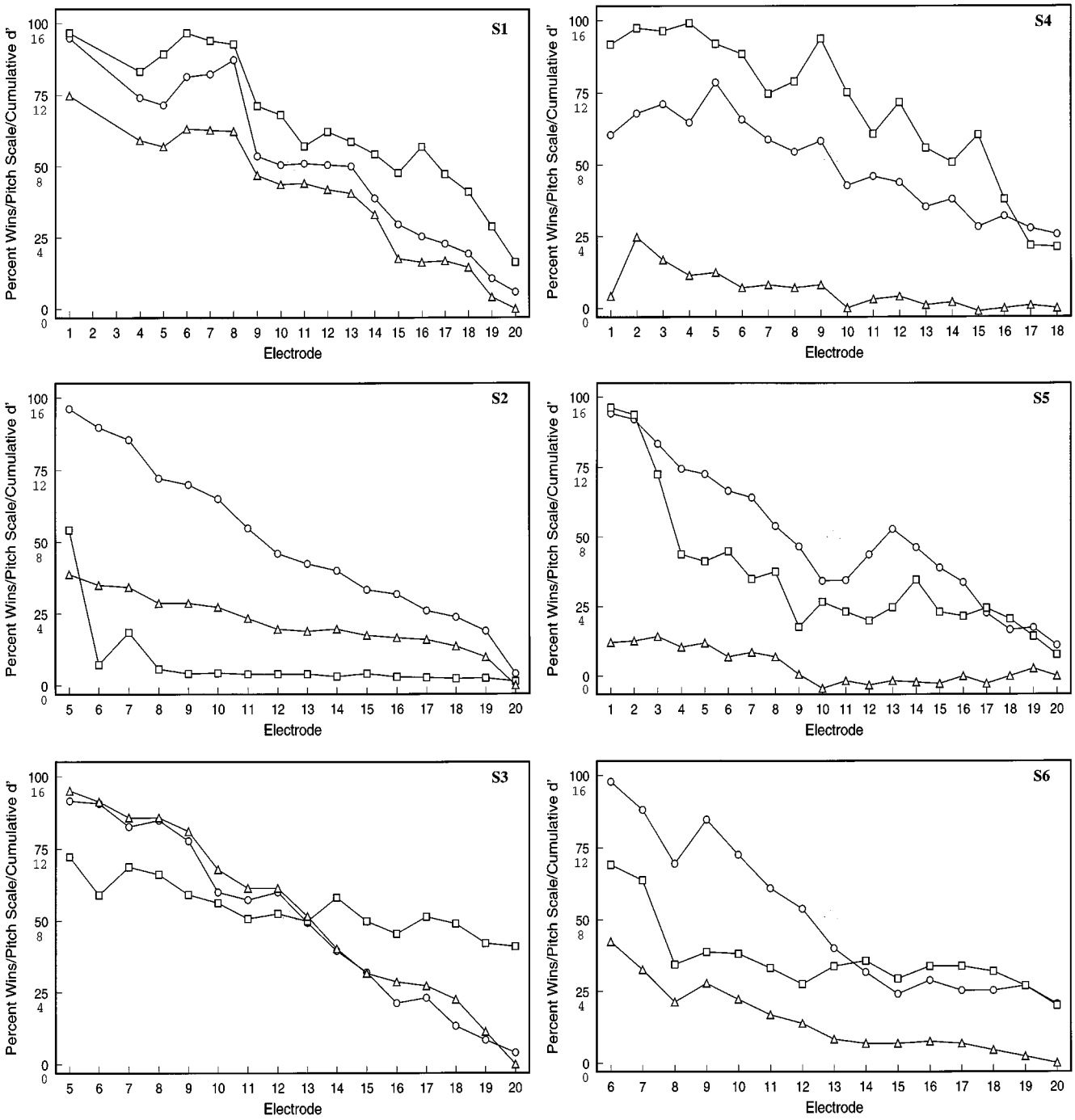


FIG. 1. Cumulative d' data for the pitch ranking task (open triangles), percent wins data for the pitch ranking task (open circles), and trimmed mean data for the pitch scaling task (open squares) for the 11 subjects who participated in this study. Cumulative d' (regular scale), percent wins (bold scale), and pitch scale (bold scale) is plotted on the ordinate, electrode number along the abscissa.

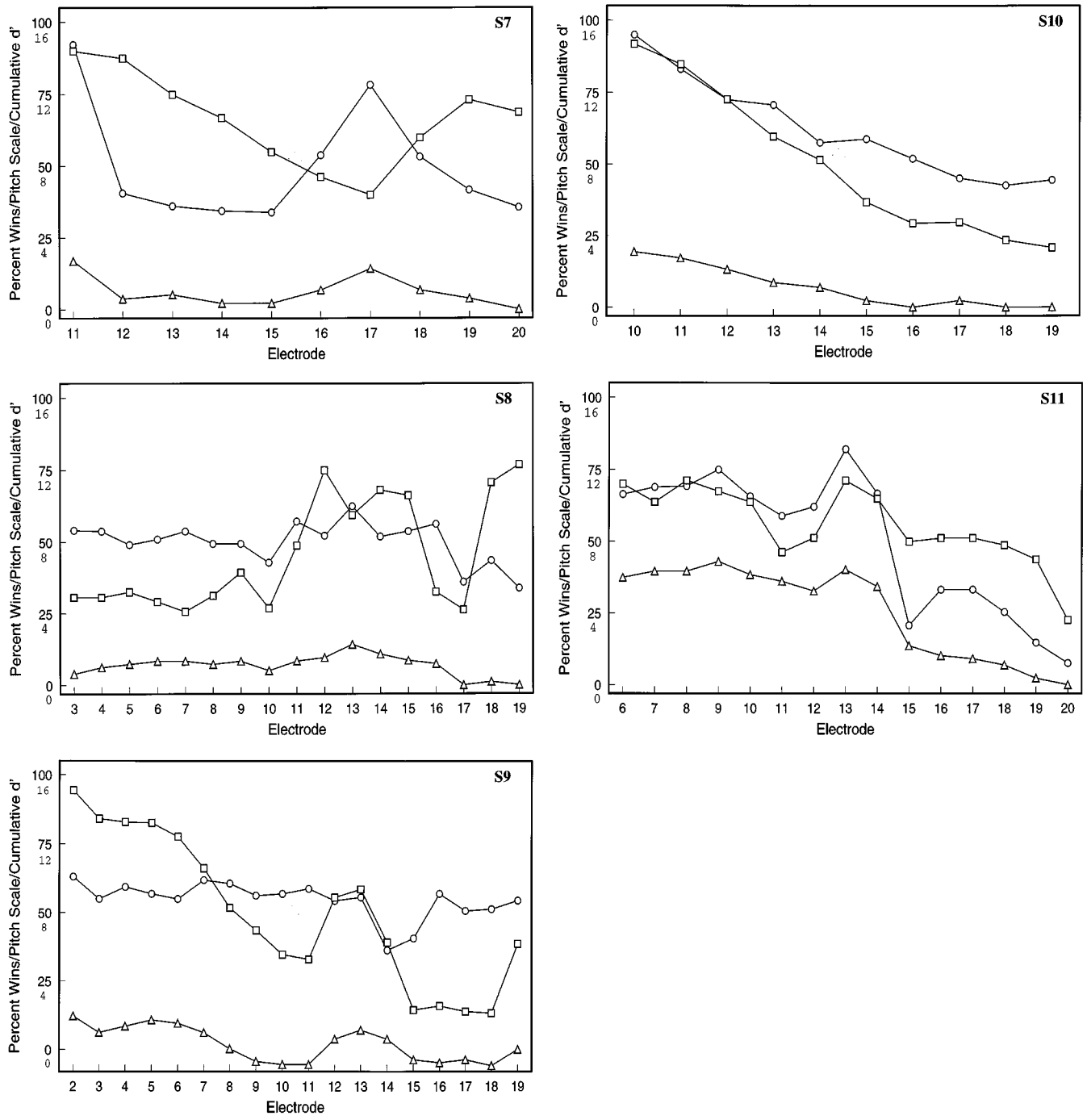


FIG. 1. (*Continued.*)

observed across subjects, and no subjects performed equivalently to an ideal observer with a tonotopic pitch structure.

A measure of the consistency of the ranking data, which is related to the number of transitivity violations within the ranking data for each of the subjects, can be calculated based on the row sum scores. This consistency measure is listed in Table VI. The measure has a value of 1 when the data obtained in the pitch ranking task is completely self-consistent and approaches a value of 0.0 as the number of inconsistencies increases. Although the consistency measure may be interpreted as a measure of the percentage of time a subject selected a stimulus at random during the task, it may also indicate that the underlying structure of the percept is not unidimensional (David, 1988; Torgerson, 1958). The range of consistency measured across the subjects in this study was 0.1 to 0.87. These values may indicate that several of the subjects are at least partially guessing when judging the stimuli based on their pitch (as might occur if the stimuli are indiscriminable), or the values may indicate a multidimensional percept associated with the stimuli.

The results of the row sum formulation scored as percent wins can be compared with the cumulative d' formulation, also shown in Fig. 1, since both the percent wins and the cumulative d' curves are a measure of the relationship be-

TABLE III. The d' matrix for subject S10 calculated from the pitch ranking data. Entries are the d' indices calculated from the data in Table II.

Electrode	11	12	13	14	15	16	17	18	19
10	0.36	2.04	1.47	2.33	2.33	2.04	2.33	2.77	2.77
11	⋯	0.64	1.47	1.32	2.04	2.04	2.77	3.3	2.33
12	⋯	⋯	0.74	0.64	2.33	2.33	2.04	1.63	2.33
13	⋯	⋯	⋯	0.27	0.85	2.04	1.81	2.77	0.95
14	⋯	⋯	⋯	⋯	0.74	0.95	1.19	0.95	1.32
15	⋯	⋯	⋯	⋯	⋯	0.36	0.27	0.36	0.74
16	⋯	⋯	⋯	⋯	⋯	⋯	−0.36	0.18	0.64
17	⋯	⋯	⋯	⋯	⋯	⋯	⋯	0.36	0.45
18	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯	0.0

tween pitch and spatial position. For all of the subjects, the relationship between pitch and spatial position is similar under both formulations, especially for the subjects with a more monotonic pitch structure. The observation that the two methods of analyzing the data result in similar estimates of the pitch structure with respect to electrode position is not surprising, since both utilize the same data to calculate the underlying pitch structure. The row sum formulation uses all of the data in the stimulus response matrices, while the cumulative d' formulation only uses data corresponding to adjacent electrodes. It can be shown that under a hypothesis of an approximately monotonic pitch structure, the two methods do in fact generate similar results. Table VI lists the percentage of the time that the relative pitches were the same between the row sum analysis and the d' analysis. The data from subjects S5 and S8 were the least consistent across the two analysis methods.

To generate a measure of electrode discriminability from the row sum formulation of the data, a two-step procedure was utilized. First, a statistical test was performed to determine whether any significant differences existed between the set of row sum scores for each subject (a significant difference existed for all subjects tested). Then, an approach known as the least significant difference method was employed to determine the minimum difference between two scores that is required for the difference to be statistically significant (David, 1988).

An example of the results of this analysis is shown in Table VII for subject S10. The first row and column list the electrode number along with the cumulative row sums associated with each of the electrodes evaluated in the procedure. The least significant difference method applied to this data set indicated that a minimum difference of 15 between row sum scores was required to achieve a d' of 1.0. Those pairs of electrodes that exceeded this minimum difference are denoted by a ''Y'' in the table, and those that did not exceed the difference are denoted by a ''N.'' In a manner similar to that used to generate a set of discriminable electrodes from the d' matrices (described above), a set of discriminable electrodes can be derived from the row sum data using this analysis. For example, the set of discriminable electrodes for subject S10 determined from the data in Table VII is {10,11,12,13,14,15,16,18}.

The subset of indiscriminable electrodes that was determined for each subject using this analysis is also provided in Table IV. In all but one case (subject S7), the number of indiscriminable electrodes that was estimated using this method exceeds the number of indiscriminable electrodes measured during the electrode discrimination experiment. However, the subset of indiscriminable electrodes that was established in the electrode discrimination experiment was usually included in the subset of indiscriminable electrodes established by this analysis procedure. In all 11 cases, the subset of indiscriminable electrodes estimated by the least significant difference analysis was smaller than the subset generated by the d' analysis method.

It is important to confirm that the least significant difference method of analysis is powerful enough to detect discriminable electrodes for an ideal observer based on the amount of data obtained under the experimental conditions of the ranking experiment. Based on the statistical analysis provided in David (1988), it can be shown that under the conditions of this experiment, the method would estimate a d' of 1.22 between adjacent, discriminable electrodes for an ideal observer with a tonotopic pitch structure. Thus it is possible to determine that adjacent electrodes are discriminable for the ideal observer under the conditions in this experiment.

The row sum method of analyzing the ranking data generated estimates of pitch order more directly than the d' analysis. Although both methods can detect discriminable electrodes for an ideal observer, the row sum method appears to be more accurate than the d' analysis method. However, neither of the methods resulted in estimates of electrode discriminability that were completely consistent with the results from the electrode discrimination experiment, and the consistency measures indicate that the data for many of the subjects in the pitch ranking task was not self-consistent.

B. Pitch scaling task

Results for the pitch scaling task are reported as the trimmed mean of the final eight values assigned by each subject to each electrode. The trimmed mean data are marked with open-square symbols in Fig. 1. The stimulated electrode is plotted along the ordinate and trimmed mean of the pitch estimates is plotted along the abscissa. The pitch scale is labeled on the left axis by the first of the two numbers at each tick mark. As was the case with the ranking data, a monotonic decrease in pitch estimates from the basal to the apical end of the array (low to high numbered electrodes) is indicative of a tonotopic pitch structure. The data from the

TABLE IV. List of indiscriminable electrodes determined for each subject and each analysis method. ED indicates electrode discrimination data presented in Collins *et al.* (1994). PR indicates analysis based on pitch ranking data, PS indicates analysis based on pitch scaling data.

Subject	Test	Indiscriminable electrodes
S1	ED	16
	PR: d'	5,6,8,10,11,12,13,16,18,20
	PR: least significant difference	5,7,10,11,12,13,17
	PS: Student's t test	8,10,14
S2	ED	9
	PR: d'	6,8,10,12,13,14,16,17,18
	PR: least significant difference	9,13,16,18
	PR: Student's t test	7,8,9,10,11,12,13,14,15,16,17,18,19
S3	ED	6
	PR: d'	6,7,8,12,16,18
	PR: least significant difference	6,8,12,15
	PS: Student's t test	8,10,12,13,18,20
S4	ED	2,3,5,8,9,11,14
	PR: d'	4,5,7,9,11,12,13,14,15,16,18
	PR: least significant difference	3,4,8,9,10,12,13,15,17
	PS: Student's t test	3,4,5,8,18
S5	ED	7,12,16
	PR: d'	2,3,5,6,7,10,11,12,13,14,15,16,18,19
	PR: least significant difference	2,5,7,11,19
	PS: Student's t test	2,5,6,8,11,12,13,14,15,17,18,19
S6	ED	10,14,15,16
	PR: d'	10,12,13,15,16,17,18,19
	PR: least significant difference	12,17,18,19
	PS: Student's t test	10,14,16,17,18
S7	ED	12,13,14,15,16,18,19
	PR: d'	12,14,15,19,20
	PR: least significant difference	14,15
	PS: Student's t test	12,20
S8	ED	4,5,6,7,8,9,10,13,14
	PR: d'	3,4,5,6,8,9,10,11,12,13,14,15,16,17,18
	PR: least significant difference	4,6,9,15
	PS: Student's t test	3,5,6,14,15
S9	ED	3,5,9,11
	PR: d'	3,5,6,8,10,11,12,13,14,16,17,18
	PR: least significant difference	4,5,11,13,16,17,18
	PS: Student's t test	3,5,8,10,11,12,13,15,18,19
S10	ED	None
	PR: d'	11,13,14,16,17,18,19
	PR: least significant difference	17,19
	PS: Student's t test	13,15,18,19
S11	ED	None
	PR: d'	6,7,8,9,10,12,14,16,17,18,19
	PR: least significant difference	7,8,12,17
	PS: Student's t test	10,12,14,15,17,18

scaling task can be used to generate a pitch order directly, simply by ranking the electrodes in decreasing order of the trimmed mean. Although some of the gross characteristics of the resulting pitch structures are the same as those generated either by the cumulative d' or row sum analyses of the data from the ranking task described above, the correspondence between methods is not exact, especially for the subjects whose pitch structure appears to be less tonotopic. In some cases (S1, S2, S4, S5, S10, S11) the pitch structure predicted by the two tasks is fairly similar. However, in other cases (S3, S6, S7, S8, S9) the discrepancy is quite striking. Table VI lists the percentage of the time that the relative pitches were the same between the pitch ranking data (row sum analysis) and the pitch scaling data. The data from subjects S7 and S8 were the least consistent across the two tasks.

The standard deviations of the scaling data for the subjects tended to be within approximately 10% of the pitch scale. This is consistent with the data reported by Shannon (1983) and Busby *et al.* (1994). One difference noted between the subjects in the two previous studies and this study

TABLE V. Electrode orders based on measured pitch for each subject and each analysis method. PR indicates analysis based on pitch ranking data; PS indicates analysis based on pitch scaling data.

Subject	Test	Electrode ordering
S1	PR: d'	1,6,7,8,4,5,9,10,11,12,13,14,15,17,16,18,19,20
	PR: least significant difference	1,8,7,6,4,5,9,11,10,12,13,14,15,16,17,18,19,20
	PS	1,6,7,8,5,4,9,10,12,13,11,16,14,15,17,18,19,20
S2	PR: d'	5,6,7,8,9,10,11,12,14,13,15,16,17,18,19,20
	PR: least significant difference	5,6,7,8,9,10,11,12,14,13,15,16,17,18,19,20
	PS	5,6,7,8,15,9,10,12,11,13,14,16,17,18,19,20
S3	PR: d'	5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20
	PR: least significant difference	5,6,8,7,9,10,12,11,13,14,15,16,17,18,19,20
	PS	5,7,8,6,9,14,10,12,17,11,18,15,13,16,19,20
S4	PR: d'	5,2,3,1,4,6,9,7,8,12,11,15,10,14,13,17,18,16
	PR: least significant difference	5,3,2,6,4,1,7,9,8,11,12,10,14,13,16,15,17,18
	PS	4,2,3,5,1,9,6,8,7,10,12,11,15,13,14,16,17,18
S5	PR: d'	1,3,2,5,4,7,6,8,13,9,14,11,12,16,15,10,17,19,18,20
	PR: least significant difference	1,2,3,4,5,6,7,8,13,9,14,12,15,11,10,16,17,19,18,20
	PS	1,2,3,6,4,5,8,7,14,10,11,13,15,12,17,18,9,16,19,20
S6	PR: d'	6,7,9,10,8,11,12,13,16,14,17,18,19,20,15
	PR: least significant difference	6,7,9,10,8,11,12,13,14,16,19,17,18,15,20
	PS	6,7,9,10,14,13,8,11,16,17,18,15,12,19,20
S7	PR: d'	11,17,16,18,15,13,12,19,14,20
	PR: least significant difference	11,17,16,18,19,12,15,14,20,13
	PS	11,12,13,19,20,14,18,15,16,17
S8	PR: d'	16,13,3,4,11,7,9,12,8,14,5,6,10,15,18,17,19
	PR: least significant difference	13,16,11,3,4,7,15,12,14,9,8,5,6,18,10,17,19
	PS	19,12,18,14,15,13,11,9,5,8,4,3,16,6,10,7,17
S9	PR: d'	2,3,4,5,6,7,13,8,12,9,14,10,11,19,15,17,18,16
	PR: least significant difference	2,3,4,5,6,7,13,12,8,9,14,19,10,11,16,15,17,18
	PS	2,7,11,8,4,5,9,16,10,3,6,12,13,19,18,17,15,14
S10	PR: d'	10,11,12,13,14,15,17,16,18,19
	PR: least significant difference	10,11,12,13,14,15,17,16,18,19
	PS	10,11,12,13,14,15,16,17,19,18
S11	PR: d'	13,9,7,6,8,10,14,11,12,16,17,18,19,15,20
	PR: least significant difference	13,9,8,7,14,6,10,12,11,16,17,18,15,19,20
	PS	8,6,13,9,7,14,10,12,16,15,17,18,19,11,20

is that there is a much wider variability in the estimated pitch structure. The data from seven of the subjects in this study were similar to that presented in the two previous studies (S1, S2, S3, S4, S5, S6, S10), while the data from the remaining four subjects (S7, S8, S9, S11) were much less orderly than that previously described. The majority of the nine subjects tested by Busby *et al.* (1994) exhibited monotonic decreases in pitch with increasing electrode number. However, eight of the nine subjects tested in that study were excellent performers on speech recognition tasks, and it is possible that this may be related to the consistency and monotonicity of their pitch scaling data.

Two methods to obtain measures of electrode discriminability from the trimmed mean data were evaluated. The first method, d' analysis, is the same as that used by Busby *et al.* (1994) to analyze pitch scaling data. The second method, Student's t test analysis, is based on a statistical theory developed specifically for analyzing the differences between the means of two distributions with equal variances, when the variance of the distribution must be estimated from the experimental data. One benefit of the Student's t test method is that, unlike the d' method, it is robust to the validity of the assumptions that the distributions of the percep-

TABLE VI. Summary of speech recognition data and additional analyses for each subject. Percent agreement is measured as the percentage of time that the relative pitches between electrodes were the same across the different measures.

Subject	Average of original speech scores	Change in speech performance	Consistency measure from row sum analysis of pitch rank data	Percent agreement between row sum and d' analysis of pitch rank data	Percent agreement between pitch rank and pitch scale data
S1	64.3	Decline (4/5)	0.78	96.7	92.8
S2	28.8	Improve (5/5)	0.76	97.5	90.8
S3	38.8	Improve (4/5)	0.87	98.3	88.3
S4	40.3	Decline (3/5)	0.31	87.6	86.3
S5	6.9	Improve (3/5)	0.66	74.7	86.3
S6	31.0	Decline (5/5)	0.73	90.5	81.9
S7	18.2	Decline (5/5)	0.42	84.4	51.1
S8	22.1	Decline (5/5)	0.10	68.4	49.3
S9	23.4	Decline (4/5)	0.75	84.3	65.4
S10	35.6	Improve (5/5)	0.59	91.1	93.3
S11	31.8	Decline (5/5)	0.66	89.5	86.7

tual responses evoked by each stimulus have equal variance and follow a Gaussian distribution.

1. d' analysis

A method was proposed by Busby *et al.* (1994) where discrimination indices in the form of d' metrics are calculated from scaling data. To use this method, it is necessary to assume that the variance associated with the pitch estimates for each electrode in the array is constant. However, the estimates of the perceptual variance for each of the electrodes in this study were not always the same (evaluated for each subject using an F test, $p<0.05$). In addition, the variance in the process renders it impossible to determine adjacent electrodes that are discriminable for an ideal observer with a tonotopic pitch structure. This can be seen using the formulation provided by Busby *et al.* (1994), and assuming both that the observer operates with a standard deviation of approximately 10% of the scale, and that the pitch values are equally spaced along the scale (which maximizes the probability of being able to discriminate between all of the adjacent electrodes). Under these assumptions, it can be shown that for an ideal receiver whose number of active electrodes ranges from 10 to 20, the value of d' that will occur for adjacent, discriminable electrodes ranges from 1.11 to 0.52, respectively. Therefore, in the majority of cases, a standard deviation of 10% of the scale will not permit detection of adjacent electrodes that are, in fact, discriminable. Based on this analysis, it was decided that the d' approach was not appropriate; therefore estimates of electrode discriminability generated using this analysis method are not presented.

2. T-test analysis

In the equal variance case, the appropriate statistical test to determine whether the means of the pitch estimates for two electrodes are significantly different is a Student's t test. It is not possible to use tables of the Gaussian distribution, since the true means and variances of the sensory distributions are not known. The estimates provided by the Student's t test are robust to both the Gaussian distribution assumption and to the equal variance assumption, especially when a large number of points are used to obtain the estimates of the means.

TABLE VII. Example of calculation of electrode discriminability using the least significant difference method for subject S10. First row and column list the electrode numbers and their associated row sums generated from the stimulus response matrix. Entries marked ''Y'' indicate that the row sums for that pair of electrodes exceed the least significant difference of 15, whereas entries marked ''N'' indicate that the least significant difference is not met for those two electrodes.

Electrode and row sum score	E11: 301	E12: 257	E13: 218	E14: 192	E15: 129	E16: 102	E17: 108	E18: 87	E19: 82
E10:324	Y	Y	Y	Y	Y	Y	Y	Y	Y
E11:301	⋯	Y	Y	Y	Y	Y	Y	Y	Y
E12:257	⋯	⋯	Y	Y	Y	Y	Y	Y	Y
E13:218	⋯	⋯	⋯	Y	Y	Y	Y	Y	Y
E14:192	⋯	⋯	⋯	⋯	Y	Y	Y	Y	Y
E15:129	⋯	⋯	⋯	⋯	⋯	Y	Y	Y	Y
E16:102	⋯	⋯	⋯	⋯	⋯	⋯	N	Y	Y
E17:108	⋯	⋯	⋯	⋯	⋯	⋯	⋯	Y	Y
E18:87	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯	N

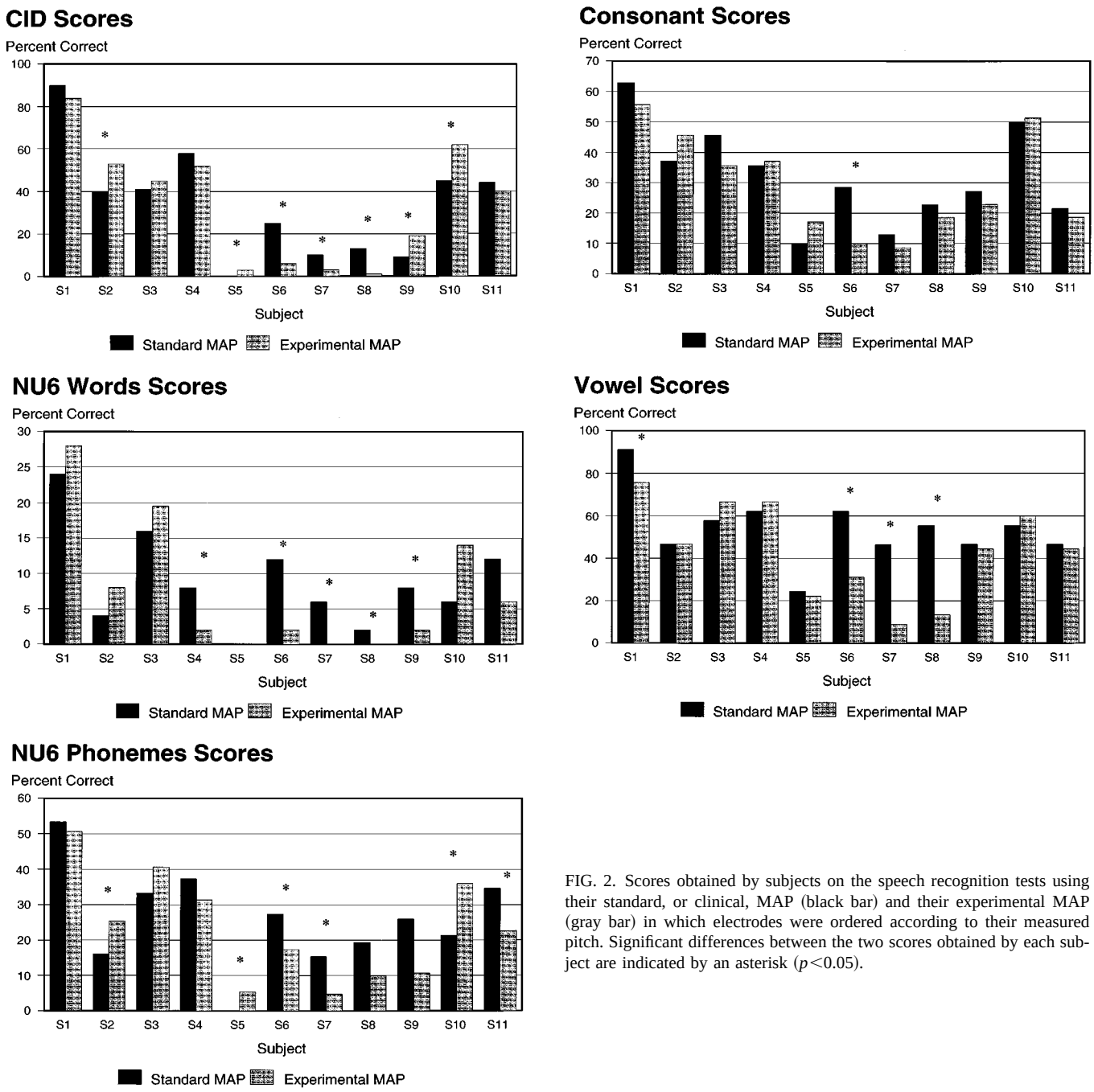


FIG. 2. Scores obtained by subjects on the speech recognition tests using their standard, or clinical, MAP (black bar) and their experimental MAP (gray bar) in which electrodes were ordered according to their measured pitch. Significant differences between the two scores obtained by each subject are indicated by an asterisk ($p<0.05$).

The results of the Student's t test analysis are also shown in Table IV. The data was generated in a manner similar to that used to generate a set of discriminable electrodes from the row sum data for the pitch ranking task. For example, the set of discriminable electrodes for subject S10 as determined using this method is {10, 11, 12, 14, 16, 17}. The Student's t test analysis allows adjacent discriminable electrodes to be determined even with a perceptual variance of 10% of the scale. As in the analysis of the d' method, if it is assumed that the pitch values are equally spaced along the scale, it can be shown that for an ideal receiver whose number of active electrodes ranges from 10 to 20, the value of the t statistic that will occur for adjacent, discriminable electrodes ranges from 2.7 to 1.27, respectively. This range of values exceeds the level required for significance; therefore adjacent discriminable electrodes can be determined using this methodology for all subjects.

C. Speech recognition task

Five different measures of speech recognition were administered to subjects while they used a MAP that employed electrodes numbered in an order determined from their pitch ranking data. This order was generated by the results of the row sum analysis for ten of the 11 subjects, and by the results of the d' analysis for subject S2. The results of the d' analysis were used for subject S2 since the row sum analysis resulted in a tonotopic ordering that matched the natural ordering, while the d' method indicated one reversal of two adjacent electrodes. Electrodes were ordered from highest

score to lowest score in the experimental MAP, instead of using the natural ordering. If a tie occurred in the row sum data, the electrodes were ordered following their natural ordering. A MAP was not created based on the pitch scaling data since the high variability in the pitch estimates produced using this method made it difficult to accurately estimate electrode pitch, and thus electrode order.

Scores obtained by the 11 subjects on the speech recognition tests while using the experimental MAP are displayed in Fig. 2. The abscissa of these plots lists the subjects; the ordinate shows the percent correct score obtained on each of the measures under the two different conditions. The standard, or clinical, MAP score is shown by the black bar and the experimental MAP score is shown by the gray bar. Few subjects improved on the speech measures when using the experimental MAP. Subjects who did improve were those who exhibited only minor deviations from a tonotopic pitch structure, such as S1, S2, and S10. It is possible that other subjects may have performed better with the experimental MAP given more experience, or if they had been programmed with a pitch-based electrode ordering when they first obtained their device. It is also possible that the results may have been different had subjects been programmed with the SPEAK speech processing strategy since it makes better use of the tonotopic location of electrodes. However, none of the subjects were programmed with the SPEAK strategy at the time of the experiment.

The binomial test described by Thornton and Raffin (1978) was used to determine if the speech recognition scores obtained while subjects used the experimental MAPs differed significantly for the 11 subjects. The results of these analyses are also shown in Fig. 2, where significance ($p < 0.05$) is indicated by an asterisk (*). The analysis revealed that there were decreases in scores between the clinical and experimental MAPs for several subjects (especially S6, S7, and S8) on all of the measures. The only improvement occurred for subjects S2 and S10 on the NU6 test scored for phonemes correct, and for subjects S2, S5, S9, and S10 on the CID everyday sentences test. Speech recognition test results are briefly summarized for each subject in Table VI, which lists the average of subjects' five original speech recognition scores, and the number of speech measures in which subjects' performance either declined or improved (not necessarily significantly) using the experimental MAP.

D. Correlation analysis

Correlation analyses were performed to statistically examine the relationship between the pitch data and speech recognition, and the results are listed in Table VIII. Spearman rank order correlation coefficients were calculated between each of the original speech measures, along with the slopes of the regression lines for the pitch rank cumulative d' curves, the row sum analysis consistency measure, and the percent agreement between the relative rankings of the pitch rank and pitch scale data. Correlations that are significant at a 0.01 level are listed in bold text, and correlations that are significant at a 0.05 level are italicized. The results are similar to those presented in Nelson *et al.* (1995), and indicate that speech recognition is slightly more correlated with the ranking than with the scaling data. Although speech recognition is somewhat correlated with the agreement between the two data sets, it does not appear to be correlated to the dimensionality of the percept elicited in the pitch ranking task.

TABLE VIII. Results of the Spearman rank order correlation coefficient calculations. Bold entries indicate significance levels of 0.01; italicized entries indicate significance levels of 0.05.

Measure	Slope of ranking data	Consistency	Percent agreement
NU6 words	**0.78**	0.57	0.31
NU6 phonemes	*0.65*	0.28	0.30
CID sentences	*0.65*	0.13	*0.65*
Consonants	*0.62*	0.48	*0.70*
Vowels	0.50	0.16	0.22
Average speech	**0.75**	0.34	0.61
Slope of ranking data	⋯	*0.69*	*0.74*
Consistency	⋯	⋯	0.53

E. Theoretical analysis

The previous subsections have described experiments concerning the issues of electrode-site/channel-independence and electrode-site/pitch-relationships in subjects implanted with a cochlear prostheses. Results from these experiments indicate that the three measures do not provide equivalent estimates for either pitch or electrode-site discriminability. In fact, the relationships between the discriminability estimates obtained in these experiments are contrary to those established for normal-hearing subjects in both frequency and intensity discrimination tasks. This section examines the decision processes that would be employed by an ideal receiver for each of the experiments.

The theoretical and experimental relationships between paired comparison (or 2IFC) and same/different tasks have been analyzed by several researchers. There are several ways to model the subject's decision processes in these procedures. However, all are based on the assumption that the sensation evoked by each stimulus can be modeled as a single, unidimensional random variable that follows a Gaussian distribution, and that the random variable associated with each stimulus has the same variance. These are the standard assumptions that have historically been used in intensity or frequency discrimination experiments in normal-hearing subjects. The elicited sensation forms the basis of the perceptual decisions that are required under each of the experimental procedures.

Depending on the model of subject behavior adopted, the performance of an ideal observer can vary by a factor of $\sqrt{2}$ on a same/different task. In one formulation of the signal detection model, the presence of a noise-free reference is assumed in the same/different task (Swets, 1959; Sorkin, 1962; Durlach and Braida, 1969; Green and Swets, 1974), which results in a discrimination metric for the paired comparison task that is $\sqrt{2}$ times larger than the same/different metric for the same stimuli. If the noise-free reference is not assumed in the same/different task, the discriminability met-

ric associated with that task is reduced by a factor of $\sqrt{2}$, and the ratio of the d' values between the paired comparison and the same/different tasks becomes 2 (Sorkin, 1962; Viemeister, 1970). To determine which of the formulations is accurate, the discriminability estimates across each of the experimental procedures have been measured for detection tasks, as well as both intensity and frequency discrimination tasks. In a study where paired comparison and same/different procedures were compared on both intensity and frequency discrimination tasks in the same subjects, Jesteadt and Bilger (1974) found that the ratio between the d' values in a paired comparison procedure and the same/different procedure was close to 2.

Both the theoretical and the experimental results for normal-hearing subjects on intensity and frequency discrimination tasks are at odds with the data obtained in the experiments described in the previous three subsections, where electrode discrimination measured using a same/different task was much better than that measured using a paired comparison task. Regardless of the model adopted, paired comparison performance is consistently higher than same/different performance in normal-hearing subjects on frequency and intensity discrimination tasks.

In each trial of the electrode discrimination experiment, subjects were asked to determine whether two stimuli were the same or different. Subjects could use any cues that were available to perform the discrimination, e.g., pitch, timbre, or even loudness if the loudness balancing procedure was ineffective. Therefore, the perceived signal is more appropriately modeled using a multidimensional random vector, where each element of the vector represents the perceived value of one of the dimensions of the signal. This is simply a multidimensional extension to the modeling approach described previously. Although Nelson *et al.* (1995) concluded that the percept elicited by electrical stimulation varied across place of stimulation is unidimensional, this conclusion contradicts anecdotal remarks and experimental data from several other studies where it has been suggested that a multidimensional percept is associated with place of stimulation (Eddington *et al.*, 1978; Shannon, 1983; Soli, 1990; McKay *et al.*, 1996). When it is assumed that the perceptual variance in each of the dimensions is constant, and that the dimensions are uncorrelated, performance can still be quantified in terms of a d' statistic. This statistic is equivalent to the square root of the sum of the squared value of each of the individual discriminability indices. Therefore, if the percept is in fact multidimensional, the discriminability index of the multidimensional signal is larger than the discriminability index for any of the dimensions taken by itself.

III. DISCUSSION

Shannon (1983) suggested that irregular nerve survival patterns and irregular current distributions due to variability in impedance paths could generate complex, unpredictable patterns of neural activity that in turn generate a complex and unpredictable structure of the pitch percept. Since speech coding strategies for cochlear implants require the frequency bands in speech to be mapped to electrodes in an orderly progression, measuring the pitch structure along the electrode array and assigning frequency bands to electrodes based on the results was expected to improve speech recognition. However, the results of the speech recognition task from this experiment indicated the surprising result that for the majority of subjects, such a remapping was harmful instead of beneficial. Since the subjects' electrodes were not initially mapped for a tonotopic pitch structure, it is possible that subjects learned to understand speech in a juxtaposed frequency space. Then, when a true tonotopic organization was put in place, subjects may have needed time to adjust to and learn the new pitch structure, even though the new structure might be closer to what they experienced before they were deafened (Busby *et al.*, 1994; Tyler *et al.*, 1986). To investigate this possibility, subjects would have to use the experimental MAP for a period of time, or the difference between the two mapping approaches could be assessed at the time of the initial programming of the device.

Previous studies investigating the percepts evoked by individual stimulation of each electrode in the array of implanted subjects have concluded that the percept is related to the pitch dimension of acoustical stimulation. This percept is assumed to be unidimensional, and is thus used in some cases to generate measures of electrode discriminability. The structure of the pitch percept has been reported to vary in an orderly, or monotonic, fashion with position along the electrode array in the majority of subjects, and thus it is commonly assumed each electrode corresponds to an independent channel of information.

Results from both the pitch ranking and pitch scaling tasks in this study are more variable than the results reported in the majority of the previous studies. Although subjects could perform the task based on the instructions to use the pitch characteristic of the signal, the structure of the results often did not vary in an orderly fashion with position along the electrode array. In addition, the analyses commonly applied to pitch ranking or scaling data to infer electrode discriminability did not correspond to the results from the electrode discrimination experiment. Thus it appears that electrode discrimination cannot be inferred based on data from either a paired comparison procedure to measure pitch ranking or a scaling procedure to measure pitch structure. In addition, these analyses indicated *more* indiscriminable electrodes than the electrode discrimination task, which in turn indicates that electrodes do not correspond to independent channels of information.

Since the standard deviation of the pitch estimates in the pitch scaling task was on the order of 10% of the scale, the power of the statistical estimation of the electrode order and electrode discriminability based on the d' analysis is extremely limited, and is not the best approach. In the equal variance case, the appropriate test to determine whether the means of the pitch estimates for two electrodes are significantly different is a Student's t test. The estimates provided by the t test are robust when the Gaussian distribution assumption and the equal variance assumption are invalid. Since the pitch scaling data often violates both of these assumptions, this is an additional rationale for using the t-test formulation. In addition, some other procedures for generating a pitch order, such as a sorting algorithm, should be

investigated in order to evaluate the validity of the electrode orders generated using the two procedures evaluated in this study.

There are several factors that may reduce the validity of the analyses applied to the pitch ranking data. An investigation of the self-consistency of the data obtained from the subjects in this study indicated that it is possible that the percept is multidimensional. McKay *et al.* (1996) found similar results. Shannon (1983) suggested that pitch evaluation may be difficult if each electrode has a characteristic timbre in addition to its characteristic pitch. If this is so, subjects may find it difficult to make comparisons based solely on the pitch dimension when the sounds differ significantly in a qualitative fashion. In addition, if the percepts elicited by stimulation of the electrodes in the array vary in multiple dimensions, the assumptions required for the validity of the d' analysis are violated, and thus the results will not be accurate.

The dimensionality of the evoked percept also affects the row sum method of analyzing the pitch rank data, since it may introduce circularity into the preference structure, and result in an invalid electrode ordering. There is considerable evidence that rank order is invariant to experimental method under certain stimulus conditions. Thus the two procedures used in this study should generate the same estimates of the rank order of electrode pitch as long as a single attribute of the stimulus, in this case pitch, can be used to make the required judgments. However, if transitivity in the stimulus ordering determined by the ranking procedure is violated, the rank order generated by the two methods will not necessarily be the same. It is possible that both the multidimensionality of the percept and the high levels of variance in the pitch estimates inherent in the pitch scaling task contributed to the differences between the scaling and ranking data.

Although the electrode discrimination task appeared to be more sensitive to discriminable electrodes than the paired comparison task, the analysis of the decision processes indicated that for a unidimensional percept, the paired comparison task should be more sensitive to discriminable electrodes. However, the structure of the electrode discrimination task allows subjects to utilize any stimulus dimension where discriminable information may exist. Since the consistency analysis of the ranking data indicated that the data might be multidimensional, the disparity between the single-dimensional theoretical results and the experimental results may be reconciled using this multidimensional approach. It will be important in future studies to investigate the structure of the perceptual space using multidimensional scaling to determine the appropriate model.

ACKNOWLEDGMENTS

The authors would like to thank the subjects who participated in this study for their time, effort, and feedback. We also would like to thank Dr. R. Shannon and Dr. R. Tyler for their constructive comments on a previous version of this manuscript. This research was supported by NIH Grant No. NIDCD NS 21440.

[1]A MAP is the set of data that is used to translate acoustic signal features to stimulus parameters for subjects implanted with the Nucleus/Cochlear Corporation device. Among other things, a MAP contains the mapping between the frequency bands associated with speech and the electrodes that are available for stimulation.

[2]These subjects also participated in the electrode discrimination experiment described in Collins *et al.* (1994). The same numbering scheme was used for the subjects in both experiments. There was approximately a 4-month interval between the conclusion of the electrode discrimination experiment and the beginning of this experiment.

[3]All possible electrodes were not used in subjects' clinically programmed MAP's for a variety of reasons, including unpleasant sound associated with stimulation, partial insertion of the electrode array, and facial nerve stimulation.

Braida, L. D., and Durlach, N. I. . (**1972**). ''Intensity Perception. II. Resolution in One-Interval Paradigms,''J. Acoust. Soc. Am. **51,** 483–502.

Busby, P. A., Whitford, L. A., Blamey, P. J., Richardson, L. M., Clark, G. M. (**1994**). ''Pitch perception for different modes of stimulation using the Cochlear multiple-electrode prosthesis,''J. Acoust. Soc. Am. **95,** 2658–2669.

Clark, G. M., Shepherd, R. K., Franz, B. K., Dowell, R. C., Tong, Y. C., Blamey, P. J., Webb, R. L., Pyman, B. C., McNaughton, J., and Bloom, D. M (**1988**). ''The histopathology of the human temporal bone and auditory central nervous system following cochlear implantation in a patient. Correlation with psychophysics and speech perception results,'' Acta Otolaryngol. Suppl. (Stockholm) **448**, 1–65.

Cochlear Corporation (**1993**). *Audiologist's Handbook* (Cochlear Corp., Englewood, Colorado).

Collins, L. M., Zwolan, T. A., O'Neill, J. C., and Wakefield, G. H. (**1994**). ''Analysis of electrode pair confusions and implications for speech recognition in cochlear implant subjects,'' Assoc. Res. Otolaryng. Abstracts, No. 642.

Crosby, P. A., Seligman, P. M., Patrick, J. F., Kuzma, J. A., Money, D. K., Ridler, J., and Dowell, R. (**1984**). ''The Nucleus multi-channel implantable hearing prosthesis,'' Acta Otolaryngol. Suppl. (Stockholm) **411**, 111–114.

David, H. A. (**1988**). *The Method of Paired Comparisons* (Oxford U. P., New York, 1988).

Davis, H., and Silverman, S. R. (**1978**). *Hearing and Deafness* (Holt, Rinehart and Winston, New York).

Dorman, M. F., Smith, L., McCandless, G., Dunnavant, G., Parkin, J., and Dankowski, K. (**1990**). ''Pitch scaling and speech understanding by patients who use the Ineraid cochlear implant,'' Ear Hear. **11** (4), 310–315.

Durlach, N. I., and Braida, L. D. (**1969**). ''Intensity Perception. I. Preliminary Theory of Intensity Resolution,''J. Acoust. Soc. Am. **46,** 372–383.

Eddington, D. K., Dobelle, W. H., Brackman, D. E., Mladejovsky, M. G., and Parkin, J. (**1978**). ''Place and periodicity pitch by stimulation of multiple scala tympani electrodes in deaf volunteers,'' Trans. Am. Soc. Artif. Intern. Organs **XXIV**, 1–5.

Green, D. M., and Swets, J. A. (**1974**). *Signal Detection Theory and Psychophysics* (Krieger, New York).

Jesteadt, W., and Bilger, R. C. (**1974**). ''Intensity and frequency discrimination in one- and two-interval paradigms,''J. Acoust. Soc. Am. **55,** 1266–1276.

Jesteadt, W. (**1980**). ''An adaptive procedure for subjective judgements,'' Percept. Psychophys. **28**(1), 85–88.

McKay, C. M., McDermott, H. J., and Clark, G. M. (**1996**). ''The perceptual dimensions of single-electrode and nonsimultaneous dual-electrode stimuli in cochlear implantees,''J. Acoust. Soc. Am. **99,** 1079–1090.

Nelson, D. A., Van Tasell, D. J., Schroder, A. C., Soli, S., and Levine, S. (**1995**). ''Electrode ranking of 'place pitch' and speech recognition in electrical hearing,''J. Acoust. Soc. Am. **98,** 1987–1999.

Pauka, C. K. (**1989**). ''Place-pitch and vowel-pitch comparisons in cochlear implant patients using the Melbourne–Nucleus cochlear implant,'' J. Laryngol. Otol. Suppl. **19**, 1–31.

Shannon, R. V. (**1983**). ''Multichannel electrical stimulation of the auditory nerve in man: Basic psychophysics,'' Hear. Res. **11**, 157–189.

Soli, S. D. (**1990**). ''Perceptual Evaluation of a Neurally Based Encoding Strategy for Cochlear Implants,'' in *Cochlear Implants*, edited by J. F. Miller and F. A. Spelman (Springer-Verlag, New York).

Sorkin, R. D. (**1962**). ''Extension of the theory of signal detectability to matching procedures in psychoacoustics,''J. Acoust. Soc. Am. **34,** 1745–1751.

Swets, J. A. (**1959**). ''Indices of signal detectability obtained with various psychophysical procedures,'' J. Acoust. Soc. Am. **31,** 511–513.
Thornton, A. R., and Raffin, M. J. (**1978**). ''Speech discrimination scores modeled as a binomial variable,'' J. Speech Hear. Res. **21**, 507–518.
Tillman, T. W., and Carhart, T. (**1966**). An expanded test for speech discrimination utilizing CNC monosyllabic words: Northwestern University Auditory Test No. 6, Tech. rep. no. SAM-TR-66-55, USAF School of Aerospace Medicine, Brooks Air Force Base, Texas.
Tong, Y. C., and Clark, G. M. (**1985**). ''Absolute identification of electric pulse rates and electrode positions by cochlear implant patients,''J. Acoust. Soc. Am. **77,** 1881–2429.
Tong, Y. C., Clark, G. M., Blamey, P. J., Busby, P. A., and Dowell, R. C. (**1982**). ''Psychophysical studies for two multiple-channel cochlear implant patients,''J. Acoust. Soc. Am. **71,** 153–160
Tong, Y. C., Millar, J. B., Clark, G. M., Martin, L. F., Busby, P. A., and Patrick, J. F. (**1980**). ''Psychophysical and speech perception studies on two multiple channel cochlear implant patients,'' J. Laryngol. Otol. **94**, 11241–1256.
Torgerson, W. S. (**1958**). *Theory and Methods of Scaling* (Wiley, New York).
Townshend, B., Cotter, N., Van Compernolle, D., and White, R. L. (**1987**). ''Pitch perception by cochlear implant subjects,''J. Acoust. Soc. Am. **82,** 106–115.
Tyler, R. S., Preece, J. P., Lansing, C. R., Otto, S. R., and Gantz, B. J. (**1986**). ''Previous experience as a confounding factor in comparing cochlear implant processing schemes,'' J. Speech Hear. Res. **29**, 282–287.
Tyler, R. S. (**1991**). ''What can we learn about hearing aids from cochlear implants?,'' Ear Hear. **12**(6), 177S–186S.
Viemeister, N. F. (**1970**). ''Intensity Discrimination: Performance of Three Paradigms,'' Percept. Psychophys. **8**, 417–419.

Acoustic characteristics of the piriform fossa in models and humans

Jianwu Dang
ATR Human Information Processing Research Laboratories, 2-2 Hikaridai Seika-cho, Soraku-gun, Kyoto, 619-02 Japan

Kiyoshi Honda
ATR Human Information Processing Research Laboratories, 2-2 Hikaridai Seika-cho, Soraku-gun, Kyoto, 619-02 Japan and University of Wisconsin, 1500 Highland Avenue, Madison, Wisconsin 53705-2280

(Received 22 March 1996; accepted for publication 12 September 1996)

The piriform fossa forms the bottom part of the pharynx and acts as a pair of side branches of the vocal tract. Because of its obscure form and function, the piriform fossa has usually been neglected in the current speech production models. This study examines the geometric and acoustic characteristics of the piriform fossa by means of MRI-based mechanical modeling, *in-vivo* experiments and numerical computations. Volumetric MRI data showed that the piriform fossa is 2.1 to 2.9 cm^3 in volume and 1.6 to 2.0 cm in depth for four Japanese subjects (three males and one female). The results obtained from mechanical models showed that the piriform fossa contributes strong troughs, i.e., spectral minima, to speech spectra in a region of 4 to 5 kHz. The antiresonances were identified with increasing frequency when water was injected into the piriform fossa of human subjects in *in-vivo* experiments. Antiresonances obtained from the experiments and simulations were confirmed to be consistent with those in natural speech within 5%. Acoustic measurements and simulations showed that the influence of the piriform fossa extends to the lower vowel formants in addition to the local troughs. This global effect can be explained by the location of the fossa near the glottal end of the vocal tract.
[S0001-4966(97)04301-4]

PACS numbers: 43.70.Bk, 43.70.Aj, 43.72.Ct [AL]

INTRODUCTION

The piriform fossa, also referred to as the sinus pyriformis or Recesses Piriformis in its singular forms, consists of a pair of bilateral cavities in the hypopharynx located just above the esophageal entrance. Therefore, this space clearly belongs to the vocal tract. Except for a few previous studies (e.g., Fant, 1960; Sundberg, 1974; Fant and Båvegård, 1995), however, this structure has not been widely recognized as a functional part of the vocal tract in speech production. There seem to be at least two reasons for this neglect: one is the lack of morphological data for acoustic modeling; the other is its small size, which would appear to relegate it to minor importance as a side branch of the vocal tract.

In earlier studies, Fant (1960) made a valuable observation on the role of the piriform fossa with respect to vowel formants. He used an x-ray-based simulation to demonstrate that the acoustic effect of the piriform fossa can be seen as causing significantly lowered formant frequencies of vowels. Flach and Schwickardie (1966) reported an experimental result indicating that the piriform fossa causes sound attenuation in the frequency region above 1500 Hz. On the other hand, Mermelstein (1967) argued that theoretically the piriform fossa should not cause any major effect below 4 kHz. However, Sundberg (1974) pointed out that the piriform fossa played a significant role in forming the singing formant between 2 and 3 kHz. Lin (1990) showed that acoustic effects of the piriform fossa can vary with vowels' articulations by using a numerical model of a side branch: the effects were most prominent in $F1$ of open vowels. Fant and Båvegård (1995) found that the piriform fossa could significantly alter the density of the pole in the 3 to 5 kHz region and introduce a zero at about 5200 Hz. To obtain a realistic transfer function, they set the volume of the piriform fossa to be 50% larger than the measured volume in order to compensate for the open end correction of the piriform fossa due to the area discontinuities near the open end.

The advent of the magnetic resonance imaging (MRI) technique has made it possible to display the vocal tract, including the piriform fossa, as a three-dimensional image. A number of MRI-based studies have investigated the 3-D configuration of the vocal tract for vowels and sustained consonants (Baer *et al.*, 1991; Moore, 1992; Dang *et al.*, 1994; Narayanan *et al.*, 1995). These studies aimed at obtaining a precise area function of the vocal tract and tested the accuracy by matching the MRI-derived transfer functions to real speech spectra for the same subjects. Baer *et al.* (1991) estimated vocal tract transfer functions with and without the piriform fossa and compared the estimated formants with those from the natural utterances. They demonstrated a decrease in vowel formants due to the piriform fossa, in agreement with Fant's work, even though in their study the piriform fossa was treated not as a side branch but as an additional volume of the pharyngeal tube. Davies *et al.* (1993) used the data by Baer *et al.* (1991) to compute the transfer functions of the vocal tract with and without the

fossa. The $F1$ and $F2$ of the vowel /a/ decreased by about 5% when the fossa was incorporated in the vocal tract as a side branch. The above studies suggest that the piriform fossa does cause significant effects on speech spectra; furthermore, the studies imply that estimation of vocal tract transfer functions from the area function could be erroneous if consideration is not given to the piriform fossa. In accordance with this view, many studies have shown that an MRI-derived transfer function tends to exhibit slightly higher formant frequencies than those from real speech. Yang *et al.* (1994) explained the discrepancy between estimated and measured formants by underestimation of the vocal tract length in the MRI data. Judging from the earlier studies, however, the piriform fossa might be a critical factor in accounting for the discrepancy.

The above studies have provided conclusions resulting in both agreements and disagreements with respect to the detailed effect of the piriform fossa, and the exact role of this structure in natural speech is not yet very clear. The present study aims to explore the acoustical characteristics of the piriform fossa by means of geometric and acoustic measurements. Volumetric MRI data were obtained for four subjects. Mechanical models were constructed to replicate a partial vocal tract from the glottis to the bend of the velum based on the volumetric data from one of the subjects. Those models were used to examine the acoustical effects of the piriform fossa on the radiated sounds by artificially changing the air volume of the cavities. A similar approach was repeated on human subjects by injecting water into the fossa during sustained vowel production. Antiresonances derived from the geometric data were compared with those from natural speech produced by the subjects.

I. MORPHOLOGICAL MEASUREMENT OF THE PIRIFORM FOSSA

Determining the acoustic characteristics of the piriform fossa requires detailed morphological information. In this study, we made volumetric MRI measurements to obtain the dimensions of the piriform fossa and then developed mechanical and numerical models based on the results of the volumetric analysis.

A. Method for MRI measurement of the piriform fossa

Volumetric MRI data of the vocal tract were collected from four Japanese subjects (three males and one female) by the standard spin echo method. The relaxation time (TR) was 800 ms and the excitation time (TE) was 18 ms. A 25 cm×25 cm field of view for a slice of an image was digitally represented by a 256×256 pixel matrix (see Dang *et al.*, 1994). The MRI data consisted of 26 slices for the transverse plane and 24 slices for the coronal plane. The slice thickness was 0.5 cm for both orientations. Each slice was resampled into an image of 250×250 pixels so that each pixel represented a 0.1×0.1 cm square. The resampled slices were further processed to form volumetric data consisting of 0.1 cm^3 voxels by means of interslice image interpolation. The procedure was performed with a commercial software (VoxelView) on a workstation (IRIS Indigo).

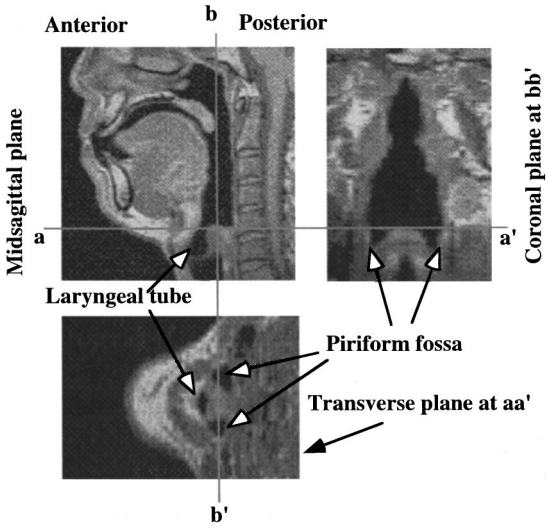


FIG. 1. A three-plane image of the volumetric MRI data of the vocal tract. The coronal and transverse planes show the gross shape of the piriform fossa. Line aa' indicates the arytenoid apex plane.

B. Three-dimensional shape of the piriform fossa

Figure 1 shows a three-plane image of the volumetric MRI data of the vocal tract. The coronal and transverse planes show that the piriform fossa consists of two bilateral cavities located on either side of the laryngeal tube. The transverse section indicates that each cavity opens to the pharynx at the horizontal plane passing through the apex of the arytenoid prominence, which is marked by the line aa' in the figure.

Figure 2 shows a sketch of the vocal tract in the vicinity of the piriform fossa. The bottom of the vocal tract divides into three small branches: the laryngeal tube (vestibule of the larynx) and the bilateral cavities of the piriform fossa. The laryngeal tube is a short conduit of about 2 cm long for our

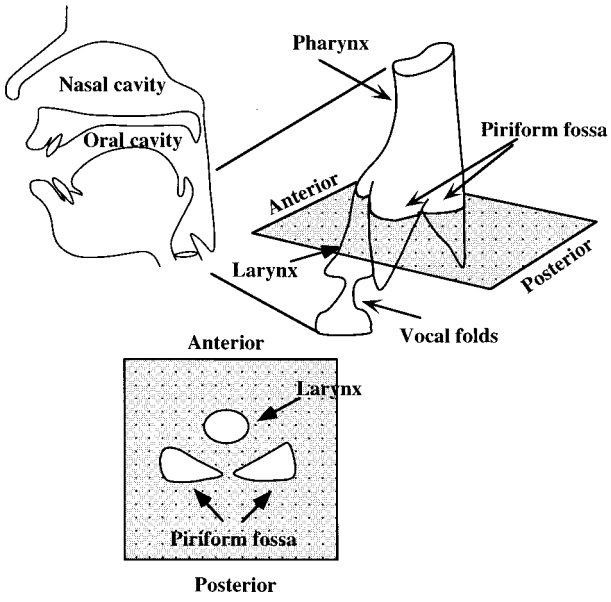


FIG. 2. A diagram of the vocal tract shape in the vicinity of the piriform fossa. The horizontal plane shows a section near the open end of the piriform fossa.

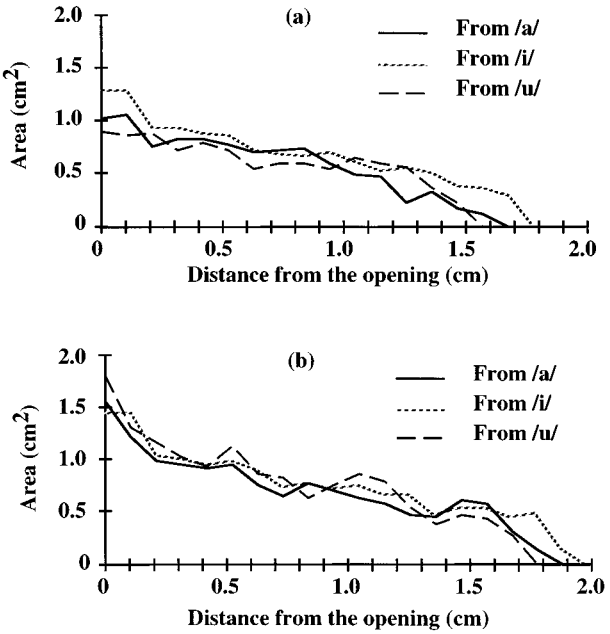


FIG. 3. Area functions of the piriform fossa based on the MRI data from subject KH. The areas are measured from the arytenoid apex plane to the bottom of the piriform fossa: (a) the left cavity and (b) the right cavity.

male subjects, bounded by the vocal folds at the bottom and by the aryepiglottic fold at the top. The piriform fossa is located behind the laryngeal tube, and forms the bottom of the pharynx. The cavities of the fossa open into the main pharynx, sharing the aryepiglottic folds as a common boundary with the laryngeal tube. However, the opening of the piriform fossa is not demarcated by a clear border because the walls of the piriform fossa continue to the vallecula medially and to the pharyngeal walls laterally and posteriorly. With reference to the direction of sound propagation, the piriform fossa comprises a pair of side branches bifurcating from the vocal tract. Since the shape of the cavities is relatively stable during phonation, the piriform fossa is expected to contribute relatively invariant antiresonances to the transmission characteristics of the vocal tract.

Dimensions of the left and right cavities of the fossa were measured for the Japanese vowels /a/, /i/, and /u/ based mainly on the transverse data with reference to the coronal data. Depth of the fossa (the distance from the bottom to the arytenoid apex plane) was determined using the coronal data. Cross-sectional shapes of the fossa were measured on the transverse slices for the entire extent from the bottom of the piriform fossa to the arytenoid apex. The results are shown as area functions in Fig. 3 for the left and right cavities of subject KH. The depth and volume of the piriform fossa are shown in Table I for the four subjects. The depth values over the three vowels range from 1.6 to 2.0 cm for all subjects. The typical value is about 1.8 cm. HH showed a relatively larger variation in depth across the vowels than the other subjects. The volume ranges from 1.0 to 1.4 cm^3 for each side of the cavities. The female subject showed a smaller cavity volume than the males. Within a subject, the variation of the volume across vowels is the smallest for RY and the largest for SM. Generally, the morphological data showed no significant asymmetry between the left and right sides.

TABLE I. Depth (D) and volume (V) of the left and right cavities of the piriform fossa obtained from volumetric MRI data for four subjects.

Subject	Vowel	Vr(cm^3)	Vl(cm^3)	Dr(cm)	Dl(cm)
HH(M)	/a/	1.401	1.346	1.7	1.6
	/i/	1.027	0.977	2	1.7
	/u/	1.248	1.167	2	2
KH(M)	/a/	1.479	1.287	1.9	1.8
	/i/	1.481	1.437	1.9	1.9
	/u/	1.345	1.065	1.8	1.7
SM(M)	/a/	1.147	1.112	1.7	1.8
	/i/	1.621	1.77	1.9	2
	/u/	1.464	1.49	1.7	1.8
RY(F)	/a/	1.159	0.921	1.7	1.8
	/i/	1.08	0.953	1.7	1.7
	/u/	1.043	0.997	1.8	1.7

C. Acoustically effective cavities of the piriform fossa

The MRI data show that the piriform fossa extends slightly above the arytenoid apex plane and continues to the aryepiglottic folds in /a/ or further to the lateral glossoepiglottic folds in /i/ and /u/, depending on the position of the epiglottis. Therefore, the extended portion above the arytenoid apex plane shapes the acoustically effective opening end of the piriform fossa. Figure 4 illustrates the coronal MRI image of the piriform fossa and the schematized cavity shape by a cone and cylinder model. Here, the piriform fossa is modeled by two cascaded portions: the lower ''cone'' portion from the bottom of the fossa to the arytenoid apex plane; and the upper ''cylinder'' portion of the oblique output end above the plane. In the figure, point A is the intersection of the effective open end and the arytenoid apex plane, and point B is the intersection of the open end and the lateral wall of the pharynx. L_B represents the length from the arytenoid apex plane to point B. According to the MRI data, L_B can be roughly evaluated by an empirical formula:

$$L_B = 1.3/D, \tag{1}$$

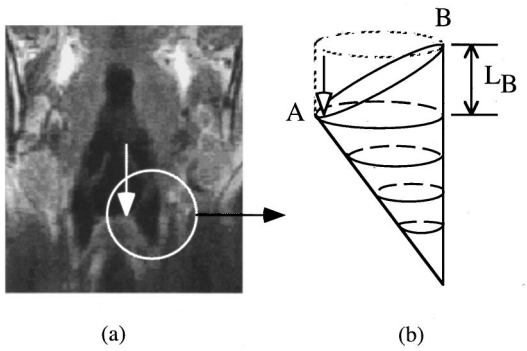


FIG. 4. A coronal MRI slice of the piriform fossa (a) and a diagram of the effective cavity of the piriform fossa (b). (White arrows indicate the arytenoid apex plane; L_B is length of the cylinder portion above the plane.)

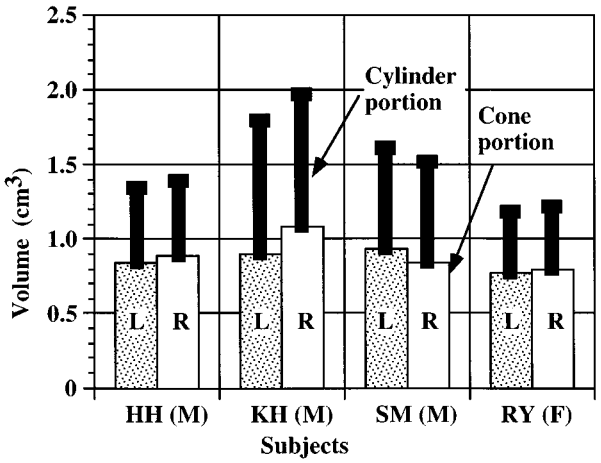


FIG. 5. Volumes of the lower ''cone'' and upper ''cylinder'' portions of the piriform fossa for four subjects (L: left cavity, R: right cavity).

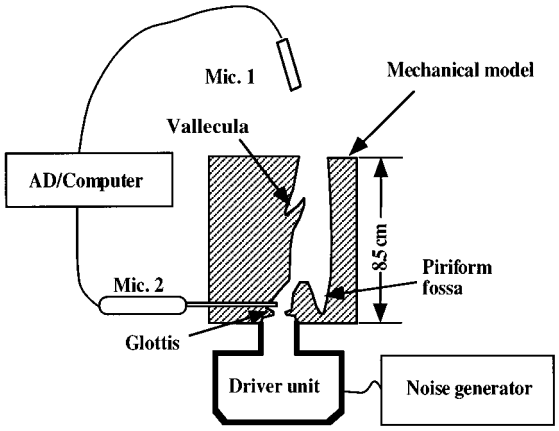


FIG. 6. Experimental setup for measuring the acoustic effects of the piriform fossa using a mechanical model constructed from volumetric MRI data.

where D denotes the diameter of the cross-sectional area of the piriform fossa in the arytenoid apex plane and the area is assumed to be circular.

Using Eq. (1), the effective cavity of the piriform fossa was calculated for the four subjects. Figure 5 shows the volumes of the cylinder and cone positions for the left and right cavities, which are averaged over three vowels /a/, /i/, and /u/. The ratio of the cylinder to the cone portions ranged from 0.4 to 0.8 for the four subjects. This result supports the suggestion of Fant and Båvegård (1995), who used an estimated volume in their study that was about 50% larger than the measured one (i.e., the cone portion).

II. ACOUSTIC MEASUREMENT USING THE MECHANICAL MODELS

The volumetric MRI data allow us to examine the acoustic characteristics of the piriform fossa by either mechanical or numerical models. The numerical model requires parameters to define the acoustic properties of the geometry, such as an open end correction for the cavities of the piriform fossa. However, the estimation of accurate parameter values is difficult because of the complex shape of the structure. For this reason, a mechanical model of the vocal tract was first used in this study. The coefficient of the open end correction is derived from the experiments on the mechanical models and used in numerical simulations in the latter sections. The mechanical models also served in a pilot experiment of injecting water into the piriform fossa of the human subjects.

A. Experimental procedure

Three mechanical models of the vocal tract for the Japanese vowels /a/, /i/, and /u/ were constructed based on the MRI data obtained from subject KH to be used for acoustic experiments. The air-tissue pattern in the transverse MRI image was carved out of a vinyl chloride plate with a size of 7 cm×7 cm×0.1 cm. The vertical part of the vocal tract was replicated by the carved plates. The models were about 8.5 cm long, from about 1 cm below the glottis to the vocal tract bend near the velum. The openings of the piriform fossa in the models were about 3.5 cm from the input end (glottal side) and about 5 cm from the radiating end (velar side). The vallecula was at about 6 cm from the input end in the models for vowels /i/ and /u/, while it was collapsed in vowel /a/. Hereafter, the mechanical model is referred to as the M model and the numerical model used in the later sections is called the N model. The M models are labeled M model /a/, M model /i/, and M model /u/, corresponding to their original vocal tract shapes, respectively.

The acoustic effects of the piriform fossa on the transfer function of the M models were examined by an experiment using the two-point sound pressure method (Dang and Honda, 1996a). The sound pressures were recorded at an inside point near the glottis and at the radiating end, while the model was excited by an external source sound. The transfer function obtained from the two-point pressure is known to exactly reflect the acoustic properties of the tube segment between the two points, since the effects of the sound source and the geometric shape upstream from the inside point are excluded from the estimation. Figure 6 shows the experimental setup for measuring the transfer function of the M models. The sound pressures inside and outside the M model were recorded by two microphones. Microphone M1, a B&K 4003, was placed about 6 cm away from the radiating end. Probe microphone M2, a B&K 4182, was used to record the pressure inside the model just above the glottis via a flexible probe tube, which was inserted through a hole in the lateral wall of the M model. The probe tube with a matched impedance to the microphone was 5 cm long, 0.165 cm in outer diameter and 0.076 cm in inner diameter. A white-noise signal was generated by an FG-143 function generator (NF Circuit Design Block Co.) and fed into the M model using an SG-505FRP horn driver unit (Goto Unit Co.). The joint between the M model and the throat of the horn driver was sealed with plasticine to prevent sound leakage.

The experiment on the M models was conducted in an anechoic room under four conditions: both cavities filled with plasticine, either the right or left cavity filled, and both cavities empty. Two pressure signals were sampled at a rate

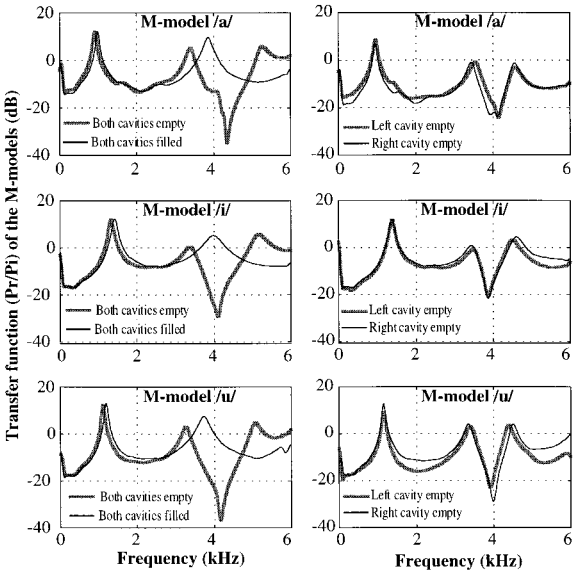


FIG. 7. Pressure-to-pressure transfer function (Pr/Pi) of the mechanical models when one or both cavities are empty, or filled with plasticine (Pr: radiated pressure; Pi: intrapressure).

of 44.1 kHz. A room temperature of 20 °C was maintained during the measurements, and the sound velocity is expected to be 34 300 cm/s. The outside sound pressure near the joint was about 25 dB lower than that at the radiating end, which implies that sound leakage from the joint was negligible in this measurement. An FFT-derived cepstrum analysis (Imai and Abe, 1979) was applied to obtain spectral envelopes of recorded signals. The cepstrum coefficients were weighted by a Hamming window of 0.05 second for white noise and 1.2 times the fundamental period for speech sound.

B. Changes in the transfer function by varying the models' cavity volume

Transfer functions of the three M models were estimated by the two-point sound-pressure method under the four conditions described above. The results are shown in Fig. 7 for the M models (left) with and without the both cavities and (right) with either the left or right cavity alone. Note that though the models represent only a part of the vocal tract, they exhibit the same antiresonances on the transfer function as they do in the whole vocal tract as far as the side branches in the models are concerned.

In Fig. 7(left), the thick lines show the spectra obtained under natural conditions, i.e., with both cavities empty; the thin lines indicate the results measured when the two cavities were filled with plasticine. Deep troughs are seen at about 4 kHz in the spectra obtained under natural conditions for all the M models of /a/, /i/, and /u/. When the cavities were filled bilaterally, the troughs were replaced by peaks, resulting in an increase of more than 30 dB in spectral level near 4 kHz. The large difference in the spectra indicates that the piriform fossa behaves as a side branch of the vocal tract and causes a significant effect on the transfer function. In Fig. 7(right), the thick lines show the spectra obtained from the M models with the left cavity empty, and the thin lines represent the

TABLE II. Antiresonance frequencies caused by the empty cavities of the piriform fossa in the transfer function of the mechanical models (Hz).

M models	Both cavities	Left cavity	Right cavity
M model /a/	4359	4148	4100
M model /i/	4054	3855	3867
M model /u/	4160	3925	3996

results with the right cavity empty. The figure shows almost identical spectral shapes for the left and right cavities except for a small discrepancy in the troughs' shape. This implies that the left and the right cavities are acoustically symmetrical in this subject. This observation is consistent with the morphological observation discussed in Sec. I.

The frequencies of these troughs were measured for each condition and are listed in Table II as the antiresonance frequencies of the models' piriform fossa. The antiresonance frequencies ranged from 4050 to 4360 Hz in the case of both cavities empty. The M model /a/ shows a slightly higher antiresonance frequency than the others. In all three M models, the antiresonance frequency caused by a single cavity was lower than that caused by both cavities together. This phenomenon can be explained by a larger open end effect of the cavity because the area ratio of the piriform fossa to the pharyngeal cavity at their boundary is larger in the case of a single cavity than it is under natural conditions. This implies that an appropriate open end correction is necessary for incorporating the piriform fossa into the transmission line model. In this regard, measurements and numerical computations have been conducted to obtain the optimum value of the end correction coefficient (Dang and Honda, 1996b). The measurements were carried out on the M models while the models' fossa was filled with plasticine from the bottom (0 cm) up to a level of 0.9 cm in 0.1-cm intervals. In the computations, the piriform fossa was schematically represented by the cone and cylinder portions as shown in Fig. 4. The length of the cylinder portion of the fossa (L_B) was derived by Eq. (1). The antiresonance frequency was then computed for each M model. The optimum value of the open end correction coefficient was estimated by matching the computed antiresonances to the measured ones. The value of 0.75 met the given condition that the computed antiresonances were consistent with those measurements within 4% for the three M models.

Figure 7 also shows a global acoustic effect of the piriform fossa in addition to the local troughs near 4 kHz. In the figure, a relatively small but non-negligible difference in the spectral peaks can be seen in the lower frequency region below 3 kHz. Peak frequencies of the model's transfer function are listed in Table III. As a general tendency, the piri-

TABLE III. Frequencies of the first two peaks in the transfer function (Pr/Pi) of the mechanical models (Hz).

	Empty cavities		Filled cavities	
M models	$P1$	$P2$	$P1$	$P2$
M model /a/	925	3398	1007	3855
M model /i/	1335	3348	1429	3978
M model /u/	1113	3246	1207	3714

form fossa lowers the frequencies of the peaks that are in the wide frequency region below the antiresonance of the fossa. The changes in the frequency of the first peak are 82 Hz for /a/, 94 Hz for /i/, and 94 Hz for /u/. For the second peak, they are 457 Hz for /a/, 630 Hz for /i/, and 468 Hz for /u/. The results show that the piriform fossa lowers the frequency of the spectral peaks about 8% in the region near 1 kHz and about 13% in the region above 3 kHz. Note that since the M models represent only a part of the vocal tract, the peaks do not correspond to true vowel formants. However, the observed frequency shifts should be representative of those of true formants that are located in the frequency region.

C. Changes in antiresonances of the piriform fossa with water injection

The direct approach used to explore the acoustic effects of the model's piriform fossa is not applicable to living human subjects. However, a similar experiment could be performed by injecting water into the fossa of human subjects during sustained phonation. Under these conditions, the effects of the piriform fossa could be viewed as increasing antiresonance frequency with a decrease in the air volume of the cavities. This hypothesis was first tested by a pilot experiment on the mechanical models.

The experimental setup for the pilot experiment was based on Fig. 6, where probe microphone M2 was removed and a membrane of polyvinylidene chloride was placed between the M model and the neck of the driver unit to prevent water from leaking into the driver unit. A piece of rubber foam was placed into the model's glottis to reduce acoustic influence of the air volume change in the piriform fossa on the driving impedance of the model. The joint between the M model and the driver unit was sealed with plasticine. The amplitude of the excitation sound source was kept constant during the measurement. Radiated sound from the model was recorded by microphone M1 while water was injected into the piriform fossa at a constant rate through a thin flexible tube.

Figure 8 illustrates running spectra of the recorded sound from M model /a/. Each curve in this figure corresponds to a frame of 46-ms duration, and the frame-to-frame interval was about 130 ms. Since the excitation source was constant during the measurements, the changes in the spectra are considered to be caused by the changes in the air volume of the fossa only. In the initial condition with no water injected, the spectral curves show a constant trough at about 4300 Hz. After about 1.2 s, water is injected into the right cavity of the piriform fossa. The trough moves toward a higher frequency as the air volume of the right cavity decreases due to injected water. Then, another trough by the left cavity is exposed and remains at about 4100 Hz. The trough caused by the right cavity disappears from the spectra as the cavity is completely filled with water after about 4 s. Furthermore, the trough associated with the left cavity begins to increase in frequency as the water flows into the left cavity. Unlike the gradual motion for the right cavity's trough, the one caused by the left cavity moves rapidly. The cause of this rapid rise of the trough was confirmed by a visual observation of water flow in the fossa: the water reserved in the right cavity by surface tension suddenly flows into the left cavity through the postarytenoid space that interconnects the two cavities. Comparing the frequencies of the peaks before and after the water injection, it can be seen that the first peak increases about 9% when the piriform fossa was completely filled with water.

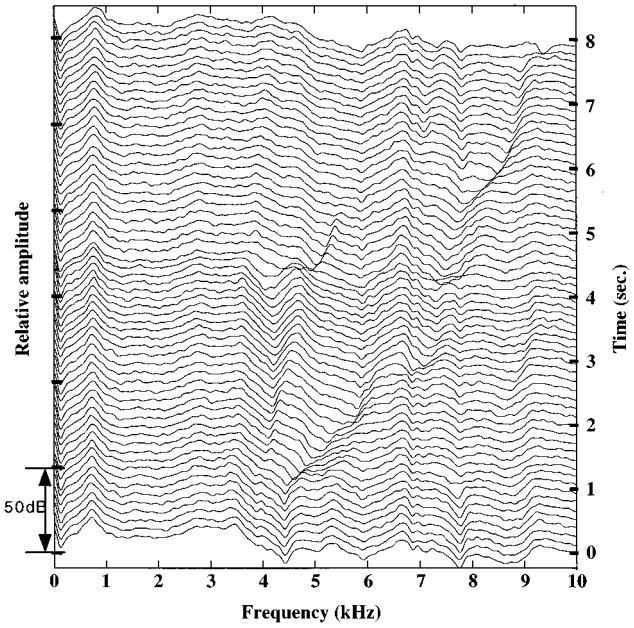


FIG. 8. Running spectra of the radiated sound from M model /a/ excited by constant sound source. Motion of the troughs is shown while water is injected into the piriform fossa.

III. ANTIRESONANCES OF THE PIRIFORM FOSSA IN SPEECH SPECTRA

The above results indicate that water injection can be used to examine the acoustic effects of the piriform fossa in humans. Water has a density close to that of muscles, and is an adequate material to fill the piriform fossa of humans if the injection is conducted carefully. In this section, the antiresonances of the piriform fossa are examined by a water injection experiment and using natural vowel utterances.

A. Water injection experiment on human subjects

The same procedure of water injection was applied to two male subjects JD and KH. A flexible tube with a 0.3-cm o. d. and a 0.2-cm i.d. was inserted along the nasal floor and passed through the nasopharyngeal port toward the piriform fossa. The tip of the flexible tube was placed above the right cavity under fiberoptic video-endoscopy. Warm water, at about body temperature, was injected into the right cavity through the flexible tube. When the right side was filled, the water was seen to flow into the left cavity. Radiated oral sounds of sustained vowels /a/, /i/, and /u/ were recorded during the water injection.

Figure 9 shows running spectra of the vowel /a/ for JD. The motion pattern of the troughs shown in this figure is similar to that seen in the M models. In the figure, a trough appears at about 5200 Hz before the water was injected. After about 0.2 s, the trough caused by the right cavity rises

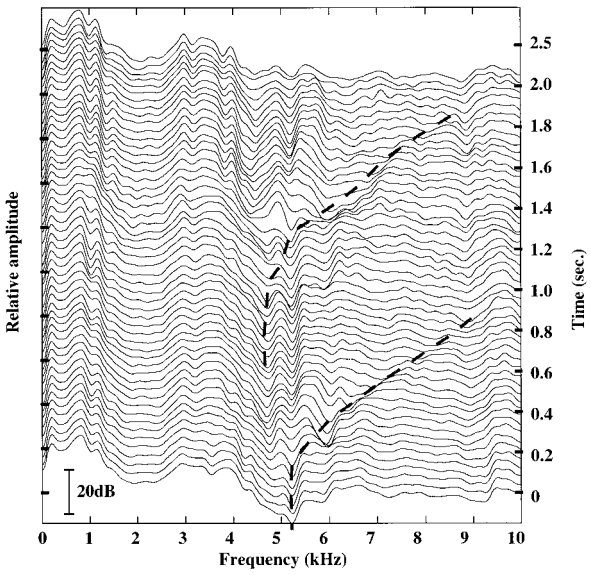


FIG. 9. Running spectra for sustained vowel /a/ produced by subject JD while water is injected into the piriform fossa.

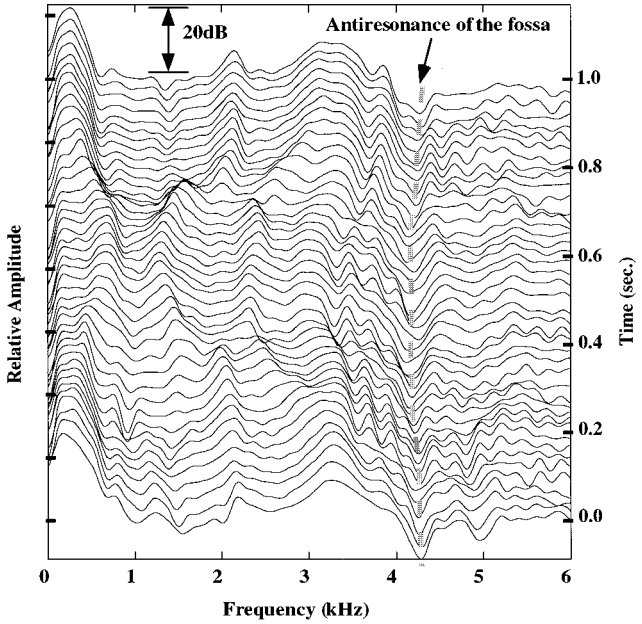


FIG. 10. Running spectra of /iai/ sequence produced by subject KH in the supine position.

as the cavity is filled with water. This trough disappears from the spectra when the right cavity is full, which is observed at about 1 s. The trough for the left cavity begins to rise as the water flows into the left side, which occurs approximately after 1 s, and disappears when both the right and left cavities become full after about 2 s. The running spectra indicate that the piriform fossa causes two antiresonances around 5 kHz in speech spectra for this subject, although the two antiresonances usually appear as one trough in speech spectra. The same trough pattern in the antiresonances was also confirmed in the speech spectra for KH.

Although both the experiments on the M models and on humans demonstrated a consistent trough pattern, the expected changes in the lower formants were not obvious in humans. In particular, $F1$ remained almost constant during water injection. There seem to be at least two possible causes for the stability of $F1$. First, a physiological reflex to the injected water could result in a small change in articulatory posture for the vowels. Second, the subjects might adjust articulator positions to compensate for the acoustic effect of decreasing air volume by means of auditory feedback. An additional experiment was performed to verify the latter possibility (i.e., auditory feedback) by applying a loud masking noise during water injection. The result did not show the expected changes in the first formant, but rather supported the former explanation (i.e., physiological reflex to water injection).

B. Antiresonance of the piriform fossa in natural speech

It was shown in the above experiments that the piriform fossa causes a trough in the 4 to 5 kHz region. A question is raised as to whether such a trough can be found in natural speech. To answer this question, natural vowel utterances were analyzed for the four subjects who served for the MRI experiment. Speech materials consisted of isolated vowels /a/, /i/, /u/, and vowel sequences (/aia/, /aua/, /iai/, /iui/, /uau/, and /uiu/). Sound recording was carried out in an anechoic room at a sampling rate of 44.1 kHz, where the subjects repeated the speech materials twice at a natural speech rate in two body postures: upright and supine. Figure 10 shows running spectra of /iai/ recorded in the supine position for KH. While the formants in this figure show articulatory changes from /i/ to /a/ to /i/, a trough at about 4300 Hz remained almost constant during the vowel sequence. Referring to the results from human and model experiments, it is reasonable to judge that the trough is the antiresonance caused by the piriform fossa.

Speech samples recorded from the four subjects in the supine position were analyzed for the stable segments in the three Japanese vowels. Figure 11 shows the means and standard deviation of the antiresonance frequencies for the four subjects. The coefficient of variation (CV, the ratio of the

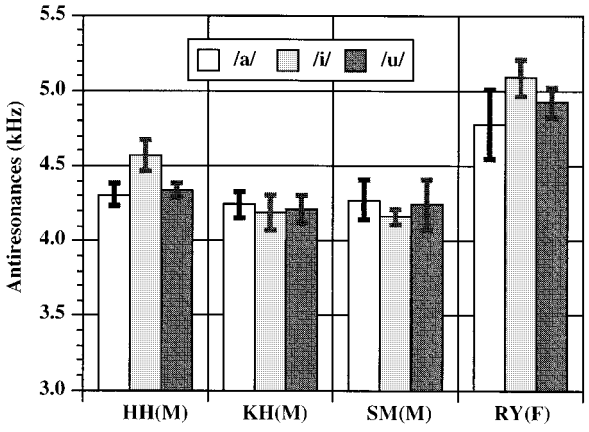


FIG. 11. Means and standard deviations of antiresonance frequencies caused by the piriform fossa measured for each vowel from vowel sequence data of four subjects.

standard deviation to the means) for each vowel was about 5% for all subjects, except that of /a/ for subject RY, which was about 10%. The CVs among the male subjects were within 5%, whereas it was about 15% between the female and the males. The female subject showed higher antiresonance frequencies than the males. The results indicate that antiresonances caused by the piriform fossa appear constantly in the spectra of the natural vowel sequences.

IV. SIMULATIONS USING NUMERICAL MODELS

The MRI data and M-model experiments in this study provide a clear view of the geometric and acoustic characteristics of the piriform fossa. These findings were adopted to design a numerical model (N model) of the vocal tract based on a transmission line model. The performance of the N model was tested in a comparison of simulation results and real speech spectra. The local and global acoustic effects of the piriform fossa were examined by model simulation and discussed on the basis of acoustic theories.

A. Comparison of the antiresonances from the N model and natural speech

The numerical model is a transmission line model designed according to the morphological data of our subjects. The basic design of the model is a pair of side branches attached to the main tract at 2 cm above the glottis. The cone and cylinder model of the piriform fossa as shown in Fig. 4 is represented by area functions with section lengths of 0.1 cm. The open end correction coefficient of the fossa was 0.75, as discussed in Sec. II. The radiation impedance of the vocal tract was approximated by a cascade concatenation of radiation resistance and radiation inductance (Caussé *et al.*, 1984), which is valid for a wide frequency region of $kr<1.5$, where k is the wave number and r is the radius of the radiating end.

The accuracy of the N model was tested by a comparison of the computed and measured transfer functions of the M models. The N model computations were confirmed to be consistent with the measured M model spectra within 2% for both resonances and antiresonances below 6 kHz. Then, the antiresonance of the piriform fossa was calculated using N models of the entire vocal tract for the four subjects. They are shown in Fig. 12, along with those obtained from natural speech for each subject. The antiresonance frequencies predicted by the N models were consistent with those from the speech data within 5%, except for a vowel /i/ of the subject HH. The result shows that the model gives a realistic description of the antiresonances of the piriform fossa. Reviewing the results above, the antiresonance frequency (F) of the piriform fossa and the depth (D) of the cone portion approximately satisfy the relation $F=c/4D$ after taking into account the cylinder portion, where c denotes the sound velocity.

B. The global and local effects of the piriform fossa

The global effect of the piriform fossa on the lower formants has been demonstrated by earlier studies (Fant, 1960; Baer *et al.*, 1991; Davies *et al.*, 1993; Fant and Båvegård, 1995). The purpose of the numerical simulation in this section is to examine the causality of the global effect of the piriform fossa. The result of simulation is discussed in terms of the plausibility of two possible treatments of the piriform fossa in vocal tract modeling: as side branches or as an additional volume.

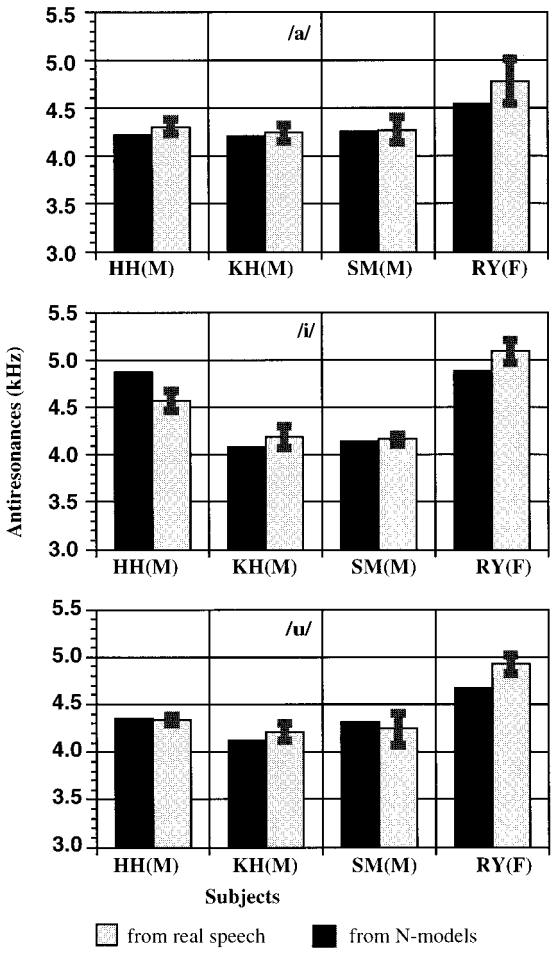


FIG. 12. Comparison of the antiresonance frequencies derived from N models and real speech for four subjects.

N model simulations were carried out for the vowels /a/ and /i/ under three conditions in which the piriform fossa was treated as a side branch (SB), as an additional volume (AV), and with no fossa (NF) as a control condition. The vocal tract configuration of the N model was based on volumetric MRI data of KH. Real speech spectra from the subject were obtained by homomorphic analysis (Markel and Gray, 1976). Figure 13 shows the speech spectra and computed velocity-to-velocity transfer functions from the glottis to the lips of the N models. In the case of /a/, $F1$ decreases by 9% in the AV treatment, and by 10% in SB treatment in comparison with the NF condition. In contrast, the effects on $F2$ and $F3$ of /a/ are relatively small. The differences in $F2$ and $F3$ were 2% and 3% in SB, while they were not seen in AV. In the case of /i/, the first two formants were lowered 3% in both AV and SB for $F1$, and 2% in AV and 6% in SB for $F2$, respectively. In comparison with the condition with-

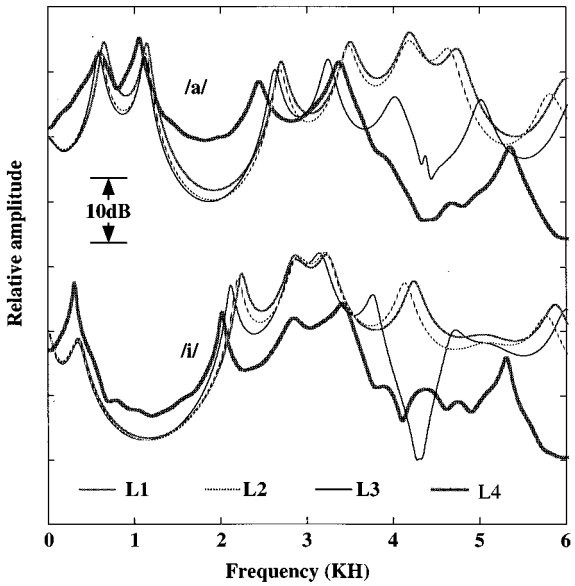


FIG. 13. Velocity-to-velocity transfer function from the glottis to the lips computed using N models, whose area functions were obtained from KH for vowels /a/ and /i/ under conditions: without the piriform fossa (L1), treating the fossa as an additional volume (L2) and as a side branch (L3). L4 shows speech spectra from the same subject.

out the piriform fossa, both AV and SB treatments demonstrated a global effect on the lower formants. The extent of the changes in the lower formants is higher than the different limens shown by Kewley-Port and Watson (1994) and Hawks (1994), who pointed out that formant-frequency discrimination is 14 Hz for $F1$ ($<$800 Hz), and 1.5% for $F2$. Therefore, it can be expected that the global effect causes some perceptual difference. This expected effect was confirmed in our informal listening test.

A few theoretical discussions are found in the literature regarding the acoustic effects of local volume change in the vocal tract. According to the perturbation theory (Schroeder, 1967), the effect of an additional volume of the vocal tract on formant frequencies is approximately proportional to the volume increment and the square of the sound pressure at the location of the additional volume. The same phenomenon can also be explained by the transmission line theory. The effect of a side branch on vowel formants is proportional to the admittance of the branch and the square of the sound pressure at the branch location in the vocal tract. Both theories agree in their view that the manner of formant frequency change depends on the sound pressure at the location where a small volume change takes place. At the closed end of the vocal tract sound pressure is near a maximum for all formant frequencies. Since the piriform fossa is located near the closed end of the vocal tract, it plays an important role for almost all formants. In the region much lower than the antiresonance frequency of the piriform fossa, the distribution of sound pressure in the vocal tract is expected to be the same in AV and SB treatments. Accordingly, the global effect of the fossa on the lower formants can be represented equally well regardless of whether the fossa is treated as a side branch or as an additional volume. However, the AV treatment demonstrates an incorrect behavior in the frequency region close to the antiresonance frequency of a branch because the perturbation theory cannot predict the pressure disturbance caused by the antiresonance of the branch.

Summarizing our experimental data and the above theoretical accounts, SB treatment demonstrates much more realistic performance than AV treatment in the region above 2 kHz, though they both can represent real speech spectra relatively well in the region below 2 kHz. Taken together the local and global features of the acoustic effects, it is reasonable to conclude that the piriform fossa should be modeled as side branches of the vocal tract.

V. CONCLUSIONS

This study examined the acoustic effects of the piriform fossa on the vocal tract transfer function by conducting both human experiments and model simulations. Our results indicate that the piriform fossa forms an important part of the vocal tract in realizing natural speech spectra. Furthermore, the results suggest that the function of the piriform fossa must be incorporated in any realistic model of speech production.

The morphology of the piriform fossa was investigated using volumetric MRI images for four subjects. Mechanical models were constructed from the volumetric data of one of the subjects and were used for acoustic investigation. The effective cavity of the modeled piriform fossa consists of two cascaded portions: the cone portion from the bottom of the fossa to the arytenoid apex plane and the cylinder portion above the plane. The ratio of the cylinder portion to the cone portion ranged from 0.4 to 0.8 for the four subjects.

The acoustic characteristics of the piriform fossa were investigated on the mechanical models by manipulating the air volume of the fossa. The results showed that the fossa causes troughs in the transfer function of the vocal tract in the frequency region between 4 to 5 kHz. The fossa not only affects spectral shape in the vicinity of its antiresonance, but also decreases resonance frequencies in the lower frequency region. Observations of an *in-vivo* experiment on human subjects and in natural vowel sequences showed that the antiresonance of the fossa constantly appears in natural speech utterances as well as in sustained vowels.

Numerical models were designed to simulate the acoustic behaviors of the piriform fossa. Computed antiresonances of the piriform fossa were consistent with those obtained from natural speech, generally within 5% for the four subjects. The numerical model demonstrated that the piriform fossa causes global effects on the lower formants because it is located near the closed end of the vocal tract. Considering both global and local effects, the piriform fossa should be treated as a side branch attached near the glottal end of the vocal tract rather than as an additional volume of the pharynx.

ACKNOWLEDGMENTS

The authors would like to express their appreciation to Christine Shadle for her helpful discussions. The authors also would like to thank Gunner Fant and Osamu Fujimura for

their instructive comments. We also thank Naoki Kusakawa and Hiroyuki Hirai for their assistance in the experiments.

Baer, T., Gore, J., Gracco, L. C., and Nye, P. W. (**1991**). ''Analysis of vocal tract shape and dimensions using magnetic resonance imaging: vowels,'' J. Acoust. Soc. Am. **90,** 799–828.

Caussé, R., Kergomard, J., and Lurton, X. (**1984**). ''Input impedance of brass musical instruments—comparison between experiment and numerical model,'' J. Acoust. Soc. Am. **75,** 241–254.

Dang, J., Honda, K., and Suzuki, H. (**1994**). ''Morphological and acoustical analysis of the nasal and the paranasal cavities,'' J. Acoust. Soc. Am. **96,** 2088–2100.

Dang, J., and Honda, K. (**1996a**). ''A new method for measuring antiresonances of the vocal tract transmission characteristics: an experiment study of acoustic tubes,'' J. Acoust. Soc. Jpn. (E) **13**, 93–99.

Dang, J., and Honda, K. (**1996b**). ''Acoustical modeling of the vocal tract based on morphological reality: Incorporation of the paranasal sinuses and the piriform fossa,'' Proc. of 4th Speech Production Seminar, 49–52, Grenoble.

Davies, L., McGowan, R., and Shadle, C. (**1993**). ''Practical flow duct acoustics applied to the vocal tract,'' in *Vocal Fold Physiology: Frontiers in Basic Science*, edited by I. Titze, pp. 93–142.

Fant, G. (**1960**). *Acoustic Theory of Speech Production* (Mouton, The Hague), 2nd ed., 1970.

Fant, G., and Båvegård, M. (**1995**). ''Parametric model of VT area functions: Vowels and consonants,'' EU report, Speechmaps (Esprit/BR 6975) Deliv. 28 WP2.2.

Flach, M., and Schwickardie, H. (**1966**). ''Die Recessus Piriformes unter phoniatrischer Sicht,'' Fol. Phoniat. **18**, 153–167.

Hawks, F. W. (**1994**). ''Difference limens for vowel formant frequency,'' J. Acoust. Soc. Am. **95,** 1074–1084.

Imai, S., and Abe, Y. (**1979**). ''Spectral envelope extraction by improved cepstrum,'' IEICE, J62-A, 4, 217–228 (in Japanese).

Kewley-Port, D., and Watson, C. S. (**1994**). ''Threshold for formant-frequency discrimination for isolated English vowels,'' J. Acoust. Soc. Am. **95,** 485–496.

Lin, Q. (**1990**). ''Speech production theory and articulatory speech synthesis,'' Ph.D. thesis of KTH.

Markel, J. D., and Gray, A. H. (**1976**). *Linear Prediction of Speech* (Springer-Verlag, New York).

Mermelstein, P. (**1967**). ''On the piriform recesses and their acoustic effects,'' Fol. Phoniat. **18**, 153–167.

Moore, C. A. (**1992**). ''The correspondence of vocal tract resonance with volumes obtained from magnetic resonance images,'' J. Speech Hear. Res. **35**, 1009–1023.

Narayanan, S., Alwan, A., and Haker, K. (**1995**). ''An articulatory study of fricative consonants using magnetic resonance images,'' J. Acoust. Soc. Am. **98,** 1325–1347.

Schroeder, M. R. (**1967**). ''Determination of the geometry of the human vocal tract by acoustic measurements,'' J. Acoust. Soc. Am. **41,** 1002–1020.

Sundberg, J. (**1974**). ''Articulatory interpretation of the singing formants,'' J. Acoust. Soc. Am. **55,** 838–844.

Yang, C., Kasuya, H., Kanou, S., and Satou, S. (**1994**). ''An accurate method to measure the shape and length of the vocal tract for the five Japanese vowels by MRI,'' Jpn. J. Logopedics Phoniatrics **35** 317–321 (in Japanese).

Glottal characteristics of female speakers: Acoustic correlates

Helen M. Hanson[a)]
Research Laboratory of Electronics, Massachusetts Institute of Technology, Room 36-549, 50 Vassar Street, Cambridge, Massachusetts 02139

(Received 17 January 1996; accepted for publication 10 September 1996)

The aim of the research reported in this paper is to formulate a set of acoustic parameters of the voicing source that reflect individual differences in the voice qualities of female speakers. Theoretical analysis and observations of experimental data suggest that a more open glottal configuration results in a glottal volume-velocity waveform with relatively greater low-frequency and weaker high-frequency components, compared to a waveform produced with a more adducted glottal configuration. The more open glottal configuration also leads to a greater source of aspiration noise and larger bandwidths of the natural frequencies of the vocal tract, particularly the first formant. These different attributes of the glottal waveform can be measured directly from the speech spectrum or waveform. A set of acoustic parameters that are likely to indicate glottal characteristics is described. These parameters are measured in the speech of a group of female speakers, and the glottal configurations of the speakers are hypothesized. This research contributes to the description of normal variations of voicing characteristics across speakers and to a continuing effort to improve the analysis and synthesis of female speech. It may also have applications in clinical settings. © *1997 Acoustical Society of America.* [S0001-4966(97)03001-4]

PACS numbers: 43.70.Gr, 43.70.Aj, 43.72.Ar [AL]

INTRODUCTION

Certain acoustic characteristics of speech give a voice its quality and individuality, and are the means by which listeners identify or distinguish speakers. These characteristics are complex, having contributions that range from those of the speech production mechanism, that is, differences between individual sound sources and the natural frequencies of the vocal tract, to the effects of prosody and dialect. The work reported in this paper focuses on individual variations in the glottal source waveform, particularly those associated with degree of glottal adduction. These variations are of interest for speech applications such as speaker verification and identification, speech recognition, and speech synthesis by computer, and also have applications for speech disorders and therapy, and in the field of forensics.

The values of the parameters that describe the glottal waveform can vary depending on the glottal configuration, and it is expected that these variations may lead to different voice qualities and intensities. Some voice qualities are usually associated with disordered voice, such as harshness, but our main concern for this paper are those that occur for voices that are not considered to be disordered. Voice qualities that occur frequently in normal speech are pressed, modal, and breathy, described in Laver (1980) and Titze (1995). There are other voice qualities that may occur in normal conversational speech (Laver, 1980), but the focus of this paper is restricted to a few different glottal vibratory patterns (to be described in Sec. I) that may be associated with pressed, modal, and breathy voice qualities. These patterns may be considered regions along one or more continua; we consider degree of glottal adduction, which is greatest for pressed voice, least for breathy voice, and somewhere in between for modal voice.

Direct examination of the vocal folds during normal phonation has revealed that female speakers are more likely than male speakers to have incomplete closure of the vocal folds (Hertegård *et al.*, 1992; Linville, 1992; Peppard *et al.*, 1988; Södersten and Lindestad, 1990; Södersten *et al.*, 1991). The degree of incomplete closure can vary, and posterior openings of the glottis may extend beyond the vocal processes to the membranous part of the folds. The implications of incomplete closure for the glottal waveform are that there is an airflow bypass even during the so-called closed phase of the glottal vibratory cycle, and that an abrupt discontinuity of the airflow derivative is not possible due to the mass of air in this pathway. As we will see later in this paper, the effects of this bypass on the glottal waveform parameters increase with its size, providing one source of variability in voicing characteristics.

The aim of our work has been to formulate a set of acoustic parameters that reflect individual and gender differences in voice quality and glottal configuration. These parameters are derived from measurements of the acoustic spectra of vowels produced by speakers. In this paper we provide the theoretical background for these parameters, and present acoustic data for female speakers. Physiological and perceptual data that support the acoustic data, data for male speakers, and comparison of data across gender will be described in future papers. The decision to initially study the male and female data separately was influenced by the gender differences in glottal configuration described above.

A. Related work

Previous studies of individual variations in glottal characteristics have been based on inverse filtering, visual in-

[a)]Electronic mail: hanson@speech.mit.edu

spections of the vocal folds, acoustic measures made directly on the speech spectrum or waveform, and perceptions of voice quality, particularly breathiness.

Holmberg and her colleagues have made aerodynamic measures of male and female voice (Holmberg *et al.*, 1988, 1989, 1994a, 1994b, 1995; Perkell *et al.*, 1994), and extracted glottal parameters directly from the glottal waveform. These studies included relatively large groups of subjects (from 15 to 45) phonating in different speech conditions. The main findings are the existence of parameter differences between females and males, and across speech conditions. Their studies have also included acoustic measures made on the speech spectrum, and they related these spectral measures to the aerodynamic results (Holmberg *et al.*, 1995), finding that adduction quotient, as measured from the glottal waveform, has a strong relationship with the relative amplitudes of the first two harmonics. In addition, the speed of glottal closure is in some cases reflected by the relative amplitudes of the first and third formants.

Karlsson (1986, 1988, 1989, 1990, 1991a, 1991b, 1992a, 1992b) also studied the glottal waveform obtained by inverse filtering, focusing on a small group of female speakers. Glottal parameters were derived by fitting a theoretical model to the resulting glottal waveform, and varied widely across speakers (Karlsson, 1988). She also attempted to correlate the glottal parameters with the voice qualities of her subjects, as judged by speech therapists. Different voice qualities were separated by the degree of spectral tilt of the voice source, and the presence of noise excitation at mid to high frequencies (Karlsson, 1988, 1992a). Speakers perceived to be breathy had higher minimum flows, steeper tilts, and more aspiration noise (Karlsson, 1988, 1992b). Voices with a tight, strained quality had weaker lower harmonics (Karlsson, 1992b).

Södersten and her colleagues observed glottal closure via fiberscopy and related the degree of closure to perceptions of breathiness. They also used an acoustic measure, the amplitude of the first harmonic relative to the amplitude of the first formant peak. Their results show that female speakers have a higher degree of incomplete closure and perceived breathiness than male speakers (Södersten and Lindestad, 1990), and that significant correlations exist between perceived breathiness and the relative amplitude of the first harmonic (Södersten *et al.*, 1991).

Gobl (1989) examined voice quality correlates by extracting glottal parameters from the inverse-filtered waveform and making measurements on the glottal spectrum. The acoustic measures were the average of the harmonic amplitudes in four frequency bands. Breathy voice was found to have a steeper spectral tilt than modal voice. Gobl and Ní Chasaide (1988) and Ní Chasaide and Gobl (1993) extracted glottal parameters both from the glottal waveform and from vowel spectra. They found that as glottal abduction increased, so did the downward spectral slope of the vowel spectrum, and that formant amplitudes, especially that of $F1$, decreased as well.

Klatt and Klatt (1990) studied variations in voice quality, from pressed to breathy, for both male and female speakers, using reiterant and synthesized speech. Through perception tests, they too found that female speakers were perceived to be breathier than male speakers. Unlike the work discussed above, they relied entirely on acoustic measures made directly on the speech spectrum and waveform. The relative amplitude of the first harmonic was higher for female speakers, who also showed more evidence of aspiration noise at high frequencies. Relevant cues to perception of breathy voice were increases in the amplitude of the first harmonic, aspiration noise, and lower formant bandwidths, with aspiration noise as the most important cue.

There have been other attempts to derive glottal parameters based on measurements of the spectrum of the glottal waveform, or of the speech waveform and spectrum. In particular, Fant (Fant, 1979, 1993; Fant *et al.*, 1985, 1994; Fant and Lin, 1988). Ananthapadmanabha (1984), and Gauffin and Sundberg (1989) have pioneered in developing these techniques.

B. Summary

Several researchers have studied variations in glottal vibratory patterns using inverse filtering or fiberscopic and endoscopic examination of the vocal folds. However, inverse filtering has several problems associated with it (Holmberg *et al.*, 1995), and examination of the vocal folds is necessarily invasive. Thus, there is a need to develop methods of measuring glottal characteristics directly from the acoustic sound pressure which do not require special equipment to record data. In the work to be reported here, we also estimate the spectrum of the glottal source, but rather than filtering out the effects of the vocal tract filter, we take them into account in our analysis. Besides avoiding the problems of inverse filtering, we are able to examine the effect of the glottal source on the filter, especially its bandwidths, thus gaining further information about the glottal configuration.

The contributions of this work are several. First, it adds to research efforts aimed at finding quantitative measures that describe dimensions along which normal voices vary across speakers. It also contributes to the analysis, synthesis, and modeling of female voice and speech. In addition, the work may have clinical applications.

We begin the next section by reviewing the theoretical basis for measurements made on the speech spectrum and waveform. As a result of this theoretical development, several measures of glottal characteristics are suggested. Acoustic data for 22 female speakers are then given, and we attempt to interpret these data in terms of the theoretical models and to classify individual differences based entirely on the inferences derived from the measurements of sound pressure.

I. THEORETICAL BACKGROUND

In this section we discuss several ways in which the configuration of the vocal folds and glottis may vary during vowel production. Specifically, we consider four types of configurations: (1) the arytenoids are approximated and the membranous part of the folds close abruptly; (2) the arytenoids are approximated, but the membranous folds close nonsimultaneously along the length of the folds; (3)

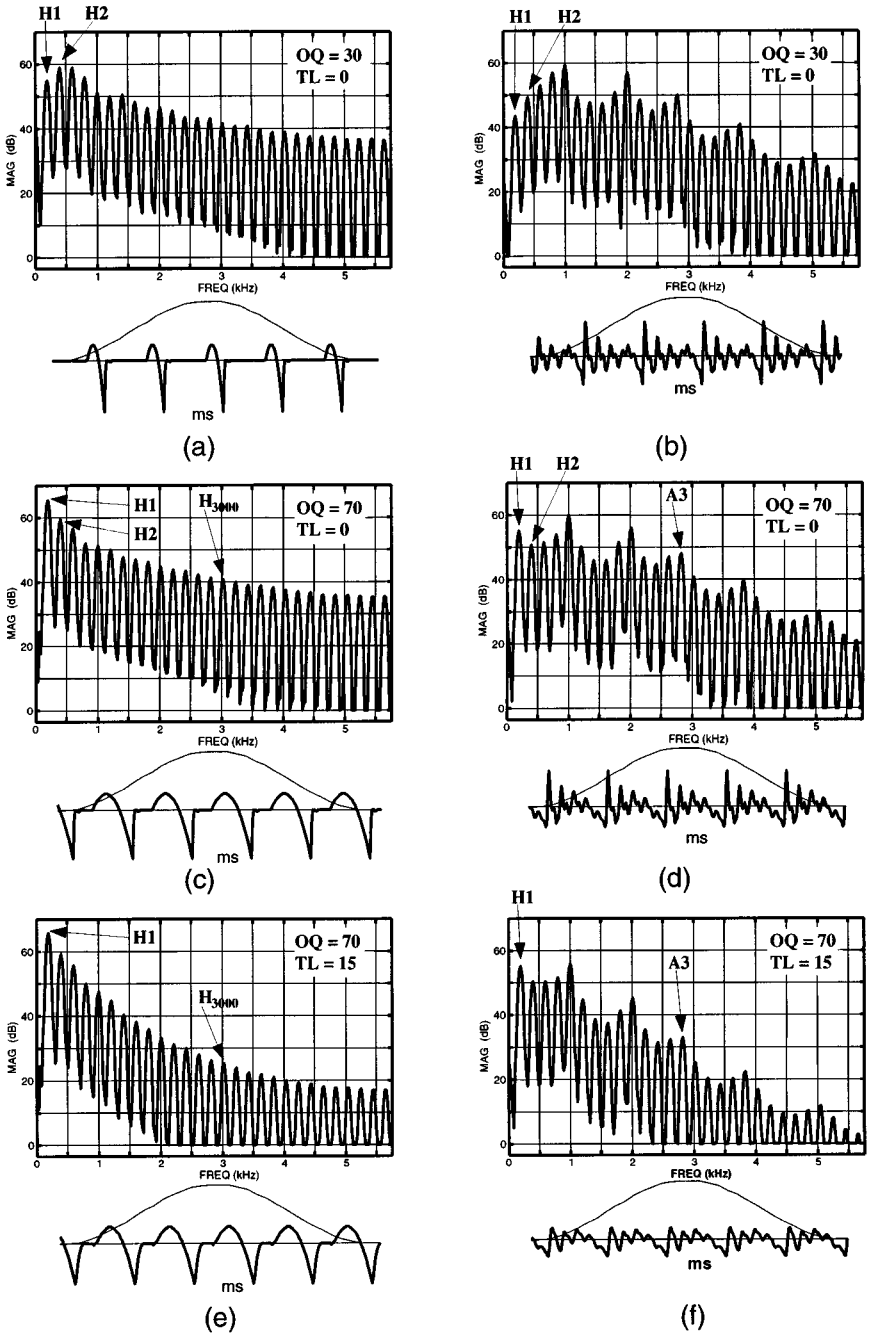


FIG. 1. Waveforms and spectra of the periodic glottal volume-velocity source corresponding to various manipulations of the glottis. The fundamental frequency is in the range for an adult female speaker. Panels (a), (c), and (e) show spectra and derivatives of the volume-velocity sources, while panels (b), (d), and (f) show the spectra of the vowel /æ/ synthesized using those volume-velocity sources. (a)–(b): Open quotient (OQ) is 30%, additional spectral tilt (TL) is zero; (c)–(d): OQ is 70%, TL is zero; (e)–(f): OQ is 70%, TL is 15 dB (i.e., spectrum is 15 dB lower at 3 kHz). The waveforms were generated with the KLSYN88 synthesizer (Klatt and Klatt, 1990).

there is a posterior glottal opening, or ''chink,'' at the arytenoids that persists throughout the glottal cycle, but the folds close abruptly; (4) a posterior opening extends into the membranous portion of the folds throughout the glottal cycle, forcing the folds to close nonsimultaneously. Through a combination of observation and modeling, we suggest several ways in which these various configurations affect the glottal airflow and are manifested in the speech spectrum or waveform. Note that there may be other glottal configurations in addition to the four that we consider.

A. Complete glottal closure during a vibratory cycle

For the case in which the glottis closes completely during a part of the glottal cycle, the glottal waveform can show several kinds of differences from one individual to another. For example, if a speaker modifies her production such that it results in a glottal waveform with a larger open quotient but the same rate of decrease of volume velocity at closure, the spectrum of the source undergoes a change only at low frequencies, with essentially no change in the spectrum ampli-

tude at high frequencies (Klatt and Klatt, 1990). Figure 1(a) and (c) shows the derivatives of the volume-velocity waveform[1] and the spectra of these derivatives for two synthesized waveforms[2] having different values of open quotient (OQ) (30% and 70%, respectively). When open quotient varies from 30% to 70%, the amplitude of the first harmonic relative to that of the second ($H1-H2$) changes by about 10 dB. Figure 1(b) and (d) shows spectra for the vowel /æ/ synthesized using these glottal waveforms. The difference between the values of $H1-H2$ that were observed in the glottal spectra is also evident in the spectra of the synthesized vowels. Other researchers have used $H1-H2$ as an indication of open or adduction quotient (see, for example, Holmberg *et al.*, 1995).

The spectrum of the derivative of the glottal waveform at middle and high frequencies, when the derivative has a discontinuity at the time of closing, has a downward slope of 6 dB/octave. This spectrum is influenced by the abruptness with which the flow is cut off when the membranous part of the vocal folds closes during the vibration cycle. This abruptness can be affected in two ways, for a given open quotient, when there is complete closure of the glottis during some part of the vibratory cycle (Fant *et al.*, 1985; Klatt and Klatt, 1990). One mechanism that leads to a change in abruptness is a glottal closing that does not occur simultaneously at all points along the anterior–posterior length of the vocal folds. Closing is a type of ''zipper'' action, with initial closure at the anterior end of the glottis and the closure sliding back along the length of the glottis (cf. Ananthapadmanabha, 1993). This type of closure leads to a more gradual cutoff of flow, resulting in a derivative of the glottal waveform that does not have a discontinuity. As illustrated with synthesized glottal waveforms in Fig. 1(e), the effect on the spectrum is to introduce an additional downward tilt, or slope, at high frequencies. If we define T_D as the time from initiation of the anterior closure to the time of closure at the posterior end, and if we approximate the gradual cutoff as an exponential, then a reasonable approximation to the time constant T of this exponential is roughly one-half of the time of the sliding closure,[3] that is,

$$T \approx \frac{T_D}{2}. \tag{1}$$

The breakpoint for the change in spectral slope is then given by

$$f_T = \frac{1}{2\pi T} = \frac{1}{\pi T_D}. \tag{2}$$

Above this frequency, the slope of the spectrum of the derivative of the waveform increases to 12 dB/octave if an exponential approximation is assumed (see, for example, Siebert, 1986). For f_T less than about 2000 Hz, the resulting increase in the tilt at 2750 Hz, an average location of $F3$ for female speakers, is

$$20 \log_{10} \frac{2750}{f_T}. \tag{3}$$

For example, if T_D is 0.5 ms, f_T is 637 Hz, and the increase in tilt at 2750 Hz is 13 dB. Likewise, if T_D is 1.0 ms, the increase in tilt at 2750 Hz is 19 dB. The glottal waveform of Fig. 1(e) is synthesized such that the amplitude of the source spectrum at 3000 Hz (H_{3000}) is 15 dB lower than the spectra in Fig. 1(a) and (c). The spectrum of a vowel synthesized using the glottal waveform of Fig. 1(e) is shown in Fig. 1(f). Note that the amplitude of the third formant ($A3$) drops by about 15 dB compared with the spectrum in Fig. 1(d). Thus, the amplitude of the third formant relative to that of the first harmonic ($H1-A3$) appears to be a reasonably accurate indication of source spectral tilt, except if $H1$ is weak, as in Fig. 1(b). Holmberg *et al.* (1995) have also used this measurement as an indication of how abruptly airflow is cut off.

Another way in which the abruptness at closure can be influenced is by manipulating the rate of decrease of flow at the instant of closure. For a given open quotient, this rate of decrease is related to the amount of skewness of the glottal pulse, as shown with synthesized glottal waveforms in Fig. 2(a) and (b). As the slope of the closing phase becomes faster relative to the slope of the opening phase, the spectrum amplitude at middle and high frequencies increases relative to the amplitude at low frequencies. In this example, the difference between the amplitudes of the first harmonic ($H1$) and the harmonic at 3000 Hz (H_{3000}) increases by about 10 dB with a change in speed quotient from 140% to 320%. Speech spectra corresponding to these glottal waveforms are shown in Fig. 2(c) and (d), and again, the difference $H1-A3$ for the two spectra provides an indication of tilt.

The amplitude of the third formant is also influenced by other factors, one being the locations of $F1$ and $F2$. Another is that the bandwidth of $F3$ is affected by the radiation characteristic to a greater extent than are the lower formants, and the degree of this influence varies with the configuration of the vocal tract for the vowel. Corrections must be made for these effects if values of $A3$ are to be compared across vowels or speakers. Finally, the value of $A3$ will vary depending on whether or not $F3$ is centered on a harmonic. A minimum value of $H1-A3$ that can be expected for a neutral vowel with equally spaced formants when there is no additional spectral tilt due to nonabrupt airflow cutoff has been estimated to be about 9 dB for female speakers (Hanson, 1995), based on calculations made with the KLGLOTT88 source model (Klatt and Klatt, 1990).

B. Incomplete glottal closure during a vibration cycle

Glottal configurations exhibiting an airflow bypass can modify the spectrum of the glottal waveform and the transfer function of the vocal tract relative to those that would exist for the configuration in which the entire glottis is closed over part of the cycle of vibration. In this section we show that among the modifications introduced are: (1) an increase in the bandwidth of the first (and possibly the second) formant; (2) an increase in the tilt of the glottal spectrum at high frequencies; and (3) emergence of a turbulence noise source in the vicinity of the glottis that may be comparable in amplitude (at high frequencies) to the spectrum amplitude of the periodic source.

Cranen and Schroeter (1995) have also studied the acoustic consequences of glottal openings. Although their model exhibits an increase in source spectral tilt only when

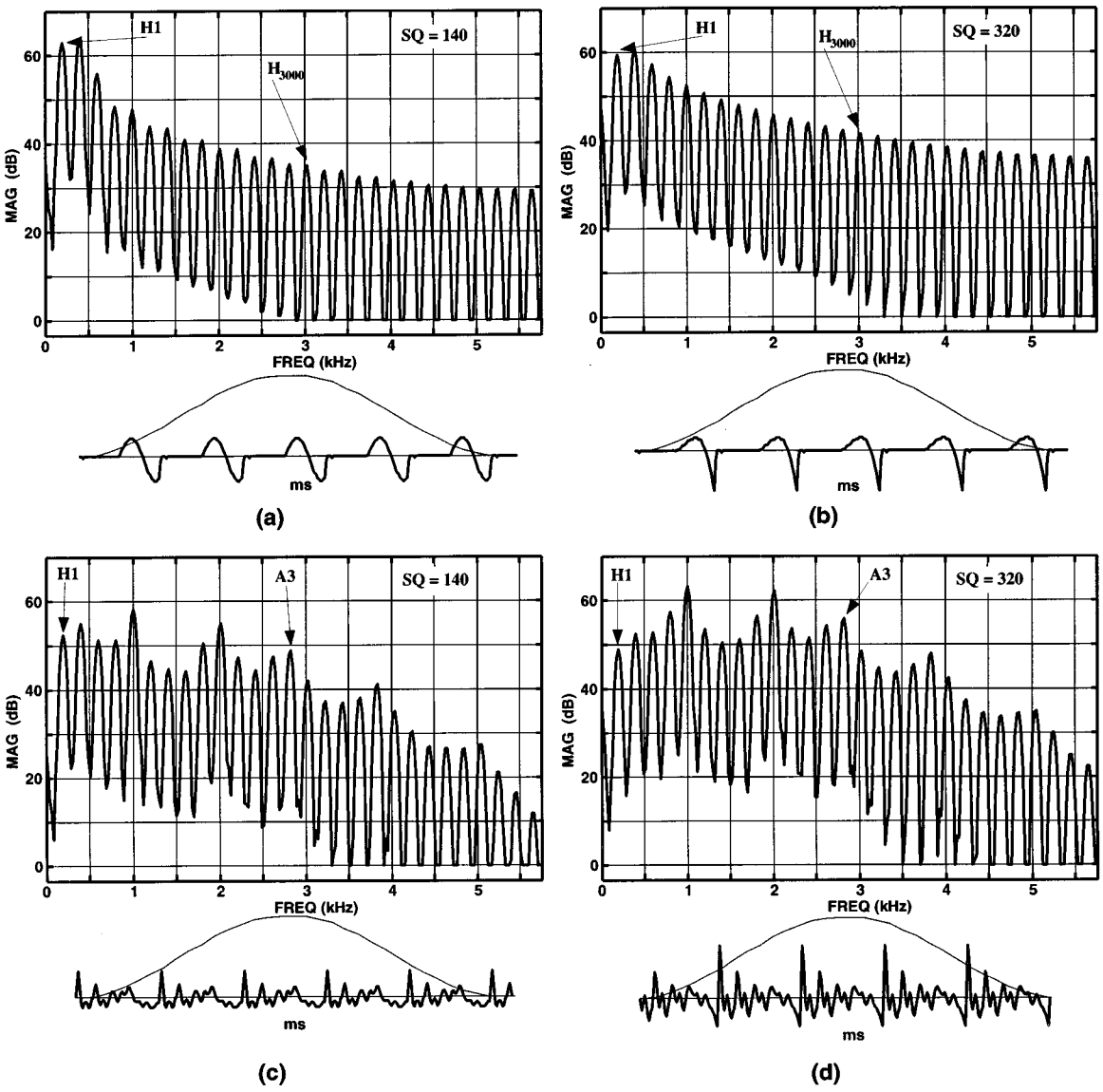


FIG. 2. Waveforms and corresponding spectra of the derivative of the periodic glottal volume-velocity source with different speed quotients (SQ), the speed of the closing phase relative to that of the opening phase. (a) Skewing of waveform decreased to give a SQ of 140%; (b) SQ is 320%; (c) The vowel /æ/ synthesized with the glottal waveform of (a); (d) The vowel /æ/ synthesized with the glottal waveform of (b).

the vocal processes are abducted, we show that all of the above-mentioned acoustic manifestations occur whether the bypass is a glottal chink or an abduction extending beyond the vocal processes. However, the second of these properties, an increase in tilt, is expected to be more marked when there is also abduction of the arytenoids at the vocal processes.

1. Effect on first-formant bandwidth

Formant bandwidths are related to the rate of acoustic energy loss in the vocal tract. The energy losses in the frequency range of the first formant come from several sources, including the resistance of the yielding walls of the vocal tract, and heat conduction and frictional losses at the walls. In earlier work (House and Stevens, 1958; Fant, 1962; Fujimura and Lindqvist, 1971), bandwidths due to vocal-tract losses were measured by exciting a subject's vocal tract while the subject held his or her glottis closed. A first-formant bandwidth of 40–95 Hz was found for female speakers in the closed glottis condition (Fujimura and Lindqvist, 1971; Fant, 1972). When the glottis is open and there is airflow through it, the glottal resistance can contribute further energy loss, particularly at low frequencies, thus adding significantly to the first-formant bandwidth. House and Stevens (1958) also measured bandwidths for the open glottis condition for their male subjects, and found that bandwidth did indeed increase under this condition. In fact, measurement of the $F1$ bandwidth can provide an indirect indication of the degree to which the glottis fails to close completely during a cycle of glottal vibration.

Bandwidths, particularly the first-formant bandwidth ($B1$), can also be estimated from the speech waveform. If the $F1$ oscillation is assumed to be of the form $e^{-\alpha t}\cos 2\pi f t$, that is, a damped sinusoid, where f is the frequency of the first formant, then the constant α (in s^{-1}) is

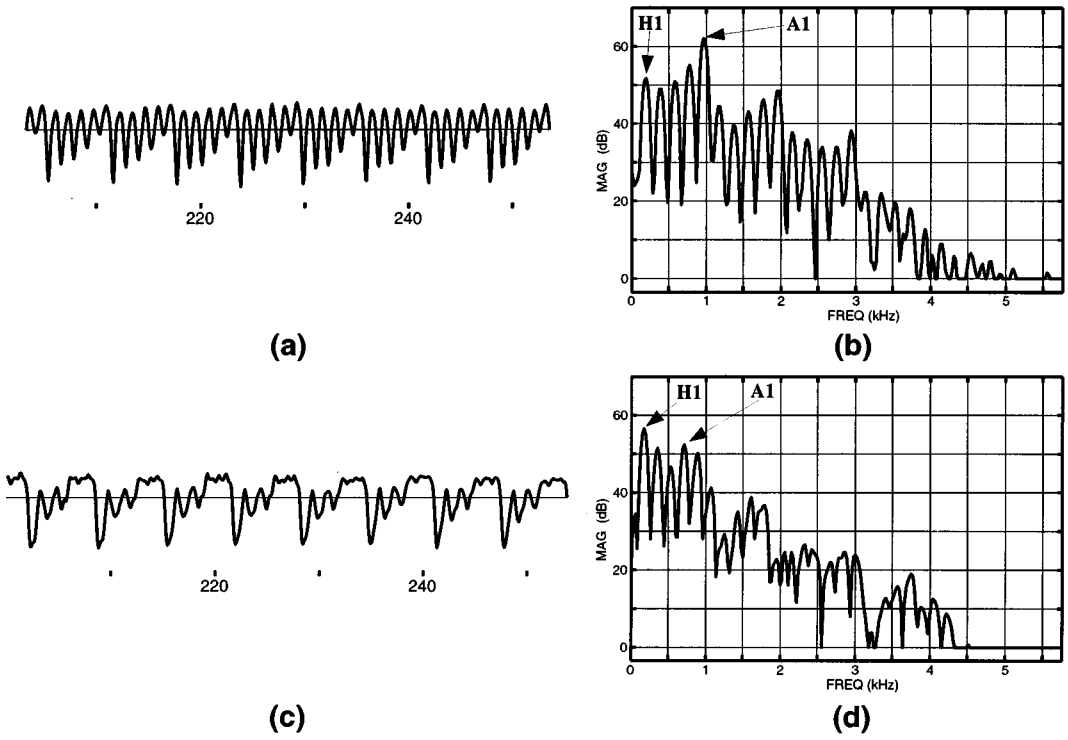


FIG. 3. Examples of waveforms and spectra of the vowel /æ/ produced by two different adult female speakers. The waveforms illustrate decay rates that are (a) slow and (c) rapid, corresponding to narrow and wide first-formant bandwidths, respectively. As estimated from the decay of the speech waveform, the bandwidths during the first part of the cycle are about 60 Hz for (a) and 275 Hz for (c). From the corresponding spectra in (b) and (d), we see that a narrow first-formant bandwidth results in a stronger, more prominent first-formant peak.

related to the bandwidth $B1$ by the equation $B1=\alpha/\pi$ Hz (see, for example, Siebert, 1986). Then by measuring the decay rate of the first-formant waveform during the early part of the glottal period, where the glottal area is expected to be smallest, one can estimate the first-formant bandwidth. Care must be taken to avoid the initial excitation spike when using this method, although bandpass filtering to isolate the first-formant waveform may alleviate this problem. To obtain an accurate estimate, there must be a high enough first-formant frequency and a long enough pitch period to get at least two oscillations during the closed part of the cycle.

As an example of how $F1$ bandwidth is manifested in the acoustic sound pressure, waveforms of the radiated sound pressure for the vowel /æ/ produced by two different female speakers are shown in Fig. 3(a) and (c). The sound-pressure waveform during the initial part of each glottal period is a damped oscillation, the largest component of which is at the frequency of the first formant. The increased rate of decay of the $F1$ oscillation during the last part of each cycle [particularly in Fig. 3(c)] reflects the increased losses at the glottis, and hence the increased bandwidth, during the open phase (Fant, 1979). If the glottis remains open throughout the cycle of vibration, the decay rate during the first part of the glottal cycle will also be increased relative to that for the closed-glottis condition. For the waveforms in Fig. 3, the first-formant bandwidths during the first part of the cycle, as estimated from the decay rate, are about 60 Hz for (a) and 275 Hz for (c).

Another indication of $F1$ bandwidth is the amplitude of the $F1$ peak in the speech spectrum. As predicted by theory, this amplitude should be proportional to the inverse of the bandwidth. Given the bandwidths estimated for the waveforms in Fig. 3, we might expect a difference in relative amplitudes of the $F1$ peaks in the corresponding spectra to be about 12 dB. From the spectra in Fig. 3(b) and (d), we see that the larger bandwidth in (c) results in a reduced $F1$ peak amplitude, making the peak less prominent relative to the amplitude of the first harmonic. The values of $H1-A1$ for these two spectra are about -10 dB for (b) and 4 dB for (d), resulting in a difference of about 14 dB between the two spectra, close to that predicted. This difference is mainly due to the difference in first-formant amplitude ($A1$), but is also partially due to the variation in the relative value of $H1$ across the two speakers.

This example suggests that the amplitude of $F1$ relative to that of the first harmonic ($H1-A1$) may also be a suitable indication of bandwidth. However, this method gives an average bandwidth over the entire glottal cycle, including those times when the glottis is open. During the open phase both $F1$ and its bandwidth increase. Thus, this method may give a larger value of bandwidth than that estimated from the waveform near the beginning of the glottal cycle. Variation across speakers in the relative amplitude of the first harmonic will also add some uncertainty to this acoustic parameter. In addition, $A1$ will vary depending on whether or not $F1$ is centered on a harmonic. The values expected for $H1-A1$ have been estimated to range from a minimum of about -11 dB to a maximum of 5 dB for female speakers (Hanson, 1995), based on theoretical analysis using the KLGLOTT88 glottal source model (Klatt and Klatt, 1990) and the range of

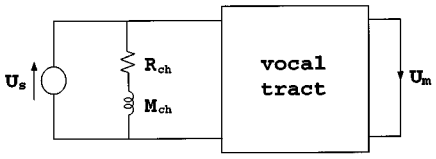


FIG. 4. Model of speech production when the membranous part of the folds have come together, but an opening remains at the arytenoid cartilages, the vocal processes, or both. R_{ch} and M_{ch} represent the resistance and mass of the glottal opening, U_s the volume velocity at the source, and U_m the volume velocity at the mouth.

minimum flows reported by Holmberg *et al.* (1994a).

The estimate of bandwidth based on the decay of the $F1$ oscillations can be used to estimate the area of the glottis during the maximally constricted part of the cycle. Theoretical estimates of the contribution of the glottal opening to the bandwidth $B1$ of the first formant can be made by calculating the value of the resistive termination at the glottis and determining the acoustic energy loss in this resistance (Fant, 1960). An equivalent circuit for calculating the losses is given in Fig. 4. The glottal impedance is represented by an acoustic resistance and an acoustic mass. If we assume that the glottis terminates a uniform vocal tract of length l_v and cross-sectional area A_v, the contribution B_g to the bandwidth of a formant is

$$B_g = \frac{\rho c^2}{\pi A_v l_v R_{\mathrm{ch}}(1+4\pi^2 f^2 M_{\mathrm{ch}}^2/R_{\mathrm{ch}}^2)}, \tag{4}$$

where ρ is the density of air in the vocal tract (g/cm^3), c is the speed of sound (cm/s), f is the formant frequency (Hz), R_{ch} is the glottal resistance due to the chink (dyn s/cm^5), and M_{ch} is the acoustic mass of the glottal chink (g/cm^4) (Stevens, in preparation). From this equation we see that, for a given glottal opening, as f increases, B_g decreases, so the glottal opening has its greatest effect on the bandwidth of $F1$.

Assuming that the pressure drop across the glottis is equal to the subglottal pressure P_s, it has been shown that

$$R_{\mathrm{ch}} = \frac{d\Delta P}{dU_{\mathrm{ch}}} \approx \frac{\rho U_{\mathrm{ch}}}{A_{\mathrm{ch}}^2} \tag{5}$$

and

$$U_{\mathrm{ch}} = A_{\mathrm{ch}} \sqrt{\frac{2P_s}{\rho}}, \tag{6}$$

where U_{ch} is the airflow through the glottal chink and A_{ch} is the area of the chink (Fant, 1960). M_{ch} can be expressed as

$$M_{\mathrm{ch}} = \frac{\rho l_g}{A_{\mathrm{ch}}}, \tag{7}$$

where l_g is the vertical thickness of the glottis (Fant, 1960). Thus, for a given subglottal pressure and thickness of the glottis, we can calculate B_g and U_{ch} as functions of A_{ch}, using Eqs. (4)–(7). Table I lists a range of A_{ch} values and the corresponding values of B_g, $B1$, and U_{ch}, where $B1 = B_g + B_v$, B_v being the $F1$ bandwidth due to vocal-tract losses (with a closed glottis). For /æ/, B_v is approximately 50 Hz for female speakers of Swedish (Fujimura and Lindqvist, 1971; Fant, 1972). From this table we see that bandwidth increments B_g up to 200 Hz might be expected for glottal openings in the range up to 8 mm^2, when the subglottal pressure P_s is assumed to be 5500 dyn/cm^2. This minimum opening corresponds to a minimum flow of about 249 cm^3/s, which is about the upper limit observed by Holmberg *et al.* (1994a) for 15 female speakers of American English.

TABLE I. Range of glottal chink areas (A_{ch}) and corresponding estimations of: glottal contribution to first formant (B_g); bandwidth of first formant ($B1$); flow through chink (U_{ch}); time constant (T) of the rate of change of flow near closure; and resulting increase in spectral tilt at 2750 Hz. We assume a subglottal pressure of 5500 dynes/cm^2, and vocal tract losses of 50 Hz. See text for equations for B_g, U_{ch}, and T.

A_{ch} (cm^2)	B_g (Hz)	$B1$ (Hz)	20 $\log_{10}$ $B1$ (dB)	U_{ch} (cm^3/s)	T (ms)	Tilt (dB)
0.00	0	50	34	0	0	0
0.01	25	75	38	31	0.13	7
0.02	50	100	40	62	0.16	9
0.03	76	126	42	93	0.20	11
0.04	101	151	44	124	0.23	12
0.05	126	176	45	155	0.27	13
0.06	151	201	46	186	0.30	14
0.07	176	226	47	217	0.33	15
0.08	202	252	48	249	0.37	16
0.09	227	277	49	280	0.40	17
0.10	252	302	50	311	0.43	18

2. Effect on spectral tilt

When there is a glottal chink with the arytenoid cartilages approximated at the vocal processes, the pattern of mechanical vibration of the vocal folds should be approximately the same as it is when there is no glottal chink. The shape of the airflow waveform will, however, be influenced by the bypass through the interarytenoid space, particularly at the time when the vocal folds come together. Although there may be a discontinuity in the rate of closure of the glottis, the acoustic mass of the airway and the presence of the bypass path prevent this discontinuous change from being present in the modulated portion of the glottal airflow.

The glottal flow in the vicinity of the time of closure when there is a glottal chink can be approximated by the response of the circuit model in Fig. 5. When the switch is closed, the circuit represents the case where the glottis is

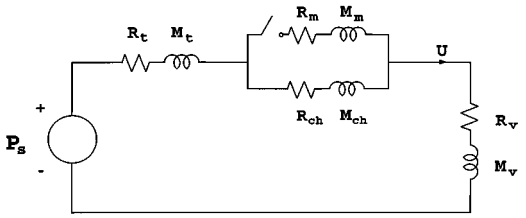


FIG. 5. Model of glottal flow when a fixed opening remains at the arytenoid cartilages during the ''closed'' phase of the glottal cycle. When the switch is closed the situation just prior to closure is modeled. At closure, the switch opens, and the rate of change of flow is limited by the time constant of the circuit. In this figure, P_s represents the subglottal pressure and U the glottal flow. R_t and M_t are the acoustic resistance and mass of the trachea, while R_v and M_v are those of the vocal tract. R_m and M_m are the acoustic resistance and mass of the opening at the membranous part of the vocal folds, and R_{ch} and M_{ch} are those of the glottal chink.

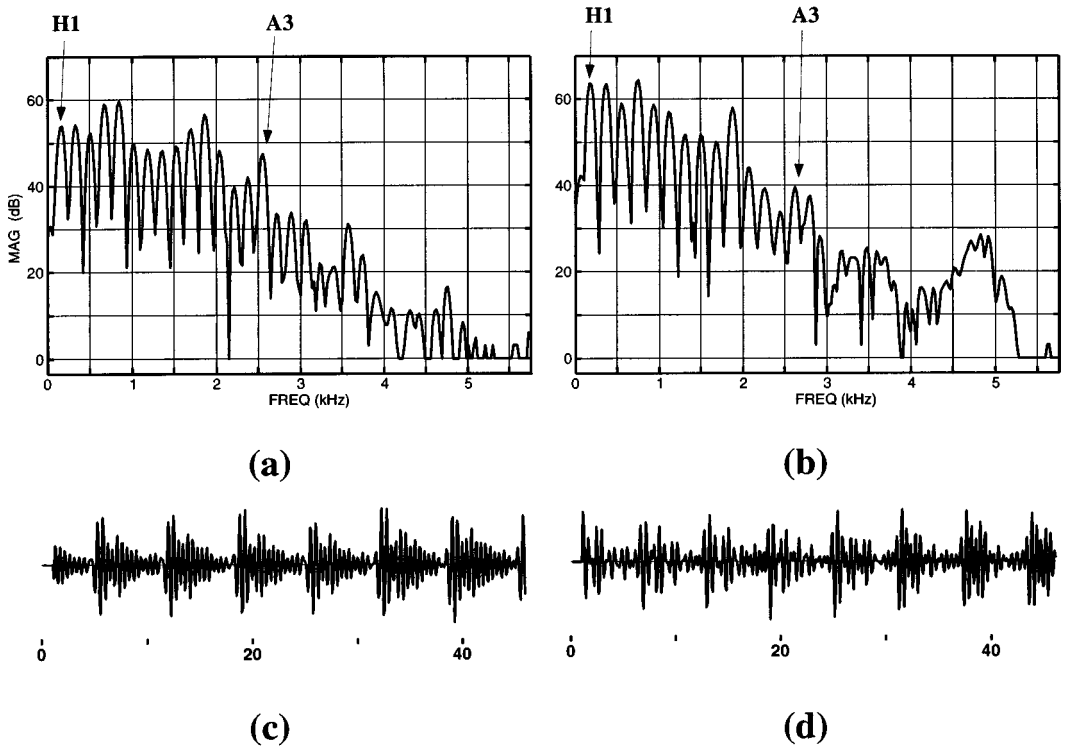


FIG. 6. (a)–(b) Spectra of the vowel /æ/ produced by two adult female speakers with different amounts of spectral tilt. Time window for calculating spectrum is 22.3 ms; (c)–(d) Bandpass-filtered third-formant waveforms corresponding to the spectra in (a)–(b). The center frequency of the filter was 2525 and 2650 Hz for (c) and (d), respectively, and the bandwidth was 600 Hz for both (c) and (d).

open at the folds, and there is flow in both branches of the circuit, giving a total flow U. The closure of the folds is modeled by the sudden opening of the switch. Thus, the rate of change of flow U at the instant of closure is limited by the time constant $T=M/R_{\mathrm{ch}}$, M representing the acoustic mass of the air in the trachea (M_t), glottal chink (M_c), and vocal tract (M_v),

$$M=M_t+M_c+M_v=\rho\left(\frac{l_t}{A_t}+\frac{l_g}{A_{\mathrm{ch}}}+\frac{l_v}{A_v}\right) \tag{8}$$

(Stevens, in preparation), where l_t is the length of the trachea, A_t is the cross-sectional area of the trachea, l_g is the effective vertical length of the glottis, A_{ch} is the cross-sectional area of the glottal chink, l_v is the length of the vocal tract, and A_v is the cross-sectional area of the vocal tract. R_{ch} is the acoustic resistance of the glottal chink, and we have assumed that the acoustic resistance of the trachea and the vocal tract are negligible in comparison to R_{ch}. The length and cross-sectional area of the trachea are about 11 cm and 2 cm^2, respectively, for females (Zemlin, 1988), and l_g is about 0.3 cm (based on data from Titze, 1989a, 1989b, and taking into account end effects). If we assume a vocal tract in a neutral setting, with a length l_v of 15 cm and cross-sectional area A_v of 3 cm^2, then

$$M\approx\rho\left(10.5+\frac{0.3}{A_{\mathrm{ch}}}\right). \tag{9}$$

Using Eqs. (5) and (6), the time constant is then

$$T=\frac{M}{R_{\mathrm{ch}}}=\sqrt{\frac{\rho}{2P_s}}\,(10.5A_{\mathrm{ch}}+0.3). \tag{10}$$

This time constant leads to an additional 6 dB/octave tilt in the spectrum at high frequencies (cf. Siebert, 1986), with the extra tilt beginning at a frequency $f_T=1/2\pi T$. This breakpoint can be translated into a measure of the number of decibels reduction in spectrum amplitude at 2750 Hz, which is approximately the frequency of the third formant for a female speaker. Table I summarizes some time constants T and the corresponding increases in spectral tilt that might be expected for a range of glottal chink areas. Based on minimum airflows measured from inverse-filtered waveforms (Holmberg *et al.*, 1994a) which have a range up to 256 cm^3/s, the maximum increase in tilt that one should expect due to a glottal chink is about 16 dB.

When the arytenoid cartilages remain abducted at the vocal processes throughout the glottal vibration cycle, the membranous part of the folds does not close abruptly, but rather closes nonsimultaneously along the length of the glottis. As discussed in Sec. I A, this nonabrupt closing can contribute significantly to the spectral tilt at mid to high frequencies, depending on the time it takes for the folds to close. With a glottal configuration that has both a fixed space between the arytenoids and some separation at the vocal processes, the effect on the spectral tilt in the third-formant frequency region could be considerable. Thus, depending on the positioning of the arytenoids, including the vocal processes, during phonation, one might expect variations in the $F3$

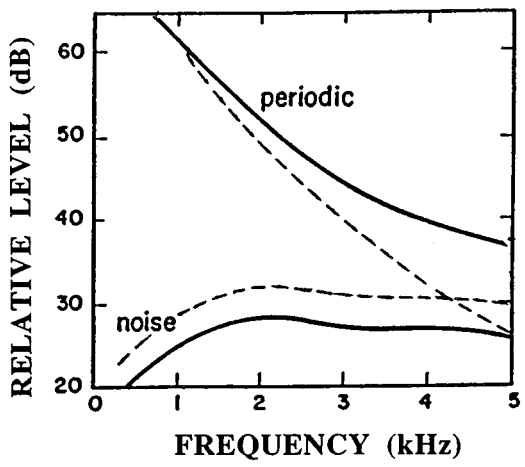


FIG. 7. Calculated spectra and relative amplitudes of periodic volume-velocity source and turbulence-noise source for two different glottal configurations: a configuration in which the glottis has a relatively small minimum opening over one-half of the cycle (solid lines), and a configuration for which the minimum glottal opening has increased (dashed lines). The spectrum for the periodic component gives the amplitudes of the individual harmonics. The noise spectrum is the spectrum amplitude in 50 Hz bands. The calculations are based on theoretical models of glottal vibration and of turbulence noise generation (Stevens, 1993; Shadle, 1985). (From Stevens and Hanson, 1995 and Stevens, in preparation.)

range of the high-frequency spectrum that are substantially greater than 16 dB.

Examples of spectra for the vowel /æ/ produced by two different female speakers are displayed in Fig. 6(a) and (b). These spectra illustrate two extremes of spectral tilt. As discussed in Sec. I A, the amplitude of the first harmonic relative to that of the third-formant peak ($H1-A3$) may be a suitable acoustic correlate of spectral tilt, if certain corrections are made. The values of $H1-A3$ in these two examples are 6 and 23 dB.

3. Turbulence noise at the glottis

Another acoustic consequence of a glottal opening is the generation of turbulence noise in the vicinity of the glottis. It is possible to make estimates of the amplitude and the spectrum of the turbulence noise source at the glottis when the glottal area and the transglottal pressure are known (Shadle, 1985; Stevens, 1993). We can then compare the spectrum of the periodic glottal source to the effective spectrum of the noise source.

When there is vocal-fold vibration with a relatively small glottal opening during half of the cycle, the comparison of the periodic and noise source spectra is shown by the solid lines in Fig. 7. These source spectra are the result of calculations based on theoretical and experimental data from turbulence noise sources and from periodic glottal sources (Shadle, 1985; Stevens, 1993). During phonation both of these sources are filtered by essentially the same vocal-tract transfer function to yield formant prominences. The ratio of the amplitude of the harmonics at 3 kHz to the noise amplitude in a 50-Hz band at the same frequency is 17 dB. Over the entire frequency range up to 5 kHz the noise spectrum is well below the spectrum of the periodic source, so that the combined spectrum is expected to show well-defined harmonics.

When the minimum glottal opening becomes larger, the spectra given by solid lines in Fig. 7 change in two ways. The spectrum amplitude of the periodic component becomes weaker at high frequencies, as noted in Sec. I B 2, and the amplitude of the turbulence noise increases because of the increased flow.[4] For a given subglottal pressure, the amplitude of the turbulence noise source at the glottis is expected to increase approximately in proportion to $A_g^{0.5}$, where A_g is the average glottal area during a cycle of vibration (Stevens, 1971). The two spectra now have the form given as dashed lines in Fig. 7, with the noise spectrum being comparable to the periodic spectrum at high frequencies.

Figure 6(a) and (b) shows the effect of turbulence noise at the glottis in the spectra of natural vowels. The harmonic structure of the spectrum in Fig. 6(b), which has a more extreme tilt, is less apparent at high frequencies (2.5 kHz and above) than that of Fig. 6(a), presumably because of the effect of the aspiration noise. The influence of aspiration noise can also be seen by examining a vowel waveform when it is bandpass filtered at $F3$, with a bandwidth of 600 Hz. The two $F3$ waveforms corresponding to Fig. 6(a) and (b) are displayed in Fig. 6(c) and (d). The waveform in Fig. 6(c), while showing signs of noise excitation, still has a strongly periodic nature. However, the waveform in Fig. 6(d) shows mainly noise, and less evidence of periodic excitation.

Numerous researchers have developed objective measures of the noise present in the speech waveform during glottal vibration (see, for example, Yumoto *et al.*, 1982; Ladefoged and Antoñanzas-Barroso, 1985; Kasuya *et al.*, 1986; Klingholz, 1987; de Krom, 1993; Hillenbrand *et al.*, 1994; Mori *et al.*, 1994). Usually these methods involve isolating the periodic component of the speech waveform from the noisy component. This separation can be done through spectral- or cepstral-based analysis, or through comparison of pitch periods in the time domain, measuring the differences between pitch periods that result from the statistical variability of noise. However, as pointed out by Ladefoged and Antoñanzas-Barroso (1985), these methods do not measure just the noise that is due to an aspiration source, but rather the noise that results from a combination of factors. These other factors include jitter and shimmer. Their solution was to use only *part* of a vibratory cycle and compare it with the corresponding part of the next cycle. However, unless the fundamental frequency is an exact multiple of the sampling period, even a perfectly periodic waveform will appear aperiodic, due to frequency components near the Nyquist frequency that are represented by only a few samples (Klatt and Klatt, 1990). This source of variation can only be remedied by significant oversampling.

Klatt and Klatt (1990) used another technique to estimate aspiration noise, avoiding the inclusion of noise due to other factors. In this method, the amount of aspiration noise in relation to the periodic component is estimated subjectively by examining the bandpass-filtered waveform in the $F3$ region, such as those in Fig. 6(c) and (d). It is also possible for an observer to make estimates of the amount of

TABLE II. Dialect history of the 22 female subjects, age 0–12 years.

	Age	Dialect history
F1	30	Massachusetts (north shore)
F2	22	Rhode Island
F3	27	0–5 North Carolina; 5–6 Brussels; 6–12 Zaire
F4	38	New York; Illinois; Maryland
F5	44	St. Louis, MO
F6	36	Atlanta, GA
F7	25	Virginia; California
F8	23	Massachusetts
F9	32	Schenectady, NY
F10	41	Massachusetts
F11	30	Attleboro, MA
F12	22	0–8 California; 8–12 New Hampshire
F13	31	Southeastern Connecticut
F14	14	0–3 western Missouri; 3–5 eastern Pennsylvania; 5–12 southeastern Iowa
F15	27	Massachusetts
F16	37	0–3 Queens, NY; 3–12 Long Island, NY
F17	30	Massachusetts
F18	38	Buffalo, NY
F19	30	Haverhill, MA
F20	26	0–3 New York, NY; 3–8 Haiti; 8–12 Brockton, MA
F21	37	0–6 Arkansas; 6–7 Berkeley, CA; 7–8 Arkansas; 8–12 Missouri
F22	49	0–3 Syracuse, NY; 4–12 Philadelphia, PA

noise in a spectral representation, such as those of Fig. 6(a) and (b). The observer makes estimates of the amount of noise on a scale from 1 to 4, where 1 means there is essentially no evidence of noise interference and 4 means that there is little evidence of periodicity. Separate estimates are made from the waveform and from the high-frequency part of the spectrum.

II. EXPERIMENTAL DATA

Based on the theoretical discussion of Sec. I, we suggested several measures that can be made directly on the spectra and waveforms of natural vowels and that may give some indication of the vocal-fold and glottal configuration during vowel production. The theory predicts relationships between these measures in some cases, particularly under conditions where the glottis does not close completely during some part of the vibration cycle. For example, we see in Table I that as the area of the glottal chink increases, both the $F1$ bandwidth and the spectral tilt are expected to increase, and we also expect the strength of the noise source to increase. In this section we describe data collected from a group of female speakers, and we attempt to interpret these data in terms of the theoretical models, using the proposed acoustic parameters.

A. Speakers and speech material

We collected recordings of a number of utterances from 22 adult female subjects. Subjects were recruited from the Speech Communication Group at MIT and a group of speech pathology students at Massachusetts General Hospital. The age and dialect history of each subject are listed in Table II. The speakers showed no evidence of voice or hearing problems, and all were native speakers of American English. Eleven speakers had experience being subjects for speech production experiments. The speakers had a wide range of dialects, but nearly half of the group grew up in New England. Subjectively, the vowel qualities produced were quite similar to one another, with a few exceptions for /æ/.

The utterances consisted of three nonhigh vowels, /æ, ɛ, ʌ/, embedded in the carrier phrase ''Say bVd again.'' These vowels were chosen because the first formant is well separated from the first harmonic, simplifying the acoustic measures to be described below. The recordings were made in a sound-isolated chamber. The subjects were instructed to speak naturally and to put emphasis on the word /bVd/. Each utterance was repeated five times, with the 15 sentences presented in random order during a single session. All the utterances were low-pass filtered at 4.5 kHz, digitized with a sampling rate of 11.4 kHz, and stored for further analysis.

B. Measurements

The acoustic measurements described in Sec. I were extracted from these utterances in the following manner:

First-formant bandwidths. For all repetitions of the vowel /æ/ the first-formant bandwidth during the initial part of the glottal cycle was estimated from the rate of decay of the waveform. Analysis was restricted to /æ/ because its first formant is usually high enough to get at least two oscillations during the closed part of the glottal cycle, and the second formant is well separated from the first. The rate of decay was determined from the change in the peak-to-peak amplitude in the first two cycles of the $F1$ oscillation. Estimates were made for eight consecutive pitch periods in a relatively stable portion of the vowel, generally at the middle. To reduce interference by the second formant, the waveforms were bandpass filtered with a filter having a bandwidth of 600 Hz centered at the first formant frequency. These 40 estimates were then averaged to obtain a mean value for each speaker.

$H1^*-H2^*$. The difference between the amplitudes of the first and second harmonics was measured for all repetitions of all three vowels. For /æ/, $H1-H2$ was measured from the spectrum obtained by centering a 22.3-ms Hamming window during the initial part of the glottal cycle, at the eight points where the $F1$ bandwidth was estimated. For /ʌ/ and /ɛ/, the measurements were taken at three points in midvowel, 20 ms apart, where the formants were relatively stable. Corrections were made for the amounts by which $H1$ and $H2$ are ''boosted'' by the effect of the first formant on the vocal-tract transfer function,[5] yielding the measure $H1^*-H2^*$. This corrected measure can be compared across vowels and across speakers. The values for each repetition were averaged to obtain a mean value for each vowel for each speaker.

$H1^*-A1$. The difference between the (corrected) amplitude of the first harmonic and the amplitude of the first formant peak ($A1$) was measured. $A1$ was estimated by measuring the amplitude of the strongest harmonic of the $F1$ peak. The measurements were taken at the same points as those for $H1^*-H2^*$, and similarly, average values were computed for the three vowels for each speaker.

$H1^*-A3^*$. The difference between the amplitudes of the first harmonic and the third formant peak ($A3$) was mea-

TABLE III. Average values of the acoustic parameters for the vowel /æ/, 22 female speakers, where $H1^*-H2^*$, $H1^*-A1$, and $H1^*-A3^*$ are given in dB, N_w and N_s are the waveform- and spectra-based noise judgments, and $B1$ is the bandwidth of the first formant, given in Hz. Numbers in boldface represent maxima or minima for each measure across speakers. The mean values and average standard deviations across speaker are also given.

Subject	$H1^*-H2^*$	$H1^*-A1$	$H1^*-A3^*$	N_w	N_s	$B1$
F1	3.5	−0.2	30.7	3.0	2.8	194
F2	1.7	0.4	32.2	2.8	2.9	244
F3	4.4	−8.0	32.1	2.7	2.8	94
F4	1.6	−5.7	**13.0**	1.6	1.6	209
F5	5.4	2.2	**35.0**	**3.8**	2.7	245
F6	2.4	−5.5	23.0	**1.1**	**1.1**	153
F7	3.8	−1.3	31.3	3.1	**3.1**	150
F8	2.1	−3.7	32.6	2.9	2.7	97
F9	2.8	−7.2	16.8	1.2	1.2	104
F10	5.0	**3.9**	26.4	2.2	2.6	184
F11	4.5	−4.4	19.5	1.8	2.1	158
F12	0.7	−5.6	31.3	2.4	2.2	217
F13	3.8	−8.9	19.4	1.7	1.2	**53**
F14	5.2	**−11.3**	16.3	**1.1**	1.2	78
F15	6.2	0.3	33.7	3.1	2.4	256
F16	**6.8**	1.2	30.4	2.3	2.5	132
F17	1.6	−2.6	22.0	2.0	1.8	**280**
F18	4.5	−2.2	21.8	2.0	2.5	163
F19	5.4	−0.5	24.3	2.0	2.0	166
F20	0.9	−6.2	14.7	1.7	1.6	178
F21	0.8	−8.5	17.9	1.5	1.4	124
F22	**0.6**	−9.2	20.8	1.4	1.2	149
Mean	3.4	−4.2	24.1	2.1	2.1	165
Mean s.d.	1.4	2.3	3.4	0.5	0.5	34

TABLE IV. Average values of the acoustic parameters for the vowel /ʌ/, 22 female speakers, where $H1^*-H2^*$, $H1^*-A1$, and $H1^*-A3^*$ are given in dB, and N_w and N_s are the waveform- and spectra-based noise judgments. Numbers in boldface represent maxima and minima for each measure across speaker. The mean values and average standard deviations across speaker are also given.

Subject	$H1^*-H2^*$	$H1^*-A1$	$H1^*-A3^*$	N_w	N_s
F1	4.8	2.8	26.4	3.0	2.8
F2	1.2	−0.3	25.2	2.7	2.9
F3	3.6	−1.7	26.0	2.7	2.7
F4	**−0.7**	−9.0	**10.9**	1.8	1.3
F5	3.7	1.5	29.1	2.3	2.4
F6	3.0	−6.6	18.9	**1.4**	1.2
F7	1.8	−1.0	28.3	3.2	**3.5**
F8	3.0	−2.7	29.2	2.5	2.3
F9	1.5	−6.4	20.6	1.7	1.8
F10	3.1	2.8	24.7	2.4	2.3
F11	3.9	−2.9	22.0	1.7	2.1
F12	2.2	−5.8	22.9	2.2	1.9
F13	2.7	−4.4	15.5	**1.4**	**1.1**
F14	5.1	−11.9	15.1	**1.4**	1.3
F15	3.6	−4.0	27.2	2.9	2.3
F16	**5.8**	**3.5**	24.6	2.0	2.3
F17	1.5	−4.0	22.7	2.4	1.7
F18	3.5	−2.8	18.5	1.7	2.0
F19	5.0	1.3	**34.1**	**3.5**	3.2
F20	−0.2	−9.9	14.9	1.6	1.7
F21	0.1	−6.8	20.5	2.5	1.6
F22	0.3	**−12.1**	14.8	2.1	1.2
Mean	2.6	−4.1	22.0	2.2	2.0
Mean s.d.	1.1	1.7	3.3	0.6	0.6

sured. As was done for $A1$, $A3$ was estimated using the strongest harmonic of the $F3$ peak. $H1$ was corrected as above, and $A3$ was corrected for the effect of $F1$ and $F2$ on the spectrum amplitude of the third formant.[6] For this normalization neutral first and second formant frequencies were set to 555 and 1665 Hz, respectively, based on the average third-formant frequency measured for all speakers. As mentioned in Sec. I A, $A3$ is also dependent on the bandwidth of $F3$. House and Stevens (1958) measured $F3$ bandwidths of male speakers of English for /æ, ʌ, ɛ/ to be 103, 64, and 88 Hz, respectively. In dB this means that /æ/ is expected to have an $F3$ amplitude that is 4 dB less than that of /ʌ/, while that for /ɛ/ is 3 dB less. For females speakers, the bandwidth values will be higher, but because data are not available for these vowels for female speakers, we made corrections based on the male data. This use of male data should result in minimal error because the ratios of the bandwidths are used to compute the difference in dB and these ratios are not expected to differ greatly across gender. Thus, the value of $A3$ measured for each token of /æ/ and /ɛ/ was increased by 4 and 3 dB, respectively. The combination of these two corrections, for the location of $F1$ and $F2$ and for the $F3$ bandwidth, yields a normalized $H1^*-A3^*$.

Noise ratings. All repetitions of the three vowels were bandpass filtered around the third formant using a filter having a bandwidth of 600 Hz. The bandpass-filtered waveforms and the speech spectra corresponding to the speech segments used in the previously described measures were given ratings for noise, as described in Sec. I B 3 and in Klatt and Klatt (1990). These judgments were made independently by two judges, who did not know which waveforms or spectra corresponded to which speaker. Their average ratings were highly correlated ($r>0.92$) and were averaged to obtain two noise judgments for each speaker, one based on the waveforms and the other on the spectra. The waveform-based ratings were found to be well correlated with the spectrum-based ratings.

C. Results

1. Mean values

The mean values of the acoustic parameters for each speaker are summarized in Tables III–V. Minimum and maximum values for each measure across speakers are given in boldface in these tables. $H1^*-H2^*$ has a range of about 10 dB. $H1^*-A3^*$ has a range of about 26 dB, indicating a wide variation in spectral tilt among the subjects. This large range of spectral tilt is assumed to be a consequence of the presence of a glottal chink or nonsimultaneous closure along the length of the glottis, or both, for some speakers. The minimum value of tilt is 8.6 dB, about what might be expected for the case where there is complete, abrupt glottal closure during some part of the glottal cycle (see Sec. I A). The range of $H1^*-A1$ is 16 dB, as predicted earlier, and the minimum and maximum values are very close to those predicted in Sec. I B 1: −11 and 5 dB. The range of values obtained suggests that first-formant peaks vary from being very prominent for some speakers to being highly damped for others, although part of this range can be due to variation

TABLE V. Average values of the acoustic parameters for the vowel /ɛ/, 22 female speakers, where $H1^*-H2^*$, $H1^*-A1$, and $H1^*-A3^*$ are given in dB, and N_w and N_s are the waveform- and spectra-based noise judgments. Numbers in boldface represent maxima and minima for each measure across speakers. Average values and average standard deviations across speaker are also given.

Subject	$H1^*-H2^*$	$H1^*-A1$	$H1^*-A3^*$	N_w	N_s
F1	6.3	1.7	28.8	3.2	2.9
F2	1.3	−2.0	27.4	2.8	2.1
F3	3.5	−3.1	31.9	3.2	3.1
F4	0.9	−11.0	**8.6**	1.7	1.1
F5	5.4	**3.7**	30.6	3.2	3.0
F6	3.3	−9.0	17.3	1.4	**1.0**
F7	3.1	−2.5	27.3	**3.6**	**3.3**
F8	2.6	−3.8	29.8	2.4	2.2
F9	3.0	−4.3	19.9	2.2	1.5
F10	6.5	2.5	22.6	2.7	2.5
F11	4.6	−5.8	18.0	1.8	1.7
F12	1.9	−5.7	26.0	2.1	1.6
F13	3.0	−5.3	16.0	**1.6**	1.1
F14	4.0	**−12.4**	16.6	1.9	1.2
F15	4.0	−1.1	30.2	2.5	1.9
F16	**6.9**	−1.6	29.4	2.9	1.9
F17	2.4	−5.3	27.1	2.7	2.3
F18	4.2	−3.7	16.5	**1.6**	1.6
F19	5.1	−3.9	**32.8**	2.8	2.5
F20	−0.8	−10.3	13.7	1.9	1.4
F21	1.5	−5.5	20.4	1.9	1.2
F22	**−2.6**	−6.7	15.5	1.7	1.2
Mean	3.1	−4.7	22.5	2.3	1.9
Mean s.d.	1.3	1.9	3.4	0.5	0.8

in the amplitude of $H1$ and how well $F1$ is centered on a harmonic across speakers. This range of first-formant amplitudes presumably arises in part due to a range of $F1$ bandwidths and in part due to differences in the degree to which spectral tilt extends to the low-frequency harmonics.

The first-formant bandwidth estimates for /æ/ vary from 53 to 280 Hz. For the speaker with the lowest value of bandwidth (53 Hz), this estimate is about what is expected for the closed-glottis condition (Fant, 1972). For speakers with higher values of bandwidth, losses must exist at the glottis. Theoretical analysis of glottal losses indicates that a first-formant bandwidth of 280 Hz corresponds to a minimum glottal opening of about 0.09 cm^2 (see Table I), while 75 Hz corresponds to about 0.01 cm^2, so we have a range of estimated glottal chink cross-sectional areas of about 0.08 cm^2. The noise judgments range from 1.0 to 3.8; that is, some of our speakers show little to no noise in the high-frequency range, while other speakers have substantial noise.

Because there were only five tokens of each vowel per speaker, it is difficult to judge the consistency of the speakers. Standard deviations of each measure were computed for each speaker, and the average value across speakers is given in the last lines of Tables III–V. We see that the speakers were generally quite consistent, and there are only small differences in consistency across vowel. Given the range of values observed for each acoustic parameter, the variations within a speaker do not seem significant.

2. Statistical analysis

All acoustic parameters (except $B1$) were subjected to repeated measures analyses of variance (ANOVA) with vowel (/æ/ vs /ʌ/ vs /ɛ/) as a within-subject factor. The results are summarized in Table VI. There was a significant effect for the parameters $H1^*-H2^*$ and $H1^*-A3^*$ [$F(2,42)=4.04$, $p<0.05$, and $F(2,42)=5.25$, $p<0.01$, respectively]. However, referring to Tables III–V and comparing mean values across vowel for these two parameters, we see that, despite the statistical significance, there are only very small differences (1–2 dB). Thus, it would seem that the corrections made to $H1$, $H2$, and $A3$ for vowel quality (see Sec. II B) were successful in minimizing differences across vowels.

TABLE VI. Results of repeated measures analyses of variance (ANOVA) performed to examine differences in the acoustic parameters across vowels. An asterisk (*) indicates statistical significance at the 0.05 level.

Measure	$F(2,42)$	p
$H1^*-H2^*$	4.04	<0.05*
$H1^*-A1$	0.85	>0.1
$H1^*-A3^*$	5.25	<0.01*
Waveform-based noise	1.97	>0.1
Spectra-based noise	2.25	>0.1

Table VII shows Pearson product moment correlation coefficients (r) for the various measures for each vowel. In the following discussion we consider a correlation with $r \geqslant 0.70$ to be strong. The strongest correlation was found between the high-frequency noise ratings and the measure $H1^*-A3^*$ ($r \geqslant 0.79$, $N=22$). As mentioned earlier, this is not unexpected given that both tilt and noise are expected to increase with the area of a fixed glottal opening (see Table I and the discussion in Secs. I B 2 and I B 3). $H1^*-A1$ tends to have a strong correlation with the spectra-based noise ratings ($r>0.70$, $N=22$ for four out of six correlations). Again, this is predicted from earlier discussion (see Table I where $B1$ increases with A_{ch}). For the vowels /ʌ/ and /ɛ/, $H1^*-A3^*$ is well correlated with $H1^*-A1$ ($r \geqslant 0.70$, $N=22$), but the correlation is only moderate for /æ/ ($r=0.62$, $N=22$). Finally, the correlation between $H1^*-A1$ and estimated $F1$ bandwidth for /æ/ is moderate ($r=0.61$, $N=22$).

Although one might expect a larger open quotient to lead to greater losses and noise, due to an increase in average glottal area, the difference $H1^*-H2^*$ is not well correlated with any other measure ($r<0.59$, $N=22$). This result could be interpreted to mean that $H1^*-H2^*$ is not a good indication of open quotient. However, Holmberg *et al.* (1995) found $H1^*-H2^*$ to be well correlated with the adduction quotient in simultaneous observations of airflow and acoustic spectra for female speakers. Assuming that adduction quotient is 100−open quotient, $H1^*-H2^*$ is then also well correlated with open quotient. Therefore, our results may indicate that open quotient is nearly independent of other glottal parameters. For example, a speaker may adjust her glottal configuration in such a way that a larger open quotient results while the rate of decrease of flow at glottal closure remains nearly the same. Thus, $H1^*-H2^*$ would increase while $H1^*-A3^*$ remains the same.

Table VIII shows the Pearson product moment correlation coefficients for the three vowels combined. The noise measures are strongly correlated with the tilt-related measure

TABLE VII. Pearson product moment correlation coefficients (r) for the acoustic parameters for each of the three vowels /æ, ʌ, ɛ/. Numbers in boldface represent strong correlations ($r>0.70$). The notation n.s. indicates that a correlation was not significant. $N=22$.

		$H1^*-H2^*$	$H1^*-A1$	$H1^*-A3^*$	N_w	N_s	$B1$
/æ/	$H1^*-H2^*$	1					
	$H1^*-A1$	0.47	1				
	$H1^*-A3^*$	n.s.	0.62	1			
	N_w	n.s.	0.67	**0.87**	1		
	N_s	0.38	**0.72**	**0.82**	**0.88**	1	
	$B1$	n.s.	0.61	n.s.	0.45	n.s.	1
/ʌ/	$H1^*-H2^*$	1					
	$H1^*-A1$	0.57	1				
	$H1^*-A3^*$	0.51	**0.78**	1			
	N_w	n.s.	0.56	**0.81**	1		
	N_s	n.s.	**0.75**	**0.84**	**0.83**	1	
/ɛ/	$H1^*-H2^*$	1					
	$H1^*-A1$	0.59	1				
	$H1^*-A3^*$	0.49	**0.70**				
	N_w	0.45	**0.71**	**0.82**	1		
	N_s	0.48	**0.73**	**0.79**	**0.94**	1	

$H1^*-A3^*$, and the spectra-based noise measure is strongly correlated with the bandwidth-related measure $H1^*-A1$. In addition, $H1^*-A1$ has a good correlation with the measure $H1^*-A3^*$.

D. Interpretation of acoustic measurements

In order to gain a better understanding of the correlations reported in Table VIII, and to perhaps be able to interpret the acoustic measurements in terms of glottal configurations, we examined scatterplots of measures that were well correlated with each other.

Figure 8 plots the measure $H1^*-A3^*$ against $H1^*-A1$. Almost all of the data points with $H1^*-A1$ less than about -6 dB have an $H1^*-A3^*$ measure less than about 23 dB, while all of the data points with $H1^*-A1$ greater than about -2 dB have an $H1^*-A3^*$ measure greater than about 23 dB. Note that the highest $H1^*-A3^*$ measure expected for speakers with a posterior glottal opening and simultaneous closure of the membranous part of the folds is about 25 dB (see Sec. I B 2). Based on this observation, we divided the data points into two groups, depending on whether $H1^*-A3^*$ was less than or equal to 23 dB (group 1) or greater than 23 dB (group 2). Analysis of the two groups revealed that for 19 speakers, all three data points fell into either one group or the other, but not both. Data points for the other three speakers (F10, F12, F17) fell into both groups. Because subjects F10 and F12 had only one point each in group 1, they were assigned to group 2. Speaker F17 had only one point in group 2, so she was assigned to group 1.

In Fig. 8 data points for group 1 speakers are represented by filled circles and those for group 2 are represented by open circles. The 11 speakers in group 1 have relatively low values of $H1^*-A3^*$ and $H1^*-A1$; that is, speakers in this group have shallow spectral tilts and prominent first-formant peaks. Therefore, this group can be hypothesized to have abrupt glottal closures. Some of these speakers may also have posterior glottal chinks, leading to a range of $H1^*-A3^*$ (about 15 dB) and $H1^*-A1$ (about 11 dB).

Speakers in group 2, indicated by open circles, have much higher values of $H1^*-A3^*$, that is, steeper spectral tilts. Because these values are close to or greater than 25 dB, we surmise that the glottal closure is not simultaneous along

TABLE VIII. Pearson product moment correlation coefficients (r) for the acoustic parameters for the three vowels /æ, ʌ, ɛ/ combined. Numbers in boldface represent strong correlations ($r>0.70$). $N=66$.

	$H1^*-H2^*$	$H1^*-A1$	$H1^*-A3^*$	N_w	N_s
$H1^*-H2^*$	1				
$H1^*-A1$	0.53	1			
$H1^*-A3^*$	0.46	0.68	1		
N_w	0.30	0.63	**0.80**	1	
N_s	0.40	**0.73**	**0.80**	**0.86**	1

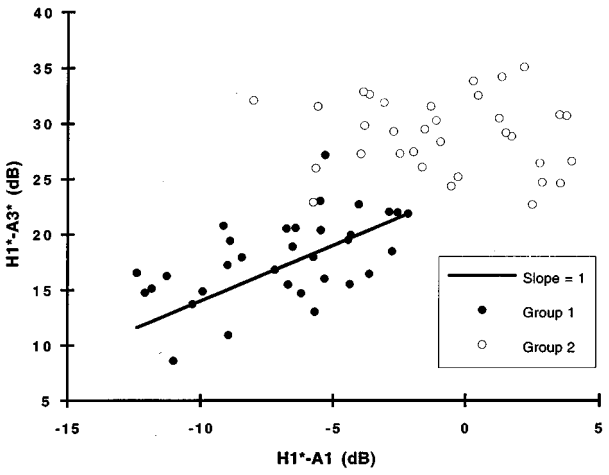


FIG. 8. Relation between $H1^*-A3^*$ and $H1^*-A1$. Each point represents data for one vowel by one speaker. Data points for group 1 are displayed as filled circles and data points for group 2 are displayed as open circles (see text). A line of slope one has been drawn through the data points for group 1, showing the theoretically predicted relationship between spectral tilt and the amplitude of the first formant. (After Stevens and Hanson, 1995.)

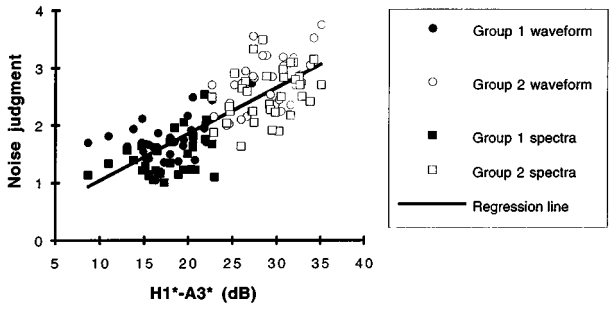


FIG. 9. Relation between noise judgments and $H1^*-A3^*$, together with a regression line ($r^2=0.62$). Points represented as circles are judgments based on waveforms and the squares are based on spectra. Filled points represent group 1 data, while open points represent group 2 data. (After Stevens and Hanson, 1995.)

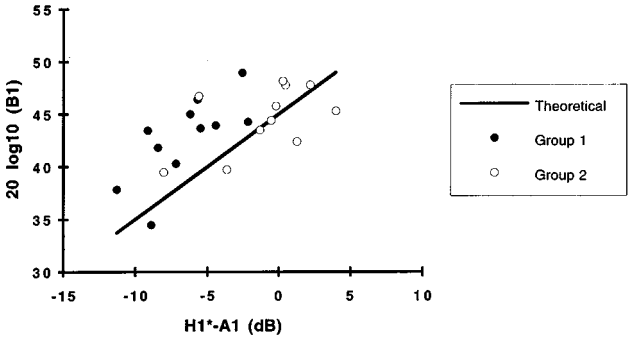


FIG. 10. Relation between $H1^*-A1$ and $F1$ bandwidth (on a log scale) as measured from the waveform. The data are from speakers producing the vowel /æ/. Data points for group 1 members are represented by filled circles, while those for group 2 members are represented by open circles. A straight line representing the slope of the theoretical relationship has been drawn through the data. (After Stevens and Hanson, 1995.)

the length of the membranous part of the vocal folds and that incomplete glottal closure also occurs. This nonsimultaneous closure is probably due to the glottis being abducted at the vocal processes, although the folds could also close nonabruptly when the vocal processes are approximated. The higher values of $H1^*-A1$ for group 2 speakers are due to two influences on $A1$: (1) the first formant has an increased bandwidth because there are greater losses associated with the glottal configuration in which the vocal processes are abducted, and (2) the spectral tilt is so steep that its influence extends down into the first-formant range. $H1^*-A3^*$ does not appear to be correlated with $H1^*-A1$ for group 2 speakers, possibly because the effect of the nonsimultaneous closure is independent of the effect of the posterior glottal opening.

From Table I we see that if the bandwidth of the first formant is expressed on a log (dB) scale, then it should increase with tilt with a slope of 1 for speakers who have abrupt glottal closure. Since $H1^*-A1$ is an indicator of bandwidth and $H1^*-A3^*$ is an indicator of tilt, a similar linear relationship should also exist between these two parameters for speakers with abrupt glottal closure. In Fig. 8 a line with slope 1 has been drawn through the data and is seen to fit nicely with the group 1 points. This result is evidence that group 1 speakers have abrupt glottal closure and posterior glottal openings that range in size across speakers, and that as first-formant bandwidth increases, $H1^*-A1$ increases.

Figure 9 shows the relation between the two types of noise judgments and the tilt parameter $H1^*-A3^*$. Recall that there was a high correlation between these quantities. This figure is also divided into the two groups of speakers of the previous figure. Speakers with greater degrees of tilt show greater amounts of noise in their speech signals, as predicted from the theoretical discussion of Sec. I. A regression line ($r^2=0.62$) has been drawn through the points in Fig. 9.

In Fig. 10 the parameter $H1^*-A1$ is plotted against $F1$ bandwidth (on a log scale) as measured in the first part of the glottal cycle ($B1$) for the 22 speakers producing the vowel /æ/. The data are presented to indicate which points belong to group 1 and group 2 speakers. A line of slope 1 is drawn through the data to represent the relationship expected based on the theoretical development. There is a trend toward a decrease in $F1$ prominence (that is, a decrease in $A1$) as the $F1$ bandwidth increases, but the correlation is only moderate ($r=0.61$, $p<0.01$). The relatively weak correlation may be due to the fact that the prominence of $A1$ depends on the entire glottal cycle, whereas the bandwidth estimate $B1$ is based only on the closed (or minimum glottal area) part of the glottal cycle. Thus, $A1$ is influenced by the open quotient and the glottal aperture during the open phase, but the bandwidth estimate $B1$ is not. In addition, other factors, such as spectral tilt, may reduce $A1$. In fact, given these influences, it is not surprising that the group 1 data in Fig. 10 appear to be better correlated than the group 2 data.

For one speaker (F13) the bandwidth is sufficiently small (53 Hz) that complete glottal closure can be assumed during a portion of the glottal cycle. This speaker is from group 1. For speakers with higher bandwidth and $H1^*-A1$ measures, it is reasonable to assume that the source of loss is an incomplete glottal closure. Two speakers from group 2 (F3 and F8) have fairly narrow bandwidths (94 and 97 Hz), although this would not be expected given our hypothesis that group 2 members have abduction at the vocal processes. The $H1^*-A1$ measure for these speakers indicates that $A1$ is indeed quite prominent, consistent with the narrow bandwidth. The findings for these speakers may indicate that their glottal closure is characterized by adducted vocal processes with no posterior glottal chink, but nonsimultaneous closure within the membranous portion. This interpretation might explain the narrow first-formant bandwidths, high first-formant amplitudes, and steep spectral tilts that these two speakers exhibit (see Tables III–V).

Our results are satisfying in that the ranges of observed values and the relationships between these values are in line with the predictions based on our theoretical development. However, our classification of the subjects by glottal configuration is only an hypothesis. It should be verified by direct observation of the vocal folds during phonation. Preliminary evidence gathered from 4 of the 22 subjects suggests that this hypothesis is correct (Hanson, 1995). These subjects were examined using fiberscopy and two subjects from group 1 were observed to have rather small glottal chinks,

while two subjects from group 2 were observed to have relatively large posterior openings. For one of the group 2 speakers, the opening extended well beyond the vocal processes and closure was possibly nonsimultaneous. However, the fiberscopic images were not collected simultaneously with the sound-pressure signals, and the resolution of the images was not high enough to allow us to make judgments regarding the abruptness of closure. Therefore, a study that includes more subjects, and simultaneous recordings of image and sound-pressure data is needed to verify our classification of glottal configuration based on acoustic parameters.

III. FUTURE WORK

The wide ranges of acoustic parameter values that we have observed in this study suggest that consideration of glottal characteristics has great importance for describing female speech and, in addition to formant frequencies and fundamental frequency, should be taken into account for applications such as speech synthesis, speech recognition, and speaker recognition. The work should be extended in several ways.

As mentioned above, simultaneous collection of physiological and sound-pressure signals is necessary to validate the use of the acoustic parameters for classification of subjects according to glottal configuration. In addition, the current study was limited to female speakers without voice disorders. A study based on male subjects is currently under way and will provide comparisons with the female data (Chuang and Hanson, to appear). These acoustic parameters should also be applied to the speech of people with voice disorders.

Another aspect of these acoustic parameters that remains to be studied is the variability for a given speaker across repeated recordings. Speaking intensity has been found to be systematically related to variations in airflow-based glottal parameters (Holmberg *et al.*, 1988), and thus should also influence acoustic parameters. Such intrasubject variation may be significant (Holmberg *et al.*, 1994a; Hanson, 1995). It should also be noted that source characteristics, and hence perhaps voice quality, change constantly throughout an utterance, due to the effects of prosody and coarticulation between vowels and voiceless consonants. The possible use of the acoustic parameters presented here as a tool to examine these changes should be explored.

Finally, the relation of the proposed parameters to perceived voice quality should be examined. In a preliminary study, speakers from group 2 were perceived to be breathier than speakers from group 1, and listeners could perceive changes in the acoustic parameters obtained by manipulating synthesized speech (Hanson, 1995).

ACKNOWLEDGMENTS

This paper is based on a portion of a Ph.D. thesis submitted to the Harvard University Graduate School of Arts and Sciences. The research was carried out at the M.I.T. Research Laboratory of Electronics. I am grateful to Kenneth N. Stevens for his guidance and encouragement, to Eva Holmberg for her careful reading of the revised paper, and to Alice Turk for her assistance with the statistical analysis. The comments and suggestions of Anders Löfqvist, Ronald C. Scherer, and an anonymous reviewer on an earlier version of the paper are greatly appreciated. This work was supported by the Harvard University Division of Engineering and Applied Sciences, a U.S. Dept. of Education fellowship, and NIH Grant No. DC00075.

[1]Due to the radiation characteristic at the mouth, the radiated sound pressure can be approximated as the derivative of the volume velocity at the mouth. Thus, the derivative of the glottal waveform is the effective excitation (Fant, 1982), and may be of more interest for analysis than the glottal airflow itself.

[2]The synthesized waveforms described in this paper were produced using the KLSYN88 synthesizer (Klatt and Klatt, 1990), which contains several glottal sources, including a representation of the source proposed by Fant *et al.* (1985). Our intention in using these waveforms is simply to illustrate how changes in glottal waveform parameters are manifested in both the glottal source and speech spectra, and not to promote a particular model of the glottal waveform.

[3]This model of nonsimultaneous closure is not based on actual data. It should be verified with high-speed image data.

[4]The amplitude of the turbulence noise increases with glottal area only up to a certain point: when the glottal area becomes large enough, a drop in transglottal pressure occurs, the flow velocity decreases, and the amplitude of the turbulence noise also decreases.

[5]The quantity $20\log_{10}[F1^2/(F1^2-f^2)]$ is subtracted from $H1$ and $H2$, where f is the frequency at which the harmonic is located.

[6]The quantity

$$20\log_{10}\left(\frac{\left[1-\left(\frac{F3}{F1}\right)^2\right]\left[1-\left(\frac{F3}{F2}\right)^2\right]}{\left[1-\left(\frac{F3}{\widetilde{F1}}\right)^2\right]\left[1-\left(\frac{F3}{\widetilde{F2}}\right)^2\right]}\right)$$

is added to $A3$, where $\widetilde{F1}$ and $\widetilde{F2}$ are the first- and second-formant frequencies of a neutral vowel.

Ananthapadmanabha, T. V. (**1984**). ''Acoustic analysis of voice source dynamics,'' Speech Trans. Lab. Q. Prog. Stat. Rep. **2–3**, Royal Institute of Technology, Stockholm, 1–24.

Ananthapadmanabha, T. V. (**1993**). ''Modeling the return phase of the derivative of glottal flow,'' Speech Communication Group Working Papers, **9**, Research Laboratory of Electronics, MIT, Cambridge, MA, 28–34.

Chuang, E. S., and Hanson, H. M. (**1996**). ''Glottal characteristics of male speakers: Acoustic correlates and comparison with female data,'' J. Acoust. Soc. Am. **100**, 2657 (A).

Cranen, B., and Schroeter, J. (**1995**). ''Modeling a leaky glottis,'' J. Phon. **23,** 165–177.

Fant, G. (**1960**). *Acoustic Theory of Speech Production* (Mouton, The Hague).

Fant, G. (**1962**). ''Formant bandwidth data,'' Speech Trans. Lab. Q. Prog. Stat. Rep. **1**, Royal Institute of Technology, Stockholm, 1–3.

Fant, G. (**1972**). ''Vocal tract wall effects, losses, and resonance bandwidths,'' Speech Trans. Lab. Q. Prog. Stat. Rep. **2–3**, Royal Institute of Technology, Stockholm, 28–52.

Fant, G. (**1979**). ''Glottal source and excitation analysis,'' Speech Trans. Lab. Q. Prog. Stat. Rep. **1**, Royal Institute of Technology, Stockholm, 85–107.

Fant, G. (**1982**). ''The Voice source—Acoustic modeling,'' Speech Trans. Lab. Q. Prog. Stat. Rep. **4**, Royal Institute of Technology, Stockholm, 28–48.

Fant, G. (**1993**). ''Some problems in voice source analysis,'' Speech Commun. **13**, 7–22.

Fant, G., and Lin, Q. (**1988**). ''Frequency domain interpretation and derivation of glottal flow parameters,'' Speech Trans. Lab. Q. Prog. Stat. Rep. **2–3**, Royal Institute of Technology, Stockholm, 1–21.

Fant, G., Kruckenberg, A., Liljencrants, J., and Båvegård, M. (**1994**). ''Voice source parameters in continuous speech. Transformation of LF-

parameters,'' ICSLP 94 (Acoustical Society of Japan, Yokohama, Japan), 1451–1454.
Fant, G., Liljencrants, J., and Lin, Q. (**1985**). ''A four-parameter model of glottal flow,'' Speech Trans. Lab. Q. Prog. Stat. Rep. **4**, Royal Institute of Technology, Stockholm, 1–13.
Fujimura, O., and Lindqvist, J. (**1971**). ''Sweep-tone measurements of vocal-tract characteristics,'' J. Acoust. Soc. Am. **49,** 541–558.
Gauffin, J., and Sundberg, J. (**1989**). ''Spectral correlates of glottal voice source waveform characteristics,'' J. Speech Hear. Res. **32**, 556–565.
Gobl, C. (**1989**). ''A preliminary study of acoustic voice quality correlates,'' Speech Trans. Lab. Q. Prog. Stat. Rep. **4**, Royal Institute of Technology, Stockholm, 9–22.
Gobl, C., and Ní Chasaide, A. (**1988**). ''The effects of adjacent voiced/voiceless consonants on the vowel voice source: A cross language study,'' Speech Trans. Lab. Q. Prog. Stat. Rep. **2–3**, Royal Institute of Technology, Stockholm, 23–59.
Hanson, H. M. (**1995**). ''Glottal characteristics of female speakers,'' Ph.D. thesis, Harvard University, Cambridge, MA.
Hertegård, S., Gauffin, J., and Karlsson, I. (**1992**). ''Physiological correlates of the inverse filtered flow waveform,'' J. Voice **6**, 224–234.
Hillenbrand, J., Cleveland, R. A., and Erickson, R. L. (**1994**). ''Acoustic correlates of breathy vocal quality,'' J. Speech Hear. Res. **37**, 769–778.
Holmberg, E. B., Hillman, R. E., and Perkell, J. S. (**1988**). ''Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice,'' J. Acoust. Soc. Am. **84,** 511–529; Erratum **85,** 1787.
Holmberg, E. B., Hillman, R. E., and Perkell, J. S. (**1989**). ''Glottal airflow and transglottal air pressure measurements for male and female speakers in low, normal, and high pitch,'' J. Voice **4**, 294–305.
Holmberg, E. B., Hillman, R. E., Perkell, J. S., and Gress, C. (**1994a**). ''Relationships between intra-speaker variation in aerodynamic measures of voice production and variation in SPL across repeated recordings,'' J. Speech Hear. Res. **37**, 484–495.
Holmberg, E. B., Hillman, R. E., Perkell, J. S., Guiod, P., and Goldman, S. L. (**1995**). ''Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice,'' J. Speech Hear. Res. **38**, 1212–1223.
Holmberg, E. B., Perkell, J. S., Hillman, R. E., and Gress, C. (**1994b**). ''Individual variation in measures of voice,'' Phonetica **51**, 30–37.
House, A. S., and Stevens, K. N. (**1958**). ''Estimation of formant band widths from measurements of transient response of the vocal tract,'' J. Speech Hear. Res. **1**, 309–315.
Karlsson, I. (**1986**). ''Glottal wave forms for normal female speakers,'' J. Phon. **14**, 415–419.
Karlsson, I. (**1988**). ''Glottal waveform parameters for different speaker types,'' Speech '88: Proceedings 7th FASE Symposium, edited by W. A. Ainsworth and J. N. Holmes (Institute of Acoustics, Edinburgh), 225–231.
Karlsson, I. (**1989**). ''A female voice for a text-to-speech system,'' *Eurospeech 89*, edited by J. P. Tubach and J. J. Mariani (CEP Consultants, Edinburgh), Vol. 1, 349–352.
Karlsson, I. (**1990**). ''Voice source dynamics for female speakers,'' ICSLP 90 Proceedings (Acoustical Society of Japan, Kobe, Japan), 69–72.
Karlsson, I. (**1991a**). ''Dynamic voice quality variations in natural female speech,'' Speech Commun. **10**, 481–490.
Karlsson, I. (**1991b**). ''Female voices in speech synthesis,'' J. Phon. **19**, 111–120.
Karlsson, I. (**1992a**). ''Analysis and synthesis of different voices with emphasis on female speech,'' Ph.D. thesis, Royal Institute of Technology, Stockholm.
Karlsson, I. (**1992b**). ''Modelling voice variations in female speech synthesis,'' Speech Commun. **11**, 1–5.
Kasuya, H., Ogawa, S., Mashima, K., and Ebihara, S. (**1986**). ''Normalized noise energy as an acoustic measure to evaluate pathologic voice,'' J. Acoust. Soc. Am. **80,** 1329–1334.
Klatt, D., and Klatt, L. (**1990**). ''Analysis, synthesis, and perception of voice quality variations among female and male talkers,'' J. Acoust. Soc. Am. **87,** 820–857.
Klingholz, F. (**1987**). ''The measurement of the signal-to-noise ratio (SNR) in continuous speech,'' Speech Commun. **6**, 15–26.
de Krom, G. (**1993**). ''A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals,'' J. Speech Hear. Res. **36**, 254–266.
Ladefoged, P., and Antoñanzas-Barroso, N. (**1985**). ''Computer measures of breathy voice quality,'' Work. Papers Phon. **61**, University of California at Los Angeles, 79–86.
Laver, J. (**1980**). *The Phonetic Description of Voice Quality* (Cambridge U.P., Cambridge).
Linville, S. (**1992**). ''Glottal gap configurations in two age groups of women,'' J. Speech Hear. Res. **35**, 1209–1215.
Mori, K., Blaugrund, S. M., and Yu, J. D. (**1994**). ''The turbulent noise ratio: An estimation of noise power of the breathy voice using PARCOR analysis,'' Laryngoscope **104**, 153–158.
Ní Chasaide, A., and Gobl, C. (**1993**). ''Contextual variation of the vowel voice source as a function of adjacent consonants,'' Language Speech **36**, 303–330.
Peppard, R. C., Bless, D. M., and Milenkovic, P. (**1988**). ''Comparison of young adult singers and nonsingers with vocal nodules,'' J. Voice **2**, 250–260.
Perkell, J. S., Hillman, R. E., and Holmberg, E. B. (**1994**). ''Group differences in measures of voice production and revised values of maximum airflow declination rate,'' J. Acoust. Soc. Am. **96,** 695–698.
Shadle, C. (**1985**). ''The acoustics of fricative consonants,'' Technical Report 506, Research Laboratory of Electronics, MIT, Cambridge, MA.
Siebert, W. M. (**1986**). *Circuits, Signals, and Systems* (MIT, Cambridge).
Södersten, M., and Lindestad, P.-Å. (**1990**). ''Glottal closure and perceived breathiness during phonation in normally speaking subjects,'' J. Speech Hear. Res. **33**, 601–611.
Södersten, M., Lindestad, P.-Å., and Hammarberg, B. (**1991**). ''Vocal fold closure, perceived breathiness, and acoustic characteristics in normal adult speakers,'' in *Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms*, edited by J. Gauffin and B. Hammarberg (Singular, San Diego).
Stevens, K. N. (**1971**). ''Airflow and turbulence noise for fricative and stop consonants,'' J. Acoust. Soc. Am. **50,** 1180–1192.
Stevens, K. N. (**1993**). ''Models for the production and acoustics of stop consonants,'' Speech Commun. **13**, 367–375.
Stevens, K. N. (**to appear**). *Acoustic Phonetics* (MIT, Cambridge).
Stevens, K. N., and Hanson, H. M. (**1995**). ''Classification of glottal vibration from acoustic measurements,'' in *Vocal Fold Physiology: Voice Quality Control*, edited by O. Fujimura and M. Hirano (Singular, San Diego).
Tize, I. R. (**1989a**). ''A four-parameter model of the glottis and vocal fold contact area,'' Speech Commun. **8**, 191–201.
Titze, I. R. (**1989b**). ''Physiologic and acoustic differences between male and female voices,'' J. Acoust. Soc. Am. **85,** 1699–1707.
Titze, I. R. (**1995**). ''Definitions and nomenclature related to voice quality,'' in *Vocal Fold Physiology: Voice Quality Control*, edited by O. Fujimura and M. Hirano (Singular, San Diego).
Yumoto, E., Gould, W. J., and Baer, T. (**1982**). ''Harmonics-to-noise ratio as an index of the degree of hoarseness,'' J. Acoust. Soc. Am. **71,** 1544–1550.
Zemlin, W. R. (**1988**). *Speech and Hearing Science: Anatomy and Physiology* (Prentice-Hall, Englewood Cliffs, NJ).

Automatic integrated analysis of jaw and lip movement in speech production

Nunzio Alberto Borghese
Istituto Neuroscienze Bioimmagini—C.N.R., Via Mario Bianco 9, 20131 Milano, Italy

Giancarlo Ferrigno, Marco Redolfi, and Antonio Pedotti
Centro di Bioingegneria, Politecnico di Milano, Fondazione ProJuventute, Pzza Leonardo da Vinci 32, 20133 Milano, Italy

(Received 16 January 1996; accepted for publication 29 July 1996)

A noninvasive methodology for studying the kinematics of speech production is presented. It is based on the tracking of very small and light passive markers attached to the subjects' face. Using a pair of TV cameras, the 3-D markers' positions are computed in real time, at a subpixel accuracy, by a dedicated hardware. From these data, the time course of a set of parameters which describe lip and jaw movement is computed; in addition, a semiautomatic procedure that identifies the exact onset and offset of the investigated sequences has been developed. To compare the results over different productions, a time normalization procedure based on a continuous inverse Fourier transform has been implemented.
[S0001-4966(97)05512-4]

PACS numbers: 43.70.Jt, 43.70.Aj [AL]

INTRODUCTION

In this work, we report on a noninvasive methodology developed to obtain reliable data on lips and jaw kinematics in speech production. In the past, different approaches have been followed. The first system, developed in the 1970s, was based on strain gauge transducers (Sussman and Smith, 1970) and was later improved in the interface with the subjects' face by Barlow and co-workers (Barlow *et al.*, 1983). A less intrusive system was introduced to the field in the early eighties by the group of Kelso (Kelso *et al.*, 1984): one TV camera detected the light emitted by LEDs attached to subjects' face. Other approaches based on magnetic transducers (Perkell *et al.*, 1992; Schönle, 1988) or ultrasounds (Sonies *et al.*, 1981) are less reliable and accurate (Burdea and Coeffet, 1994). The drawback of all these systems is the need to attach electromechanical devices to the subjects' face which, although small, do interfere with jaw and lips movements and do not allow the recording of a true free motion (Ladin, 1995). A different approach is the detection of a set of anatomical repere points directly on a stream of TV images (e.g., lips boundaries, chin, wrinkles; Essa *et al.*, 1994). Although, in principle, this approach is interesting to the field, the features cannot be identified with the high accuracy and reliability required; moreover, the sampling rate does not exceed 30 Hz, which is not sufficient for a detailed analysis.

A good compromise is the use of small, light, passive markers fixed on the subjects' repere points which can be detected by a dedicated system. In particular, the system proposed by Ferrigno and Pedotti (Ferrigno and Pedotti, 1985, 1987) allows one to detect and to reconstruct in real time, at 100 Hz, the 3-D position of hemispherical passive markers, 1 mm in diameter and 0.3 mg in weight, at a subpixel accuracy, thanks to a cross correlation between the TV images and a marker template implemented with VLSI chips. From the markers' positions, a set of 3-D parameters which decompose lip motion into opening/closing, rounding, and protrusion is computed (e.g., Bell-Berti and Harris, 1979; Ladefoged and Maddison, 1990; Lubker and Gay, 1982; Perkell, 1986). Moreover, the identification of the utterances onset and offset time (segmentation) and the statistical comparison of the different productions of the same utterance are required.

The methodology presented here is of general interest. In particular, it can be extremely useful in clinical laboratories for the quantitative diagnosis of speech motor disorders and for rehabilitation (e.g., Gracco, 1992). In computer science, it can produce a detailed knowledge on the jaw and lips kinematics in speech which can assist in the construction of artificial machines that can produce and understand speech movements (Silsbee and Bovik, 1993; Brooke and Summerfield, 1983).

I. METHODOLOGY AND RESULTS

A. Experimental setup and acquisition system

The subject comfortably sits at about 1.5 m from a pair of CCD TV cameras placed one upon the other, 1.5 m apart, with the best overlapping/accuracy trade-off (Borghese and Ferrigno, 1990). The cameras are equipped with a ring of LEDs flashing in the infrared bandwidth, creating no disturbance to the subjects, and with an infrared filter which eliminates all the light coming from the environment outside the infrared bandwidth, increasing the signal-to-noise ratio. A standard microphone is located in front of the subject, just below the chin (to avoid interference with the TV cameras' view). Eight markers are attached to the subjects' skin at the following repere points (Fig. 1): tip of the nose (1), middle of the upper (2) and of the lower (5) lip, right (4) and left corner (3) of the mouth, middle of the chin (6), lobe of the left (7) and of the right (8) ear. Markers 2–5 identify jaw and lips

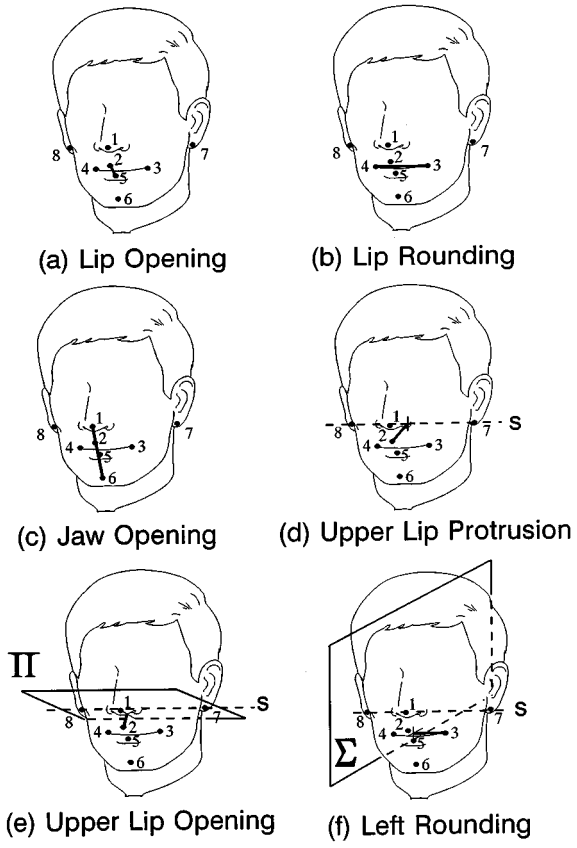


FIG. 1. The geometrical description of the analyzed kinematics parameters.

geometry, while markers 1, 7, and 8 are used to construct a coordinate reference system solid with the subject's head (cf. Sec. II A).

During speech production, the markers' coordinates and the acoustic signal are collected synchronously by a host computer. The audio signal is sent to a preamplifier which scales it to the voltage required by the acquisition board (± 1 V). It is then sampled at 16 KHz with a resolution of 12 bits. The video signal is sent to a dedicated hardware system which extracts in real time the 2-D coordinates of the markers and sends them to the host. The core of this system is two pipelined hardware processors which sequentially scan the TV image, 256×256 pixels wide, and cross correlate it with a hardware template of the markers, 6×6 pixels wide, at 100 Hz, with a rate of ≈ 15 μs per correlation ($\approx$800 Mflops).

By this operation, the cross correlation gets higher values the closer the pixel is to the center of a marker, and those pixels which belong to a marker can be extracted from the rest of the image by hardware thresholding the cross-correlation function. Moreover, computing the markers position as the average of its constituting pixels coordinates, each weighted with its corresponding cross-correlation value, increases the accuracy up to 0.1 pixels (Ferrigno and Pedotti, 1985, 1987). Another advantage offered by this kind of image processing is the possibility to use very small (1 mm of diameter) and light (about 0.3 mg) markers which, in our experiments, were made of a plastic support covered by retroreflective paper. Subjects do not feel the markers and they often have to rely on a mirror to take them off.

With the passive markers approach, the modification of the markers' shape on the TV cameras is a potential problem. For example, during the production of /u/, if the markers on the mouth corners (Fig. 1) are too internal, they rotate inward, and, if they are partially hidden by the lips, they cannot be detected anymore by the system. This is carefully checked out before the acquisition session: The subject is asked to perform maximum opening, rounding, and protrusion of the mouth while the original (not processed) images picked up by the TV cameras are analyzed on a monitor to verify that the markers are always detected.

The identification of the markers in the 2-D representations in each TV camera is automatically carried out taking into account the relative 2-D positions of the markers (e.g., the markers on the earlobes are, respectively, the rightmost and leftmost ones). This procedure is based on the fact that the subject's position, with respect to the cameras, is approximately known. The 3-D reconstruction is carried out by the host computer through a fast triangulation procedure with a 3-D experimental accuracy of 0.1 mm which is achieved throughout a high-precision calibration procedure which effectively corrects for systematic errors (distortions). The assessment of the accuracy has been carried out through the standard deviation of the measured 3-D distance between two markers attached on the extremities of a rigid bar surveyed at moderate speed inside the field of view (Borghese and Ferrigno, 1990).

B. Segmentation and interval identification

The exact time boundaries associated with a particular phonetic sequence are identified through the following semiautomatic procedure (Fig. 2) inspired by the work of Rabiner and Sanbur (1975). The choice of segmenting speech movement through the acoustic trace is suggested by the fact that the transition between phonemes can become fuzzy in the kinematic domain. The procedure will be described along with its application to the segmentation of a /tiltrù/ sequence. It is based on the analysis of a low-pass filtered version of the amplitude of the acoustic signal, $A(n)$, which is obtained as the moving average over a 1.31-ms sliding window:

$$A(n)=1/21\sum_{k=-10}^{+10}|S(n+k)|. \tag{1}$$

$A(n)$ filters out spikes accidentally recorded, which can be erroneously identified as segment boundaries. The noise mean amplitude, A_N, is estimated averaging $A(n)$ over a 100-ms period (1600 samples), where the subject was silent.

In the first step of the procedure, a gross identification of the interval boundaries is carried out manually [interval $Io[a1\ a2]$ in Fig. 2(a)] to isolate the sequence of interest from the carrier phrase. Inside Io, the mean value of $A(n)$, A_S, is computed as

$$\overline{A_S}=\frac{1}{a_2-a_1+1}\sum_{n=a_1}^{a_2}A(n) \tag{2}$$

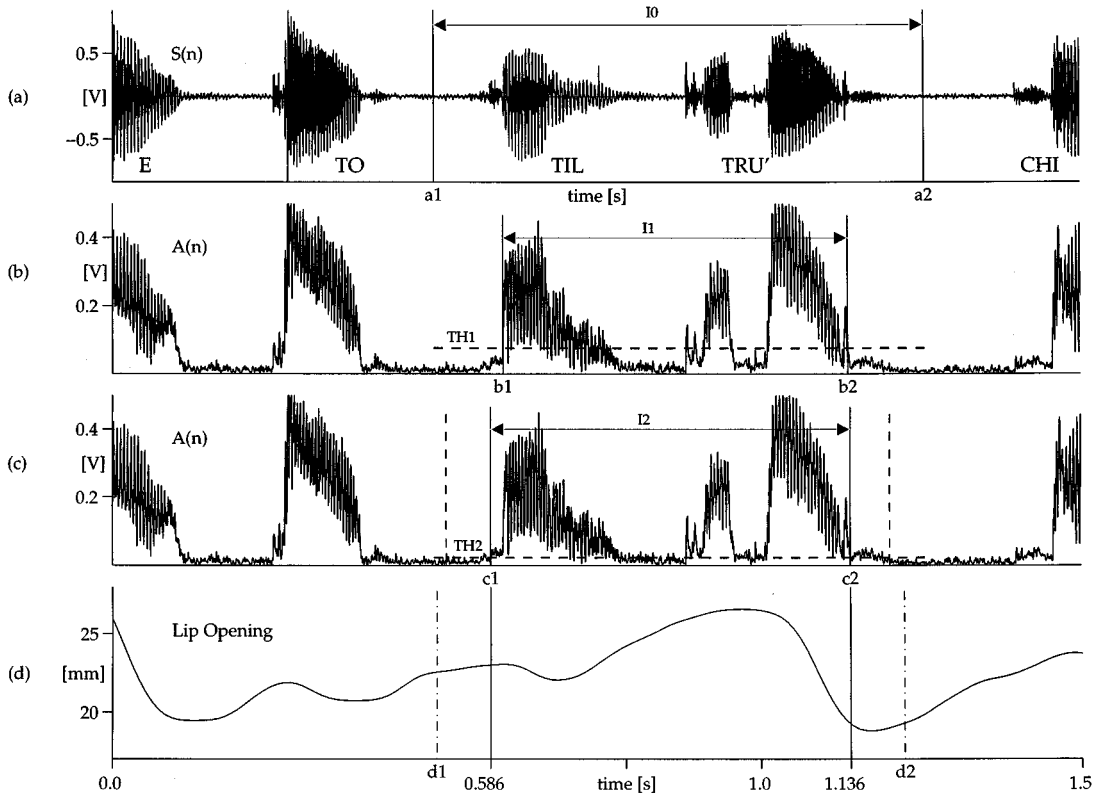


FIG. 2. The sequential steps of the segmentation procedure are reported from top to bottom (a)–(d) for the location of the boundaries of a \iltrù\ sequence. In (a), a first time interval, Io, containing the sequence is manually identified on the signal time course $S(n)$. In (b), the moving average, $A(n)$, of the acoustic signal, is plotted along with a first threshold, TH1, which is equal to 50% of the mean value of $A(n)$ computed inside $Io[a1\ a2]$. The time in which $A(n)$ goes above TH1 are the extremes of the interval $I_1[b1\ b2]$. In (c), a second threshold, TH2, which is equal to twice the noise mean amplitude, is reported. The time for which $A(n)$ goes below TH2 represent the true boundaries of the time interval associated to the analyzed sequence, $I_2[c1\ c2]$. The search for the boundaries using only TH2 would have produced wrong boundaries, which are plotted as vertical dashed lines. Finally, thanks to the simultaneous recording of kinematics and acoustics, these time boundaries are reported on the time course of the parameters [(d), segmentation of the parameter lip opening]. For subsequent normalization, the interval $I_S[d1\ d]$, 30% larger than I_2, is considered.

and a first threshold, TH_1, is set to 50% of $\overline{A_S}$ [Fig. 2(b)]. Going from the extremes of Io inward, the instants of time $[b1\ b2]$ in which $A(n)$ goes over TH_1 are detected. $I_1[b1\ b2]$ represents a new interval which, although closer to the real interval, neither contains the first /t/ nor the very last portion of the /u/. To get to the real boundaries, a second threshold, $TH_2=2A_N$, is considered [Fig. 2(c)]. Going from the extremes of the interval I_1 outward, the instants of time $[c1\ c2]$ in which $A(n)$ goes below TH_2 are detected. $I_2[c1\ c2]$ represents the true time boundaries associated to /tiltrù/. Using only TH_2 and starting from $[a1\ a2]$, false boundaries (reported as dashed vertical lines) would have been detected. At the end, thanks to the simultaneous recording of kinematics and acoustics, these temporal markers are reported on the time course of the kinematics parameters [on lip opening in Fig. 2(d)].

Although the thresholds are somehow arbitrary, variations of TH_1 of more than 100% do not shift the boundaries of the interval Io. TH_2 is more critical: the interindividual intensity of the stop consonants with respect to the background noise may require some tuning of TH_2 according to the particular speaker. Nevertheless, when the speech production results in a clear acoustic pattern, as in the laboratory environment, and the boundaries do not coincide with unvoiced–voiced boundaries, as in our study, the procedure gives the optimal result most of the time without adjustments for the level of TH_2.

C. Signal modeling and filtering

Filtering of the kinematics data is carried out in the frequency domain by multiplying the fast Fourier transform (FFT) of the parameter time course by the FFT of the ideal low-pass filter (sinc function):

$$H(n)=\frac{\sin(f_{\text{cutoff}}/f_s n)}{\pi n}. \tag{3}$$

To get a limited number of samples from $H(n)$ a polynomial windowing function is applied to the filter. To achieve the optimal filtering, the cutoff frequency, f_{cutoff}, is automatically determined for each parameter from its power density spectrum (PSD_p). This is set as the frequency at which the signal-to-noise ratio drops under a predefined threshold, TH_{cutoff}, experimentally set to 50:

$$f_{\text{cutoff}}: \quad \text{PDS}_p|_{f_{\text{cutoff}}}=\text{TH}_{\text{cutoff}}\text{PSD}_N. \tag{4}$$

PSD_N is the mean value of the noise power spectrum density which, under the hypothesis of noise being additive, white, zero mean, like in the Elite system (Ferrigno *et al.*, 1992), is equal at all the frequencies. Its value can be advan-

tageously estimated in the highest frequency region (in our case between $0.8*f_s/2$ and $f_s/2$, where $f_s=100$ Hz is the sampling frequency), whenever the signal contribution is small with respect to the noise. To avoid distortions at the borders, PSD_p is computed from the Zeta transform of an autoregressive model estimated from the parameters time course (D'Amico and Ferrigno, 1990). For the speech productions that we have analyzed, the frequency content of the markers' movement was maximal in the vertical direction, and in particular for the marker on the lower lip which reached 12.80 Hz. Whenever the signal content of an utterance would exceed 40 Hz, a standard FIR filter should be employed. This filtering scheme is particularly suitable for an adaptive implementation where different cutoff frequencies can be automatically computed for different speech segments. To obtain the derivatives of the parameters, their filtered FFT is multiplied by the derivative operators: $j\omega$ (to obtain the velocities) and $-\omega^2$ (to obtain the accelerations); and then antitransformed through the IFFT (inverse fast Fourier transform). The derivatives time course will not be reported for the sake of brevity.

D. Time normalization

As the velocity of speech changes in the different productions of the same utterance from the same subject or from different subjects, to compare different time courses of the same parameter, a time normalization of the data is required. This is a procedure which receives as input an arbitrary number of time samples (N) and outputs a fixed number of samples (N_1), not necessarily synchronous with the input ones; this implies that the time interval from two consecutive samples may differ from one production to the other depending on the actual speed of speech production. In this work, time normalization is achieved through sampling the CIFT (continuous inverse Fourier transform), computed from the filtered FFT of the kinematics parameters, into a fixed number of samples, equally spaced in time. In particular, the N input samples are all the samples contained in the interval $Is[d1\ d2]$ [cf. Fig. 2(d)] which is 30% larger than the actual interval, I_2, associated to the utterance under investigation to take into account perseveratory and anticipatory effects. The number of output samples, N_1, is set to 128 by the following considerations. The maximum value of Is was 930 ms, corresponding to 93 kinematic samples ($N=93$); the FFT is most efficient when the number of samples is equal to a power of 2.

II. EXEMPLARY DATA

A. Parameters considered

To describe speech kinematics, a simplified geometrical model is usually adopted (e.g., Bell-Berti and Harris, 1979; Ladefoged and Maddison, 1990; Lubker and Gay, 1982; Perkell, 1986) which is based on the following parameters (Fig. 1): lip opening, lip rounding, and lip protrusion. To analyze movement asymmetries, four additional parameters have been defined: upper lip opening, lower lip opening, right emi-rounding, and left emi-rounding. An additional parameter, jaw opening, was introduced to study the coordination between the lips and the jaw.

The parameters are defined as follows:

Lip opening is computed as the 3-D distance between marker 2 (upper lip) and marker 5 (lower lip), Fig. 1(a).

Lip rounding is computed as the 3-D distance between marker 3 (lips left corner) and marker 4 (lips right corner), Fig. 1(b).

Jaw opening is computed as the 3-D distance between marker 1 (tip of the nose) and marker 6 (chin), Fig. 1(c).

To compute the other parameters a reference frame is constructed. It is constituted of a straight line, s, and two planes, Π and Σ [Fig. 1(d)–(f)]. The line s connects markers 7 and 8 [earlobes; see Fig. 1(d)]; the plane Π is defined as the plane containing the line s and marker 1, and it is approximately coincident with the Camper plane [Fig. 1(e)]. The plane Σ, which is a plane of vertical symmetry, contains marker 1 and it is orthogonal to s.

Upper (lower) lip protrusion is computed as the 3-D distance between marker 2 (marker 5) and the line s, Fig. 1(d).

Upper (lower) lip opening are computed as the 3-D distance between marker 2 (marker 5) and the plane Π, Fig. 1(e).

Lip left (right) emi-rounding is computed as the 3-D distance between the marker 3 (marker 4) and the plane Σ, Fig. 1(f).

B. Results

In order to validate the system, preliminary results on a phonetic study are reported and discussed. The study was aimed to assess the coarticulation rules in Italian language (Magno Caldognetto *et al.*, 1992): six native Italian speakers had to repeat six times, in random order, each of the selected nonsense sequences /titu/, /tilù/, /tiltù/, /tiltrù/, /tisù/, /tistù/, and /tistrù/, within the carrier phrase: ''Ripeto...chiaramente'' (I repeat...clearly). A clear picture of lips and jaw coordination can be gained from Fig. 3 where the time course of all the parameters and of the acoustic signal is reported for the sequence /tiltrù/. The time boundaries of the sequence, which were identified through the procedure outlined in Sec. I B, are reported as vertical lines which segment the acoustic and parameters traces.

In this particular sequence, the variations in amplitude of the parameters are very small throughout the movement, in the range of a few millimeters. Lip movement on the horizontal plane is highly symmetrical both in the rounding [Fig. 3(a) and (b)] and in the protrusion [Fig. 3(c) and (d)] dimensions. During the production of /til/, the lip corners remain close to their resting position; this corresponds to the minimum rounding of the mouth [maximum of the parameter rounding, Fig. 3(g), cf. Fig. 1(b) and (f)]. Afterward, the lip corners are brought close to each other for the production of the vowel /u/ and then back to their resting position at the end of the production. The same temporal pattern is true for lip protrusion [Fig. 3(c) and (d)]: The lips move outward synchronously to produce the vowel /u/ and go back to their resting position afterward. On the contrary, the symmetry of

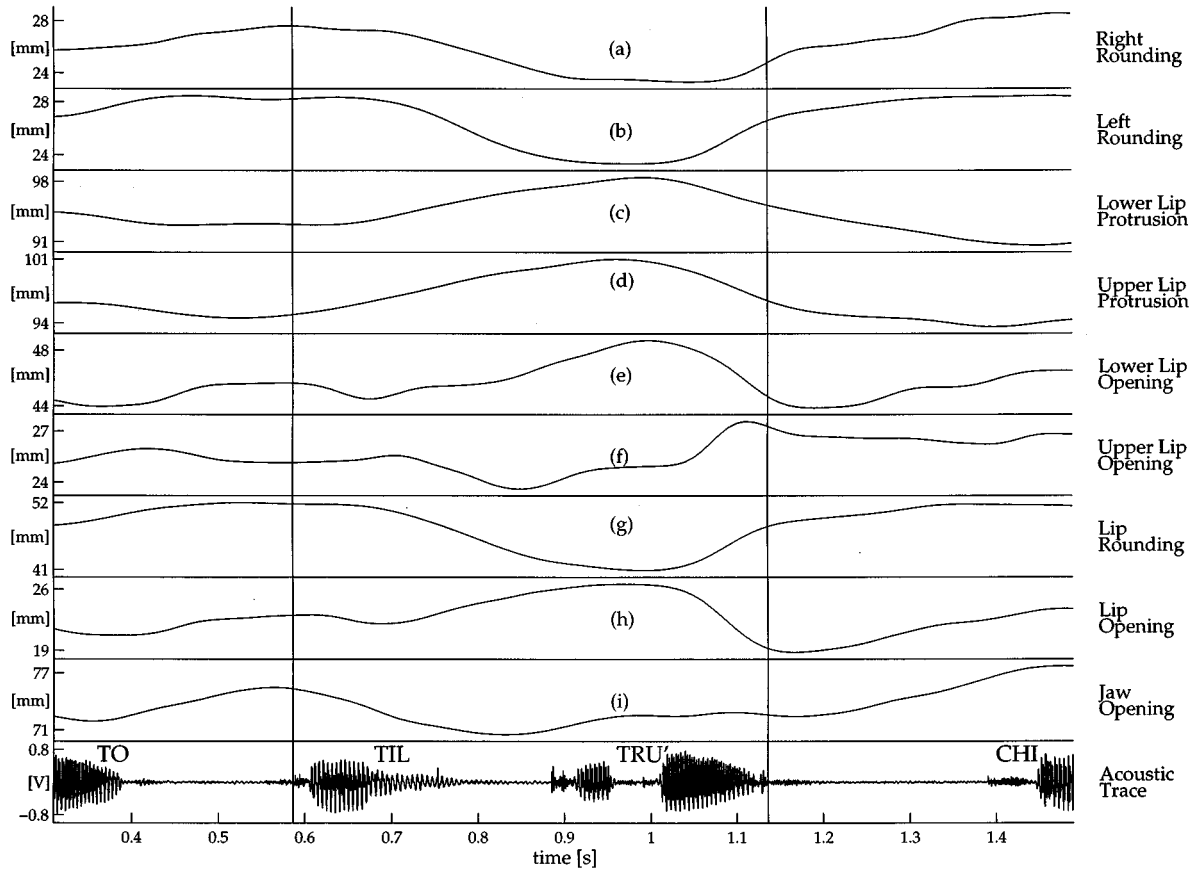


FIG. 3. The time course of all the parameters and of the acoustic trace is reported for a production of \tiltrù\. Maximum and minimum for each parameter is reported as tick markers on the vertical axes. The vertical lines bound the time interval associated to \tiltrù\.

the movement on the vertical plane is not preserved here: There is a difference in the time course of lower [Fig. 3(e)] and upper lip [Fig. 3(f)] opening for the production of /trù/. The opening of the lower lip is due not only to the lip movement per se, but also by the rising of the jaw [Fig. 3(i)]. To achieve the correct configuration for the production of /u/, this should be taken into account and it suggests a differential control of lip opening during the production of the intervocalic consonants (/ltr/ in this case) which precede the final /u/.

The comparative analysis of the time course of lip protrusion [Fig. 3(c) and (d)], lip rounding [Fig. 3(g)], and lip opening [Fig. 3(h)] shows that they do not have a similar time course throughout the entire phonetic production. For example, in the production of the vowel /u/, the lips begin to protrude in correspondence of the syllable /til/, well before the opening and rounding start; and they go back to their resting position much slower in the protrusion and rounding dimensions than in the opening one. The difference in the time course of lip rounding [Fig. 3(g)] and lip protrusion [Fig. 3(c) and (d)] is also evident in the production of the syllable /til/ where the mouth corners are set apart (low value for lip rounding) with very little opening of the mouth (small increase in lip opening). The same observations apply also to jaw kinematics which is described by the parameter jaw opening [Fig. 3(i)]. Its time course is similar to that of lip opening [Fig. 3(h)] in some productions, like in the syllable /chi/ in ''chiaramente'' [right most part of Fig. 3(h) and (i)], where the lips and the jaw open synchronously; but in general its time course is different from that of the lip parameters. For example, in the production of the syllable /trù/, while all the lip parameters exhibit their maximal variation to produce the final vowel /u/, jaw opening slightly increases. This may suggest that, for this phoneme, the jaw constitutes a stable platform with respect to which the lips movement is controlled.

III. CONCLUSION

A nonintrusive methodology for studying the kinematics of speech production has been presented. It is based on the detection of very small and light passive markers attached to a subject's face. The movement is filmed by a pair of TV cameras and the 3-D marker position is computed in real time, at a subpixel accuracy, by a dedicated hardware. From these data, the time course of a set of 3-D parameters which describe lips and jaw movement is reconstructed. These allow one to identify the differential contribution of jaw and lips to the motion, and the kinematics characteristics of each phoneme as well as of their concatenation in speech production. To determine the exact time boundaries of the investigated sequences a semiautomatic procedure which avoids manual zooming in and out the audio trace has been described. To analyze statistically the parameters' time course of the same utterances over different productions of the same

and of different subjects, a normalization procedure based on a continuous inverse Fourier transform has been implemented.

ACKNOWLEDGMENTS

The authors wish to thank Dr. E. Magno-Caldognetto and Dr. K. Vagges for the helpful and stimulating discussions, and Luciano Chiumiento for skilled artwork.

Barlow, S. M., Cole, K. J., and Abbs, J. K. (**1983**). ''A new head mounted lip-jaw movement transduction system for the study of motor speech disorders,'' J. Speech Hear. Res. **26**, 283–288.

Bell-Berti, F., and Harris, K. S. (**1979**). ''Anticipatory coarticulation: some implications from a study of lip rounding,'' J. Acoust. Soc. Am. **65,** 1268–1270.

Borghese, N. A., and Ferrigno, G. (**1990**). ''An algorithm for 3D Automatic Movement Detection by Means of Standard TV cameras,'' IEEE Trans. Biomed. Eng. **37**(12), 1221–1225.

Brooke, N. M., and Summerfield, Q. (**1983**). ''Analysis, Synthesis and Perception of Visible Articulatory Movements,'' J. Phon. **11**, 63–76.

Burdea, G., and Coeffet, P. (**1994**). *Virtual Reality Technology* (Wiley, New York), pp. 15–45.

D'Amico, M., and Ferrigno, G. (**1990**). ''Technique for the evaluation of derivatives from noisy biomechanical displacement data using a model based bandwidth-selection procedure,'' Med. Biol. Eng. Comp. **28**, 407–415.

Essa, I. A., Darrell, T., and Pentland, A. (**1994**). ''Tracking facial motion,'' in *Proceedings of the IEEE Workshop on Non-rigid and Articulate Motion*, Austin, Texas (IEEE, New York).

Ferrigno, G., and Pedotti, A. (**1985**). ''ELITE: a digital dedicated hardware system for movement analysis via real time TV signal processing,'' IEEE Trans. Biomed. Eng. **32,** 943–950.

Ferrigno, G., and Pedotti, A. (**1987**). ''Modularly expansible system for real-time processing of a TV display, useful in particular for the acquisition of coordinates of known shapes objects,'' US Patent No. 4,706,296.

Ferrigno, G., D'Amico, M., and Santambrogio G. C. (**1992**). ''Frequency Domain Implications of Time Multiplexing in Kinematic Data Capture,'' in *Proceedings of Medicon92* (Area di Ricerca-CNR, Pisa, Italy), pp. 1093–1096.

Gracco, V. L. (**1992**). ''Analysis of Speech Movements: Practical Considerations and Clinical Application,'' Haskins Laboratories Status Report On Speech Research, SR-109/110, pp. 45–58.

Kelso, J. A. S., Tuller, B., Fowler, C. A., and Vatikotis-Bateson, E. (**1984**). ''Functionally specific articulatory co-operation following jaw perturbations during speech: evidence for coordinative structures,'' J. Exp. Psychol. **10**(6), 812–832.

Ladefoged, P., and Maddison, I. (**1990**). ''Vowels of the world's languages,'' J. Phon. **18**, 93–122.

Ladin, Z. (**1995**). ''Three-dimensional instrumentation,'' in *Three-Dimensional Analysis of Human Movement*, edited by P. Allard, I. A. Stokes, and J. P. Blanchi (Human Kinetics, Champaign, IL), pp. 3–18.

Lubker, J., and Gay, T. (**1982**). ''Anticipatory labial coarticulation: experimental, biological and linguistic variables,'' J. Acoust. Soc. Am. **71,** 437–448.

Magno Caldognetto, E., Vagges, K., Ferrigno, G., and Busà, M. G. (**1992**). ''Lip rounding coarticulation in Italian,'' in *Proceeding of International Conference on Spoken Language Processing*, Banff (Priority Press, Edmonton, Canada), Vol. 1, pp. 61–64.

Perkell, J. S. (**1986**). ''Coarticulation strategies, preliminary implications of detailed analysis of lower lip protrusion movements,'' Speech Commun. **5**, 47–68.

Perkell, J. S., Cohen, M., Svirsky, M., Mathhies, M., Garabieta, I., and Jackson, M. (**1992**). ''Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements,'' J. Acoust. Soc. Am. **92,** 3078–3096.

Rabiner, L. R., and Sambur, M. R. (**1975**). ''An algorithm for determining the endpoints of isolated utterances,'' Bell Syst. Tech. J. **54**(2), 297–315.

Schönle, P. (**1988**). *Elektromagnesische Artikulographie* (Springer, Berlin).

Silsbee, P. L., and Bovik, A. C. (**1993**). ''Medium Vocabulary Audio Visual Speech Recognition,'' in *New Advances and Trends in Speech Recognition and Coding* (NATO ASI, Bubion), pp. 13–16.

Sonies, B. C., Shawker, T., Hall, T. E., Gerber, L. H., and Leighton, S. B. (**1981**). ''Ultrasonic visualization of tongue motion during speech,'' J. Acoust. Soc. Am. **70,** 683–686.

Sussman, H. M., and Smith, K. U. (**1970**). ''Transducer for measuring lip movements during speech,'' J. Acoust. Soc. Am. **48,** 858–860.

Perception of nonlinear and linear formant trajectories

Anna K. Nábělek and Alexandra Ovchinnikov
The University of Tennessee, Knoxville, Tennessee 37996-0740

(Received 2 August 1995; revised 30 May 1996; accepted 20 August 1996)

Perception of $F2$ trajectories in synthetic vowels was investigated. Perceptual boundaries (50% response points of identification functions) of 20-step /ʊ–ɪ/ continua with various shapes of $F2$ trajectories were determined and compared with the boundary for vowels with steady-state $F2$. In experiment 1, the vowels were synthesized in either /j–j/ or /w–w/ context resulting in parabolic trajectories, then the stimuli were split into halves resulting in quadratic $F2$ trajectories. All stimuli were 200 ms long. For the /wVw/, and /wV/ stimuli, the boundaries were at lower stimulus numbers than for the stimuli with steady-state $F2$, indicating that the nonlinear $F2$ trajectories were perceived as having frequencies beyond the extreme values actually synthesized in the stimuli. This type of signal processing has been termed ''perceptual compensation.'' For the /jVj/ stimuli, there was only a trend for perceptual compensation of $F2$ trajectories. For the /Vw/, /Vj/, and /jV/ stimuli, the boundaries were at stimulus numbers corresponding to frequencies of the relatively steady-state vowel-like segments. In experiment 2, the quadratic $F2$ trajectories of the /wV/ stimuli were changed to linear trajectories and $F1$ and $F3$ trajectories were either quadratic, linear, or steady state. The results indicated that the shape of $F1$ and $F3$ trajectories had no effect on the boundaries. For the linear $F2$ trajectories, as for the quadratic $F2$ trajectories in the /wV/ stimuli, the boundaries were at lower stimulus numbers than for the stimulus with steady-state $F2$, indicating presence of perceptual compensation. In the experiment 3, the $F1$ and $F3$ were steady state and the linear $F2$ trajectories had three different values of frequency difference, ΔF, between the initial and final frequencies. The perceptual compensation was found for the 200-ms stimuli with large ΔF up to 500 Hz over 200-ms stimulus duration, and perceptual enhancement of final frequencies was found for the stimuli with $\Delta F = 280$ Hz. © *1997 Acoustical Society of America.*
[S0001-4966(97)01001-1]

PACS numbers: 43.71.An, 43.71.Es, 43.66.Mk [RAF]

INTRODUCTION

Effects of temporal changes of vowel trajectories on categorization of stimuli in synthetic vowel continua have been investigated in several studies. DiBenedetto (1989); Nábělek *et al.* (1993); and Nábělek and Ovchinnikov (1994) studied $F1$ and $F2$ trajectories changing linearly over time. The results of these studies seemed to indicate that linear trajectories were perceptually averaged over time. Nonlinear vowel trajectories in synthetic vowel continua with varied $F2$ were studied by Lindblom and Studdert-Kennedy (1967) and by Nearey (1989). The results of these studies indicated that categorical boundaries were different than for continua with steady-state $F2$ as if the perceived frequencies of the nonlinear $F2$ trajectories were beyond the extreme values actually synthesized in the stimuli. This type of signal processing has been termed ''perceptual compensation'' or ''perceptual overshoot'' which might serve to offset the undershoot effects of production. In natural connected speech, vowel formant frequencies usually do not reach target values that could be reached when a talker produces sustained, monophthonged vowels. It has been hypothesized that listeners may compensate for this undershoot in articulation by ''overshooting'' the target in perception (Lindblom and Studdert-Kennedy, 1967).

Van Son (1993) has discussed ''overshoot'' versus temporal averaging theories and has come to the conclusion that both types of perception of formant trajectories can be found depending on the stimuli involved. In the Lindblom and Studdert-Kennedy (1967) study, the transitory characters of $F2$ (and $F3$) were achieved by presenting vowels in /j–j/ and /w–w/ contexts. Perceptual boundaries (50% response points of identification functions) of /ʊ–ɪ/ continua were determined. In relation to the boundary for the isolated vowels, the boundaries for the CVC stimuli were shifted. The $F2$ frequency of vowels in a /w–w/ context was perceived as higher than the maximum value of a convex upward parabola. The frequency of $F2$ of vowels in a /j–j/ context was perceived as lower than the minimum value of a concave downward parabola. Similar ''perceptual overshoot'' of $F2$ was demonstrated by Nearey (1989) for vowels in /d–d/ and /b–b/ contexts. Van Son (1993) synthesized stimuli with $F1$ and $F2$ trajectories forming either convex or concave parabolas. The stimuli were flanked by consonants /n/ or /f/, but coarticulation between consonants and vowels was deliberately not modeled. Thus according to van Son, the consonants were not integrated with vowel trajectories but were just attached to them. Subjects identified these stimuli in an open response paradigm. The results indicated that the parabolic formant trajectories were averaged over time. Van Son (1993) concluded that the perceptual compensation (''overshoot'') was related to coarticulation between consonantal and vowel segments in Lindblom and Studdert-Kennedy (1967) and Nearey (1989) stimuli. However, the difference in responses to his and others' stimuli could have been

caused by the method of data collection. While van Son required phonetic identification in an open set response, in the other two studies, the stimuli were categorized as belonging to vowels corresponding to end-point stimuli in the tested continua.

In the studies by Lindblom and Studdert-Kennedy (1967); Nearey (1989); and Nábělek and Ovchinnikov (1994), subjects categorized vowels from continua in which $F2$ was varied in steps. In the Nábělek and Ovchinnikov (1994) study the $F2$ trajectories, which changed linearly over time in either upward or downward directions, were averaged over time. In the two other studies, in which the vowel segments were synthesized in C-C contexts and the nonlinear trajectories had either maxima or minima in the middle, the trajectories were perceptually compensated (''overshoot''). The goal of the present study was to determine (1) if, in categorization experiments, the perceptual compensation is restricted to the CVC stimuli or if can be also observed for the CV and VC stimuli with quadratic $F2$ trajectories which change only in one direction, and (2) what other modifications of $F2$ trajectory are needed to demonstrate temporal averaging as in the Nábělek and Ovchinnikov (1994) experiment.

To answer the above questions, several types of stimuli were tested. The boundaries of continua with stimuli having $F2$ trajectories changing over time were compared with the boundary of a continuum with isolated vowels having steady-state $F2$ trajectories. In experiment 1, /ʊ–ɪ/ continua as used by Lindblom and Studdert-Kennedy (1967) were synthesized in /w–w/ and /j–j/ contexts. Then, continua were created which corresponded to first and second halves of the CVC stimuli (/wV/, /Vw/, /jV/, and /Vj/). In experiment 2, continua were created in which quadratic trajectories of $F1$, $F2$, and $F3$ corresponding to the /wV/ stimuli, were modified. In all modified continua, $F2$ was linearly changing in an upward direction while $F1$ and $F3$ trajectories were quadratic, linear, or steady state. In experiment 3, three /ʊ–ɪ/ continua were tested in which $F2$ was changing linearly in an upward direction with three different values of frequency difference between the initial and final frequencies.

I. EXPERIMENT 1

A. Methods

1. Test materials

Seven 20-step /ʊ–ɪ/ continua were generated with a Klatt synthesizer (Klatt, 1980). In the continua, vowels were in isolation or were combined with consonants as follows: (1) vowels in isolation, #V#; (2) in a /j–j/ context, /jVj/; (3) in a /w–w/ context, /wVw/; (4) preceded by /j/, /jV/; (5) followed by /j/, /Vj/, (6) preceded by /w/, /wV/; and (7) followed by /w/, /Vw/. The Klatt synthesizer was configured in the cascade mode using five formants. Each formant resonance was specified by two parameters: frequency (F) and bandwidth (BW).

The selected formant frequency values for vowels in the #V# stimuli and for vowels and consonant loci in the /jVj/ and /wVw/ stimuli were the same as used by Lindblom and Studdert-Kennedy (1967). For the vowels, the frequency of $F1$ was 350 Hz. The values of $F2$ varied from 1000 to 2000 Hz in equal steps of about 53 Hz. The values of $F3$ varied from 2300 to 2823 Hz in equal steps of about 28 Hz. In all stimuli BW1=60 Hz, BW2=100 Hz, and BW3=110 Hz. The remaining formants and bandwidth values were: $F4$ =3300 Hz, BW4=250 Hz; $F5$=3850 Hz, BW5=300 Hz. A fundamental frequency ($F0$) contour was specified as falling linearly from 130 Hz at the onset to 100 Hz at the offset of the stimulus.

In the consonants, the loci for /j/ were: $F1$=250 Hz, $F2$=2200 Hz, and $F3$=2900 Hz. The loci for /w/ were: $F1$=250 Hz, $F2$=800 Hz, and $F3$=2200 Hz. The frequencies in formants changed quadratically over time. In the CVC stimuli, the selected values of loci and vowel formants produced initial $F2$ and $F3$ transitions changing in downward directions for /j/ and changing in upward directions for /w/, whereas the final $F2$ and $F3$ transitions were mirror images of the initial transitions. Thus in /jVj/ stimuli, the $F2$ and $F3$ transitions formed concave downward parabolas and in /wVw/ stimuli the $F2$ and $F3$ transitions formed convex upward parabolas. The course of $F1$ was identical in both types of stimuli: it was convex upward and always reached its maximal value at the point at which the rate of change was equal to zero. The segments of stimuli with formants changing over time which were 200 ms long were preceded and followed by 20-ms steady states.

The /jV/ and /wV/ stimuli were the first halves of /jVj/ and /wVw/ stimuli, respectively. The /Vj/ and /Vw/ stimuli were the second halves of /jVj/ and /wVw/ stimuli, respectively. The durations of these halves were extended from 100 to 200 ms and were either preceded (/wV/ and /jV/ stimuli) or followed (/Vw/ and /Vj/ stimuli) by 20-ms steady states. The #V# stimuli were 200 ms long.

The synthesized stimuli were analyzed using the Signal Technology, Inc., ILS computerized system. Autocorrelation analysis was conducted. A 20-ms Hamming window, 6.4-ms frame advance and no preemphasis were used. The data points from the analyses for the seven types of stimuli are shown in Fig. 1.

For testing, the 20 stimuli in each continuum were generated ten times in random order. Therefore, there were seven tests, each containing 200 stimuli. The tests were preceded by the two end-point stimuli, each repeated two times.

2. Subjects

All subjects were college students with audiometrically normal hearing. Because most subjects were available for limited periods of time, different groups of 12 or 24 subjects were tested.

3. Procedures

Subjects were tested individually in a sound-treated room. The test tapes were reproduced from a tape recorder (Revox, model PR99) and presented monaurally through a TDH-50 earphone at a comfortable level to the right ear. Before data collection, subjects were informed about the procedures, specifically that tests contained /ʊ/ and /ɪ/ vowels in isolation or combined with the consonants /j/ or /w/. At the

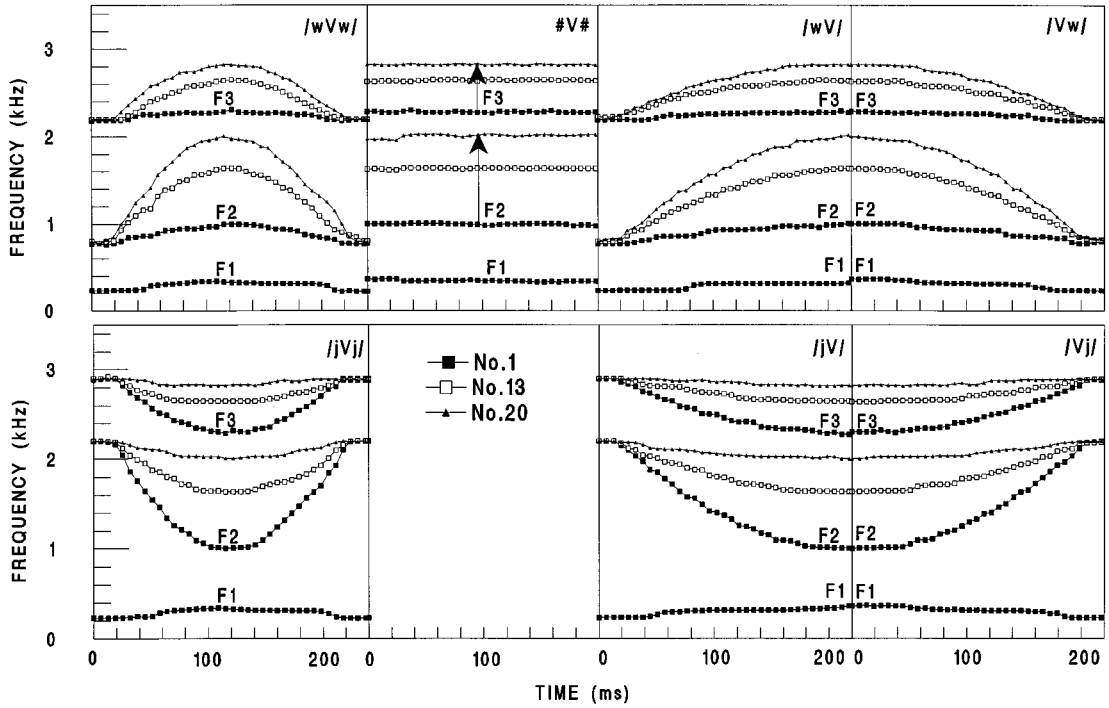


FIG. 1. Formant frequencies of stimuli from the seven continua as functions of time, $F1$ (the same for all stimuli in each continuum), $F2$ and $F3$ of stimuli Nos. 1, 13, and 20.

beginning of each continuum the end-point stimuli of that continuum were presented as examples of the contrast between the /ʊ/ and /ɪ/ stimuli that they should listen for in that condition. Subjects also had the opportunity to set the volume at a comfortable level while listening to the end-point stimuli. During data collection, subjects were asked to identify each stimulus as /ʊ/ or /ɪ/ and mark the corresponding column on a response sheet. The 200 responses were immediately computed. The two end-point stimuli, Nos. 1 and 20, were identified correctly 100% of the time as the intended vowels in all continua by all subjects. All subjects listened to the continuum with isolated vowels. #V# Group 1, consisting of 24 subjects, listened to the /jVj/ and /wVw/ continua. Group 2, consisting of 12 subjects, listened to the /jV/ and /Vj/ continua. Group 3, consisting of 12 subjects, listened to the /wV/ and /Vw/ continua. Thus each subject listened to three continua. The tests with the three continua were counterbalanced. The data collection with each subject was accomplished in 1 h.

B. Results

The percentage of /ʊ/ responses was plotted as a function of stimulus number for individual subjects for the seven stimulus types. For statistical evaluation of the data, the individual functions were smoothed using logit transformation (Anderson, 1989).[1] Two values were recorded from each function: the location of the boundary between /ʊ/ and /ɪ/, which was defined as the 50% response point (subsequently called the boundary), and the slope of the smoothed function at the boundary (subsequently called the slope). The mean boundary and slope values and their standard deviations for the tested continua and three groups of subjects are given in Table I.

1. #V# continua, groups 1, 2, and 3

Because the continua were tested with three groups of subjects, to compare performance of these groups the boundaries and slopes for the #V# continua were evaluated using one-factor analyses of variance. The boundaries and slopes

TABLE I. Means and standard deviations of boundaries and slopes for various types of stimuli and three groups of subjects.

Boundary (stimulus number)						
Group 1	__ #V#		__ jVj		__wVw	
24 Ss	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.
	13.54	1.10	14.21	1.60	12.08	1.40
Group 2	__ #V#		__ jV		__ Vj	
12 Ss	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.
	13.30	0.98	12.82	0.78	13.51	1.04
Group 3	__ #V#		__ wV		__ Vw	
12 Ss	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.
	13.43	1.32	11.93	1.28	13.15	0.77
Slope						
Group 1	__ #V#		__ jVj		__wVw	
	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.
	0.34	0.09	0.24	0.09	0.24	0.09
Group 2	__ #V#		__ jV		__ Vj	
	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.
	0.31	0.11	0.31	0.12	0.30	0.09
Group 3	__ #V#		__ wV		__ Vw	
	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.
	0.33	0.09	0.29	0.08	0.24	0.10

for the three groups of subjects were not statistically different ($p>0.05$), indicating that performance of these groups on the #V# stimuli was similar.

2. #V#, /jVj/, and /wVw/ continua, group 1

The boundaries and slopes for the three continua were evaluated using one-factor analyses of variance with repeated measures on the type of continuum. For the boundary, there was a significant effect of the type of continuum [$F(2,46)=17.30$; $p<0.0001$]. A *post hoc* Tukey test indicated that the mean boundary for the /jVj/ stimuli (14.21) was not significantly different ($p>0.05$) from the boundary for the #V# stimuli (13.54), whereas these two means were significantly greater ($p<0.05$) than the mean for the /wVw/ stimuli (12.08).

This result is in general agreement with data of Lindblom and Studdert-Kennedy (1967) who demonstrated shifts in boundaries for ʊ–ɪ continuum when the vowels were embedded in /j–j/ and /w–w/ contexts. The shifts were calculated as a difference between the boundary for a given type of stimuli (vowels in context) and the boundary for the #V# stimuli (isolated vowels). In the present study, the boundary shifts were −1.45 for the /wVw/ stimuli and +0.68 for the /jVj/ stimuli. In the Lindblom and Studdert-Kennedy (1967) study, the boundary shifts were −2.19 for the /wVw/ stimuli and −0.19 for the /jVj/ stimuli. In both studies, the absolute value of the shift for the /wVw/ stimuli was greater than for the /jVj/ stimuli.

For the slopes, there was also a significant effect of the type of continuum [$F(2,46)=13.64$; $p<0.0001$]. A *post hoc* Tukey test indicated that the slopes for the /jVj/ and /wVw/ stimuli (0.24 for both types) were shallower than the slope for the #V# stimuli (0.34) ($p<0.05$) but the slopes for the /jVj/ and /wVw/ stimuli were not significantly different ($p>0.05$). This result indicated that there were greater ranges of ambiguous stimuli in the CVC's continua than in the #V# continuum.

3. #V#, /jV/, and /Vj/ continua, group 2

The boundaries and slopes for the three continua were evaluated using one-factor analyses of variance with repeated measures on the type of continuum. The boundaries and slopes for the three types of continua were not statistically different ($p>0.10$). These results indicate that both /jV/ and /Vj/ stimuli were perceived similar to the #V# stimuli, which means that the preceding and following consonant /j/ had little effect on categorization of stimuli as /ʊ/ and /ɪ/ in the continua.

4. #V#, /wV/, and /Vw/ continua, group 3

The boundaries and slopes for the three continua were evaluated using one-factor analyses of variance with repeated measures on the type of continuum. For the boundary, there was a significant effect of the type of continuum [$F(2,22)=8.20$; $p<0.01$]. A *post hoc* Tukey test indicated that the means for the #V# (13.43) and /Vw/ (13.15) stimuli were not significantly different ($p>0.05$) but were greater than the mean for the /wV/ stimuli (11.93) ($p<0.05$). Thus it appeared that the consonant /w/ preceding vowels had a significant effect on categorization of /ʊ/ and /ɪ/ vowels, while the consonant /w/ following vowels did not have a significant effect on vowel categorization. The direction of boundary shift for the /wV/ stimuli, relative to the boundary for the #V# stimuli, was the same as for the /wVw/ stimuli, and the sizes of the boundary shifts were similar (−1.50 and −1.45 for /wV/ and /wVw/, respectively). For the slopes, the effect of the type of continuum was not significant.

5. Summary

For two types of stimuli (/wVw/ and /wV/) a perceptual compensation (''overshoot'') took place. Relative to the boundary for the #V# stimuli, the boundaries for the /wVw/ and /wV/ stimuli were shifted toward the lower stimulus numbers. The boundary for the /jVj/ stimuli was shifted toward the higher stimulus number, which also indicated the presence of a perceptual compensation, but this shift did not reach a level of statistical significance. For the stimuli /Vj/ and /Vw/ there were no significant boundary shifts. Because subject's task was to identify stimuli as /ʊ/ or /ɪ/, the initial vowel-like segments in the /Vj/ and /Vw/ stimuli appeared to be perceptually most important. The perceptual importance of the vowels might be also related to the fact that these syllables are not legal constructions in English. Because perceptual compensation was found for the /wV/ stimuli but was not found for similar stimuli in which formants changed linearly over time (Nábělek and Ovchinnikov, 1994) for the next experiment, the trajectory shapes were altered from quadratic to linear to determine if perceptual compensation can be found for these new stimuli.

For the /jV/ stimuli, the boundary shift was not significant but the boundary was slightly lower than for the #V# stimuli. This trend seems to indicate that the $F2$ trajectories were perceptually averaged over time and their frequencies perceived as slightly higher than the frequency of the steady state. Additional 12 data points were collected in the experiment 2 for the /jV/ stimuli to assess this trend.

II. EXPERIMENT 2

A. Methods

1. Test materials

Five 20-step /ʊ–ɪ/ continua in which $F2$ was varied in steps were tested. Two continua from experiment 1: (1) vowels in isolation (#V#) and (2) /wV/ stimuli were used. In addition, three new continua were generated with a Klatt synthesizer (Klatt, 1980) in which formant trajectories were either steady state or changed in upward directions. Formant trajectories in these new continua were as follows: (3) $F1$ and $F3$ changing quadratically as in the /wV/ stimuli and $F2$ changing linearly, (4) $F1$, $F2$, and $F3$ changing linearly, and (5) $F1$ and $F3$ steady state and $F2$ changing linearly. The $F1$, $F2$, and $F3$ trajectories of the /wV/ and of the three new stimuli are illustrated in Fig. 2. The linear trajectories had the same initial and final frequencies as the quadratic trajectories in the /wV/ stimuli. Thus, the frequency differences, ΔF's, for these and /wV/ stimuli were identical. The stimuli were 200-ms long except the /wV/ stimuli which were preceded

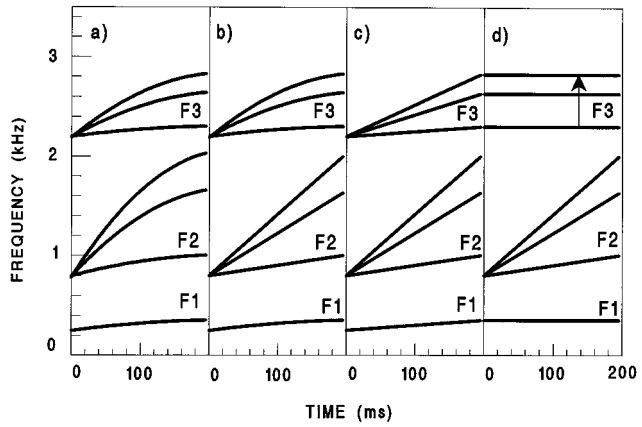


FIG. 2. Formant frequencies of stimuli: (a) /wV/; (b) $F1$ and $F3$ quadratic, $F2$ linear; (c) $F1$, $F2$, and $F3$ linear; and (d) $F1$ and $F3$ steady-state, $F2$ linear. $F1$ (the same for all stimuli in each continuum), $F2$ and $F3$ of stimuli Nos. 1, 13, and 20.

by 20-ms steady-state segments. The stimuli were numbered according to the final frequencies of $F2$ which was varied from 1000 to 2000 Hz. All other details of the new stimuli were the same as for the /wV/ stimuli.

2. Subjects

All subjects were college students with audiometrically normal hearing. The group consisted of 24 subjects.

3. Procedures

Procedures were the same as in experiment 1. Presentation of the five continua was counterbalanced. Data collection with each subject was accomplished in $1\frac{1}{2}$ h. Twelve subjects were additionally tested with the /jV/ stimuli.

B. Results

The boundaries and slopes of identification functions were established as in experiment 1. The mean boundary and slope values and their standard deviations are given in Table II. The boundaries and slopes for the five continua were evaluated using one-factor analyses of variance with repeated measures on the type of continuum. For the boundary, the effect of the type of continuum was significant [$F(4,92)=6.92$; $p<0.0001$]. A *post hoc* Tukey test indicated that the mean boundary for the #V# stimuli was significantly higher ($p<0.05$) than the boundaries for the remaining stimuli which were not significantly different ($p>0.05$) among themselves. For all types of stimuli there was an increase of /ɪ/ responses and the boundaries were shifted toward lower stimulus numbers relative to the boundary for the #V# stimuli. The lower stimulus numbers indicated that for all types of stimuli the perceived frequencies of $F2$ trajectories were beyond the extreme values actually synthesized in the stimuli, or in other words $F2$ frequencies were perceptually compensated (''overshooted'').

TABLE II. Means and standard deviations of boundaries and slopes for the following types of stimuli: (1) #V#, (2) /wV/, (3) $F1$ and $F3$ quadratic, $F2$ linear, (4) $F1$, $F2$, $F3$ linear, and (5) $F1$ and $F2$ steady state, $F2$ linear. Group of 24 subjects.

Boundary (stimulus number)									
$\overline{X}_1$	s.d.	$\overline{X}_2$	s.d.	$\overline{X}_3$	s.d.	$\overline{X}_4$	s.d.	$\overline{X}_5$	s.d.
12.94	1.09	11.60	1.48	11.82	1.30	11.52	1.54	11.46	1.56
Slope									
$\overline{X}_1$	s.d.	$\overline{X}_2$	s.d.	$\overline{X}_3$	s.d.	$\overline{X}_4$	s.d.	$\overline{X}_5$	s.d.
0.31	0.08	0.26	0.09	0.23	0.07	0.19	0.06	0.20	0.07

For the slope, the effect of the type of continuum was significant [$F(4,92)=12.98$; $p<0.0001$]. A *post hoc* Tukey test indicated that the mean slopes for the #V# and /wV/ stimuli (0.31 and 0.26, respectively) were not significantly ($p>0.05$) different between themselves but they were significantly ($p<0.05$) steeper from the slopes for the remaining stimuli with the linear $F2$ trajectories (0.23, 0.19, 0.20).

The result for the boundaries was unexpected because it was anticipated that a change from quadratic to linear $F2$ trajectories would decrease the size of perceptual compensation or would cause temporal averaging of $F2$. Such averaging was found in the Nábělek and Ovchinnikov (1994) study for a /ʊ–ɪ/ continuum in which stimuli had $F1$ and $F3$ steady state and $F2$ changing linearly in an upward direction. However, there were differences between $F2$ trajectories of the stimuli in the present and previous study. In the present study the initial frequency of $F2$ was 800 Hz, which was constant for all 20 stimuli in the continuum [Fig. 2(d)]. The final frequency of $F2$ was varied from 1000 to 2000 Hz. As a consequence, the absolute value of the frequency difference, ΔF, was 200 Hz for stimulus No. 1, about 756 Hz for the boundary stimulus, and 1200 Hz for stimulus No. 20. In the previous study, the absolute value of ΔF for $F2$ was constant and equal to 280 Hz in all 20 stimuli in the continuum. Thus ΔF for the the boundary stimulus was greater in the present than in the previous study. It is possible that the perceptual compensation (''overshoot'') observed in the present study was caused by the large ΔF and not by the shape of formant trajectories.

III. EXPERIMENT 3

A. Methods

1. Test materials

In order to determine if absolute value of ΔF of linearly changing $F2$ trajectories determines type of perception (perceptual compensation or temporal averaging), a new continuum was generated. In this continuum, $F2$ changed linearly as in the Nábělek and Ovchinnikov (1994) study but ΔF was increased from 280 to 500 Hz. This was the greatest ΔF which could be generated without overlapping with $F1$.

Four continua with the following types of stimuli were tested: (1) #V#, (2) a continuum from experiment 2 [Fig. 2(d)], with $F1$ and $F3$ steady state and $F2$ with a variable ΔF from 200 Hz in stimulus No. 1 to 1200 Hz in stimulus No. 20, shown again in Fig. 3(a), (3) $F1$ and $F3$ steady state and $F2$ with a constant ΔF of 500 Hz [Fig. 3(b)], and (4) $F1$ and $F3$ steady state and $F2$ with a constant ΔF of 280 Hz [Fig. 3(c)]. The stimuli were numbered according to the final

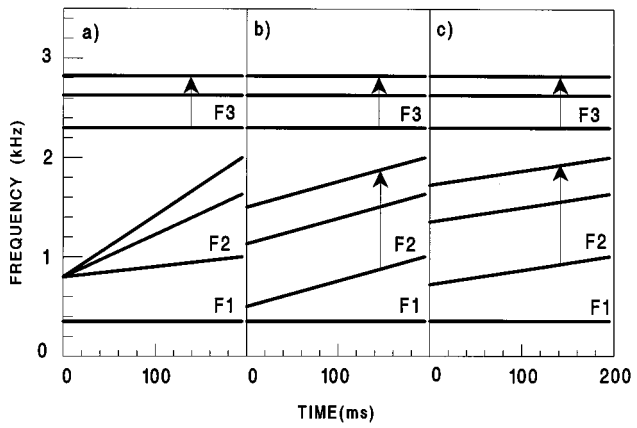


FIG. 3. Formant frequencies of stimuli with ΔF of $F2$: (a) variable, (b) constant 500 Hz, and (c) constant 280 Hz. $F1$ (the same for all stimuli in each continuum), $F2$ and $F3$ of stimuli Nos. 1, 13, and 20.

frequency of $F2$ which was varied from 1000 to 2000 Hz. The stimuli were 200-ms long. All other details of the stimuli were the same as for the /wV/ stimuli.

2. Subjects

A new group of 24 college students with audiometrically normal hearing was used.

3. Procedures

The procedures were the same as in experiment 1. Presentation of the four continua was counterbalanced. Data collection with each subject was accomplished in about $1\frac{1}{2}$ h.

B. Results

The boundaries and slopes of identification functions were established as in experiment 1. The mean boundary and slope values and their standard deviations are given in Table III. The boundary and slopes were evaluated for the four continua using one-factor analyses of variance with repeated measures on the type of continuum. For the boundary, the effect of the type of continuum was significant [$F(3,69)=11.52$; $p<0.0001$]. A *post hoc* Tukey test indicated that the mean boundary for the #V# stimuli (12.79) was significantly higher ($p<0.05$) than the boundaries for the stimuli with large ΔF, either variable (11.13) or constant (11.58). However, the boundary for the #V# stimuli was not significantly different ($p>0.05$) from the boundary for the stimuli with small ΔF (12.28). These results indicate that the stimuli with large ΔF of $F2$ were perceptually compensated while the stimuli with small ΔF of $F2$ were perceived as the final frequency of the formant.

TABLE III. Means and standard deviations of boundaries and slopes for the following types of stimuli: (1) #V#, (2) variable ΔF of $F2$, (3) constant $\Delta F=500$ Hz of $F2$, and (4) constant $\Delta F=280$ Hz of $F2$. $F1$ and $F3$ steady state in all types of stimuli. Group of 24 subjects.

Boundary (stimulus number)								
	1		2		3		4	
	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.
	12.79	1.24	11.13	1.36	11.58	1.71	12.28	1.56
Slope								
	1		2		3		4	
	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.
	0.31	0.09	0.19	0.06	0.25	0.08	0.26	0.10

For the slope, the effect of the type of continuum was significant [$F(3,69)=12.26$; $p<0.0001$]. A *post hoc* Tukey test indicated that the slope for the stimulus with variable ΔF (0.19) was significantly ($p<0.05$) shallower than the slopes for the remaining stimuli (0.31, 0.25, and 0.26) which were not significantly different among themselves ($p>0.05$).

IV. DISTRIBUTIONS OF BOUNDARY SHIFTS

The boundary shifts were calculated as the difference between the boundary of stimuli with $F2$ changing over time and the boundary of the #V# stimuli obtained by a given subject. Distributions of the boundary shifts for individual subjects were plotted in Figs. 4 and 5 for all stimuli for which at least 24 data points were collected. These were the /jVj/ and /wVw/ stimuli from experiment 1 and all stimuli from experiments 2 and 3. In experiment 2, an additional 12 data points were obtained for the /jV/ stimuli which were combined with the 12 data points for this stimulus from experiment 1. For the /wV/ stimuli, 36 data points were available, 12 from experiment 1 and 24 from experiment 2.

The distributions of the boundary shifts for the CVC and CV stimuli are shown in Fig. 4. The data points in Fig. 4 were spread indicating that various perceptual criteria might be used by subjects in categorization of the CVC and CV stimuli as /ʊ/ or /ɪ/. In order to investigate if the data points could be subgrouped, the agglomerative hierarchial clustering analysis was performed. Ward's minimum-variance method was used (Ward, 1963). The analysis indicated that the data points for the CVC stimuli could be divided into two subgroups. A multivariate secant method was used to fit the nonlinear regression model (Ralston and Jennrich, 1978), under an assumption of two normal distributions. Each of the two distributions was described by a position of a maximum, its width (σ), and a number of data points associated with this distribution. The calculated functions are plotted in Fig. 4(a) and (b).

The distribution of the data points for the /jVj/ stimuli [Fig. 4(a)] looked like a mirror image of the distribution for the /wVw/ stimuli [Fig. 4(b)]. The two maxima in the distributions seemed to indicate that two criteria were used by subjects in the categorization of these stimuli as /ʊ/ or /ɪ/. Perceptual compensation was more prevalent for the /wVw/ than for the /jVj/ stimuli (72% data points with a maximum at a boundary shift of -2.29, $\sigma=0.53$ and 48% data points with a maximum at a boundary shift of $+1.96$, $\sigma=0.37$, respectively). The maxima not greatly different than 0 (28% data points with a maximum at -0.25, $\sigma=1.04$, for the /wVw/ stimuli and 52% data points with a maximum at -0.31, $\sigma=0.77$, for the /jVj/ stimuli) indicated perceptual emphasis of the relatively steady-state, vowel-like middle segments.

The CV stimuli seemed to be interpreted differently than the CVC stimuli. It appeared that the distributions of the data points for the CV stimuli did not exhibit two distinct maxima. The clustering analysis indicated that the data

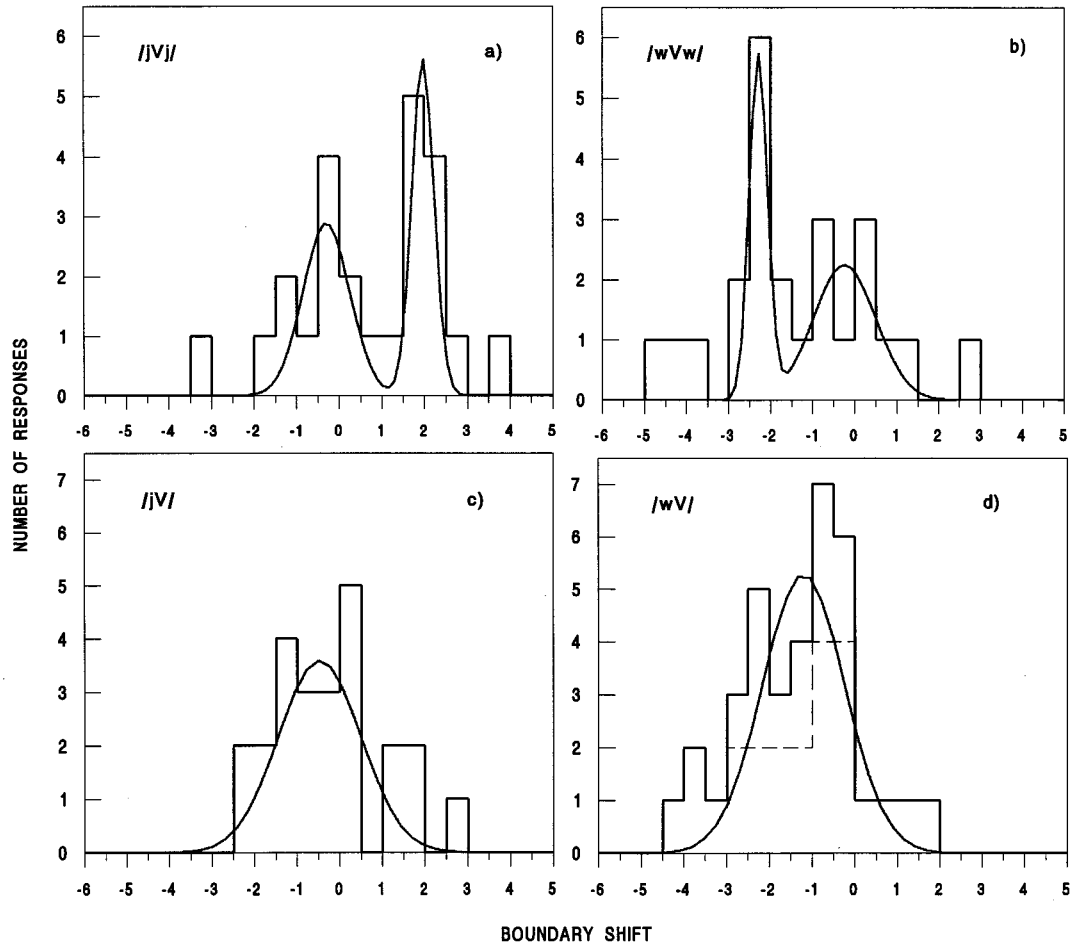


FIG. 4. Distributions of boundary shifts for individual subjects and fitted functions for the following stimuli: (a) /jVj, (b) /wVw/, (c) /jV/, and (d) /wV/ (for /wV/, data from experiment 2 are below dashed line and data from experiment 1 are above dashed line.)

points for the CV stimuli could not be divided into subgroups. For both the /jV/ and /wV/ stimuli a nonlinear regression model fitted the data under an assumption of one normal distribution. The calculated functions are plotted in Fig. 4(c) and (d). For the /jV/ stimuli, the maximum was at a boundary shift of -0.48, $\sigma=1.36$. This boundary shift was not significantly different than 0, but a trend toward negative boundary shift seemed to indicate that at least some subjects perceptually averaged $F2$ trajectories over time and perceived $F2$ slightly higher than the frequency of the final vowel-like segment. For the /wV/ stimuli, the maximum for the 36 data points was at a boundary shift of -1.20, $\sigma=1.35$ which indicated that subjects perceptually compensated these stimuli.[2)]

The distributions of the boundary shifts for the stimuli with linearly changing $F2$ trajectories from experiment 2 and 3, are shown in Fig. 5. They were arranged according to the mean values of boundaries from Tables II and III, from the smallest to the greatest. The distribution in Fig. 5(a) is for the 48 data points collected in the two experiments, the remaining distributions are for the 24 data points each. The clustering analysis indicated that the data points for these stimuli could not be divided into subgroups. Therefore a nonlinear regression model fitted the data under an assumption of one normal distribution. The calculated functions are plotted in Fig. 5. The maxima at the boundary shifts of -1.39, -1.38, -1.36, and -1.07 ($\sigma=1.42$; 1.25; 1.02; and 1.03), for the stimuli in Fig. 5(a), (b), (c), and (d), respectively, indicated perceptual compensation. The maximum at the boundary shift of -0.40, $\sigma=0.89$, for the stimulus in Fig. 5(e) indicated perceptual emphasis of the final frequencies of $F2$ changing over time. A *post hoc* Tukey test indicated that the three shifts (-1.39, -1.38, and -1.36) were not significantly different among themselves ($p>0.05$) but were different ($p<0.05$) from the -1.07 and -0.40 boundary shifts.

V. DISCUSSION

The results of the experiments of this study indicate that temporal changes of $F2$ trajectories influenced categorization of stimuli in /ʊ–ɪ/ continua. In the tested continua, perceptual compensation (''overshoot'') was found for stimuli and which transitory characters of $F2$ (and $F3$) were achieved by presenting vowels in /j-j/ and /w-w/ contexts. The /jVj/ and /wVw/ stimuli and corresponding isolated vowels, #V#, were synthesized according to specifications given in the Lindblom and Studdert-Kennedy (1967) study. The boundary for the /wVw/ stimuli, in which trajectories formed convex upward parabolas, was at a lower stimulus number than the boundary for the #V# stimuli. The differ-

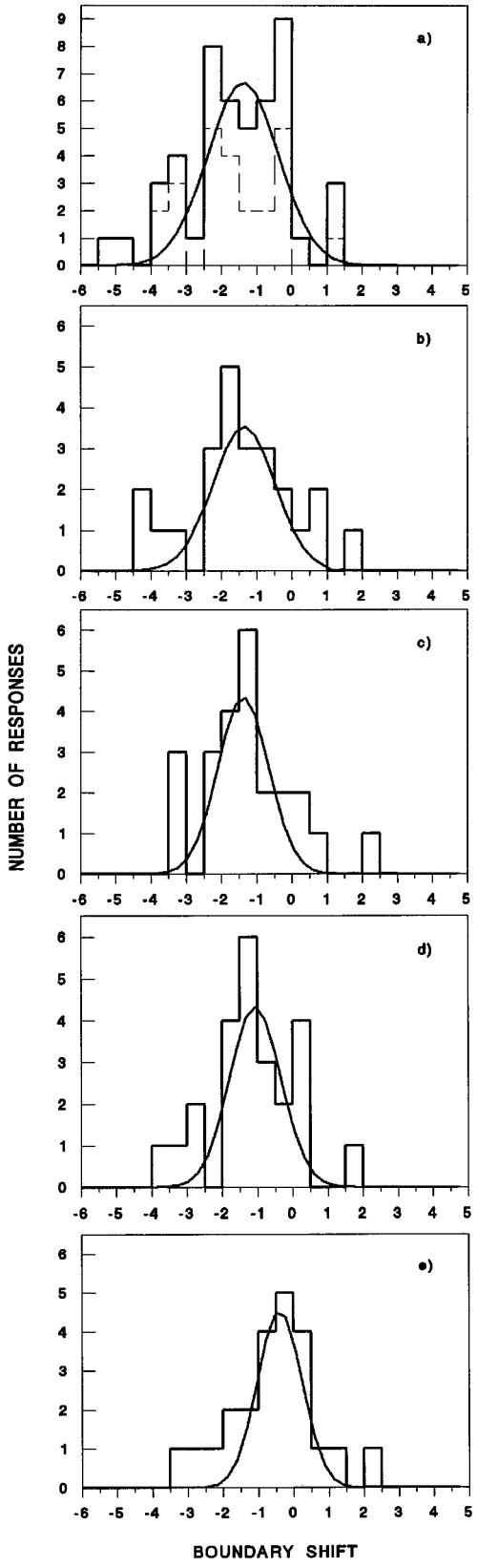


FIG. 5. Distributions of boundary shifts for individual subjects and fitted functions for the stimuli with linearly changing $F2$; (a) $F1$ and $F3$ steady-state, variable ΔF for $F2$ (data from experiment 2 are below dashed line and data from experiment 3 are above dashed line); (b) $F1$ and $F3$ linear, variable ΔF for $F2$; (c) $F1$ and $F3$ linear, constant $\Delta F=500$ Hz for $F2$; (d) $F1$ and $F3$ quadratic, variable ΔF for $F2$; and (e) $F1$ and $F3$ linear, constant $\Delta F=280$ Hz for $F2$.

ence between these two boundaries (the boundary shift) was statistically significant. The boundary for the /jVj/ stimuli, in which trajectories formed concave downward parabolas, was at a higher stimulus number than the boundary for the #V# stimuli, however, the difference between these two boundaries (the boundary shift) was not statistically significant. These results agree with the results of Lindblom and Studdert-Kennedy (1967) who found a boundary shift for the /wVw/ stimuli indicating perceptual compensation and no shift for the /jVj/ stimuli. However, there were large differences in amounts of boundary shifts for their subjects. Van Son (1993) suggested that if subjects in the Lindblom and Studdert-Kennedy study had interpreted the CVC stimuli as diphthongs, they would have used the extent of the ''glide'' segment as a cospecification of diphthong or glide identity. Van Son further suggested that diphthong or glide perception could also make the large differences between subjects more understandable. In the present study the analysis of distributions of boundary shifts indicated that various perceptual criteria were indeed used by subjects in the categorization of the CVC stimuli as /ʊ/ or /ɪ/. Whereas the data did not provide information about how subjects interpreted the CVC stimuli, Williams (1987) demonstrated that complex stimuli can be perceived as speech units or as sounds with pitch changing over time. When he instructed subjects to categorize /wVw/ stimuli as /ʊ/ or /ɪ/, he obtained perceptual compensation. However, when he instructed the subjects to judge the relative pitch of the same stimuli, he obtained either no shift (emphasis on the mid-point frequencies) or a shift in the opposite direction, suggesting perceptual averaging over time. Nearey (1989) suggested that the results of Williams (1987) provide an evidence that perceptual interpretation of the CVC stimuli depends on the mode of perception. In the present study, no instruction were given to the subjects how to interpret the stimuli. In appears that some subjects perceived the stimuli as a linguistic units and perceptually compensated $F2$ trajectories, others associated $F2$ trajectories with the frequencies of the steady states or with some frequencies resulting from perceptual averaging over time.

When the CVC stimuli in the present study were split into the CV and VC stimuli, the mean boundaries for the /jV/, /Vj/, and /Vw/ stimuli were not significantly different from the boundary for the #V# stimuli. Thus for these stimuli, the perceptual emphasis was at the relatively steady-state, vowel-like segments of $F2$ trajectories. Only for the new /wV/ stimuli was the mean boundary significantly different from the boundary for the #V# stimuli. The mean boundary shift for the /wV/ stimuli was in the same direction as for the /wVw/ stimuli which indicated that perceptual compensation took place for both types of stimuli.

For both the /Vw/ and /Vj/ stimuli, perceptual emphasis of the initial vowel-like segments, probably related to instructions required categorization of the stimuli as vowels. The perceptual compensation for the /wV/ stimuli and lack of compensation for the /jV/ stimuli may reflect differences between perception of the /wVw/ and /jVj/ stimuli (Fig. 4). There were more perceptual compensation responses for the /wVw/ than for the /jVj/ stimuli. For the /wV/ stimuli, there were many perceptual compensation responses with bound-

ary shifts ranging from very small to large. For the /jV/ stimuli, there were even fewer perceptual compensation responses, than for /jVj/ stimuli. The boundary shifts close to 0 indicated perceptual emphasis of the initial vowel-like segments and the negative boundary shifts indicated that the trajectories were perceptually averaged over time. Apparently as in the case of the CVC stimuli, for the CV stimuli different modes of perception could take place.

It was determined that perceptual compensation can take place not only for parabolic $F2$ trajectories, but also for $F2$ trajectories changing quadratically in upward directions. In order to determine if quadratic shape of $F2$ trajectories or specific shapes of $F1$ and $F3$ trajectories caused perceptual compensation, new types of stimuli were tested in experiment 2. The quadratic $F2$ trajectories were altered to linear trajectories changing in upward directions and trajectories of $F1$ and $F3$ were either quadratic, linear, or steady state. Regardless of the shape of $F1$ and $F3$ trajectories, perceptual compensation of the linear $F2$ trajectories was found. It was concluded that the perceptual compensation was related to large frequency difference for $F2$ trajectories, but it was not related to the shape of $F1$, $F2$, and $F3$ trajectories.

In experiment 3, the boundary shifts of three types of stimuli with linear $F2$ trajectories were compared. In all three types of stimuli $F1$ and $F3$ were steady states. The stimuli differed by both absolute values of ΔF and the initial frequencies of $F2$ trajectories. In the first type of stimuli, ΔF was variable and large, from 200 to 1200 Hz (in stimulus No. 1 and No. 20, respectively), and the initial constant frequency was equal to 800 Hz. In the second type of stimuli, ΔF of 500 Hz was constant and large, and the initial frequency was variable from 500 to 1500 Hz (in stimulus No. 1 and No. 20, respectively). In the third type of stimuli, ΔF of 280 Hz was constant and small, and the initial frequency was variable from 720 to 1720 Hz (in stimulus No. 1 and No. 20, respectively). The results of experiment 3 indicated that perceptual compensation took place for the stimuli with large ΔF's. The $F2$ trajectories with small ΔF's were perceived as their final frequencies. These results suggest that for 200-ms long stimuli with $F2$ trajectories changing in upward directions, perceptual compensation can take place for a large range of ΔF's. The perceptual compensation had a tendency to disappear when the ΔF was reduced to 280 Hz.

This study was limited to 200-ms long stimuli. A study with shorter stimuli could resolve if the size of perceptual compensation is dependent upon the value of ΔF or upon the rate of frequency change over stimulus duration. Although Lindblom and Studdert-Kennedy (1967) postulated that the perceptual shift would be greater for shorter than longer stimuli, they found only a trend for increased shift for their /wVw/ shortened from 200 to 100 ms but no change for the /jVj/ stimuli. It should be also considered that the relationship between the size of perceptual compensation and the rate of frequency change can be different for linear than for nonlinear $F2$ trajectories.

The absolute value of ΔF could have played role in perception of $F2$ trajectories in the CVC, CV, and VC stimuli. For the CVC stimuli, the ΔF was calculated as a difference between the initial and maximal frequencies, and for the CV and VC stimuli, the ΔF was calculated as a difference between the initial and final frequency. In the /jVj/, /jV/, and /Vj/ continua, the absolute value of ΔF for the boundary stimuli was 506, 581, and 542 Hz, respectively. In the /wVw/, /wV/, and /Vw/ continua the absolute value of ΔF for the boundary stimuli was 783, 772, and 840 Hz, respectively. The smaller boundary shift for the /jVj/ than for the /wVw/ stimuli and smaller boundary shift for the /jV/ than /wV/ stimuli might be related to the smaller absolute values of ΔF for the boundary stimuli in the continua with /j/ than in the continua with /w/.

For the 200-ms long stimuli with ΔF of 280 Hz, 54% responses (the boundary shifts greater than -0.5) indicated perceptual emphasis of the final frequency of the $F2$ trajectory. These responses agree with perceptual weighing toward the final frequencies found by Nábělek *et al.* (1970) for tone glides in which the frequency changed linearly over time. The glides investigated in their study which were changing in either upward or downward directions and had a mean frequency of 715 Hz were matched to pitch of steady tones. For the 120-ms-long glides in which the absolute value of frequency difference between the initial and final frequencies was 200 Hz, perceptual shift toward the final frequency was greater for the frequency changes in the upward than in the downward direction. In a recent study on glide perception, Schouten and Peeters (1995) reported that pitch matches to rising glides were shifted toward final glide frequencies while the matches to falling glides were close to the 1000-Hz mid-point of their stimuli. It should be noticed that in the Nabelek *et al.*(1970) and Schouten and Peeters (1995) studies in which pitch was assigned to the glides, no matches were beyond the final glide frequencies. It appears that the glides did not evoke perceptual compensations probably because they were not perceived as speech sounds and because the task was to assign a pitch. Even for the CVC stimuli which could be perceived as speech sounds, Williams (1987) eliminated perceptual compensation of responses when he asked his subjects to concentrate on pitch of the stimuli. In future studies with stimuli in which frequencies are changing over time a consideration should be given to instructions evoking specific mode of perception.

The tendency for the greater shifts toward the final frequency for the upward than downward frequency changes over time could be another reason for a significant boundary shift for the /wVw/ stimuli and not significant shift for the /jVj/ stimuli. This tendency might also have an effect on the /wV/ and /jV/ stimuli; the first were perceptually compensated and the second primarily associated with the final, relatively steady-state segments.

The results of the present study provide some support for van Son's (1993) suggestion that the presence of the perceptual compensation (''overshoot'') depends more on the perception of the context than on the actual formant track shape and for Nearey's (1989) assertion that the perceptual compensation requires speech mode perception. The ''overshoot'' was found for the /wVw/ stimuli with parabolic shape of the $F2$ trajectory, for the /wV/ stimuli with quadratic shape of the $F2$ trajectory, and for the stimuli with linear shape of the $F2$ trajectory and large frequency differ-

ence between the initial and final frequencies. While in the last types of stimuli, /w/ was not modeled, some subjects could have interpreted the $F2$ changing in upward directions as /w/.

However, perception of a consonant was not always sufficient for inducing perceptual compensation. The perceptual compensation was not found for 50% of the responses for the /jVj/ stimuli and for 80% of the responses for the /jV/ stimuli. The ''overshoot'' was not found for the /Vj/ and /Vw/ stimuli. The lack of perceptual compensation was in agreement with results of van Son (1993) for his CVC, CV, and VC stimuli in which C was either /n/ or /f/. There were differences, however, in the frequencies, which were perceptually emphasized. In the present study, the CVC, CV, and VC stimuli which were not perceptually compensated were usually perceived as their vowel-like segments, middle, final, or initial, respectively. In the van Son (1993) study, the emphasis was on the final stimulus frequencies.

The lack of perceptual compensation in the van Son (1993) study might be related not only to differences in perception of consonants in his and other studies (Lindblom and Studdert-Kennedy, 1967; Nearey, 1989) or to differences in data collection (phonetic identification versus categorization), but to relatively small values of ΔF in his stimuli. The largest frequency difference between the initial and middle frequencies in his parabolic $F2$ trajectories was either -375 or $+375$ Hz which was smaller than in the CVC stimuli in the present study. This relatively small frequency difference might not be sufficient for perceptual compensation.

Analyses of slopes of identification functions provided information about certainty of categorization of stimuli. In all continua the end point stimuli were identified 100% of the time as either /ʊ/ or /ɪ/. The steepness of the slopes was related to ranges of ambiguous stimuli in the continua. In all three experiments, the steepest slopes were found for the continua with the #V# stimuli. Significantly shallower slopes were found for the continua with the CVC stimuli and for the continua with the stimuli having $F2$ linearly changing over time with variable ΔF. The slopes for the continua with the CV and VC stimuli and with the stimuli having $F2$ changing linearly over time but with constant ΔF were not significantly different from the slope for the continuum with the #V# stimuli. It appeared that the ranges of stimuli which were unambiguously identified as /ʊ/ or /ɪ/ were not related to the type of perception of $F2$ trajectories, compensation, or emphasis of certain frequencies.

VI. CONCLUSIONS

In the 200-ms-long stimuli, the perceptual compensation for the $F2$ trajectories changing in upward directions appeared to be unaffected by the shape of the trajectories, quadratic or linear, and was similar to the compensation for the convex parabolic trajectories. For the linear $F2$ trajectories, the shape of $F1$ and $F3$ trajectories appeared to have no effect on the amount of the compensation. The only factor affecting the size and number of the perceptual compensation responses was the frequency difference between the initial and final frequencies of the $F2$ trajectories. The linear $F2$ trajectories with frequency difference up to 500 Hz were perceptually compensated and with frequency difference of 280 Hz were perceptually averaged over time with an emphasis of the final frequencies.

ACKNOWLEDGMENTS

We gratefully acknowledge funding for this study by the National Institute for Deafness and Communicative Disorders. We wish to thank Dr. Zbigniew Czyzewski for his help in synthesizing stimuli for experiment 1 and Dr. Hilary J. Crowly and Lisa Mortensen for collecting data.

[1]For each continuum and each subject, the sample fraction $f_i = s_i/m$ was calculated, where s_i is the number of /ʊ/ responses and m is the number of test repetitions. Sample logits $l_i = \ln[f_i/(1-f_i)]^{1/2}$ were regressed using a weighted regression with weight $w_i = [nf_i(1-f_i)]^{1/2}$. All points $f_i = 0$ and $f_i = 1$ were excluded except the closest ones to the boundary, i.e., the last for which $f_i = 1$ and the first for which $f_i = 0$. For these points, their values were changed from 0 to 0.05 or from 1 to 0.95.

[2]The values of boundary shifts taken as maxima of the calculated functions can be slightly different from the boundary shifts calculated from the mean boundaries given in Tables I, II, and III. The difference occurs because the experimental distributions are only sample distributions of the normal distribution class.

Anderson, E. B. (**1989**). *The Statistical Analysis of Categorical Data* (Springer-Verlag, Berlin), pp. 239–247.

DiBenedetto, M. G. (**1989**). ''Frequency and time variations of the first formant: Properties relevant to the perception of vowel height,'' J. Acoust. Soc. Am. **86,** 67–77.

Klatt, D. (**1980**). ''Software for a cascade/parallel formant synthesizer,'' J. Acoust. Soc. Am. **67,** 971–995.

Lindblom, B. E. F., and Studdert-Kennedy, M. (**1967**). ''On the role of formant transitions in vowel recognition,'' J. Acoust. Soc. Am. **42,** 830–843.

Nábělek, A. K., and Ovchinnikov, A. (**1994**). ''Perception of linear formant trajectories in vowels,'' J. Acoust. Soc. Am. **95,** 2978(A).

Nábělek, I. V., Nábělek, A. K., and Hirsch, I. J. (**1970**). ''Pitch of tone bursts of changing frequency,'' J. Acoust. Soc. Am. **48,** 536–553.

Nábělek, A. K., Czyzewski, Z., and Crowley, H. J. (**1993**). ''Vowel boundaries for steady-state and linear formant trajectories,'' J. Acoust. Soc. Am. **94,** 675–687.

Nearey, T. M. (**1989**). ''Static, dynamic, and relational properties in vowel perception,'' J. Acoust. Soc. Am. **85,** 2088–2113.

Ralston, M. L., and Jennrich, R. I. (**1978**). ''DUD, a derivative-free algorithm for nonlinear last squares,'' Technometrics **20,** 7–14.

Schouten, M. E. H., and Peeters, W. J. M. (**1995**). ''Matching frequency glides with steady tones,'' OTS-WP-FON-95-002, Research Institute for Language and Speech, Utrecht University.

van Son, R. J. J. H. (**1993**). *Spectro-temporal Features of Vowel Segments* (ICG-Printing, Dordrecht, The Netherlands).

Ward, J. H. (**1963**). ''Hierarchial grouping to optimize an objective function,'' J. Am. Stat. Assoc. **58** 236–244.

Williams, D. R. (**1987**). ''Judgements of coarticulated vowels are based on dynamic information,'' J. Acoust. Soc. Am. **81,** S17.

Effect of reducing temporal intensity modulations on sentence intelligibility

Ingrid M. Noordhoek
Experimental Audiology, Dept. of Otolaryngology, University Hospital VU, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands

Rob Drullman[a)]
Laboratory of Experimental Audiology, Dept. of Otorhinolaryngology, University Hospital Utrecht, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands

(Received 19 December 1995; revised 7 July 1996; accepted 13 August 1996)

Modulations in the temporal intensity envelope of 24 1/4-octave bands were reduced by *proportionally* raising the troughs and lowering the peaks relative to the mean intensity in each band. The effect on intelligibility of various degrees of modulation reduction was investigated by measuring the speech-reception threshold (SRT) in noise. For conditions of severe modulation reduction, the number of correctly received sentences in quiet was scored. The effect of this *deterministic* modulation reduction was compared to the effect of stochastic modulation reduction obtained with addition of noise. Results for 12 normal-hearing subjects show that in the case of deterministic modulation reduction, intelligibility is reduced to 50% when the modulation-transfer factor equals 0.10, whereas in the case of modulation reduction by addition of noise, this intelligibility is reached already at a modulation-transfer factor of 0.27. This confirms that the effect of additive noise on intelligibility cannot be understood completely as a result of only modulation reduction. As suggested by Drullman [J. Acoust. Soc. Am. **97**, 585–592 (1995)] two other factors associated with the addition of noise have to be taken into account: (1) the introduction of nonrelevant modulations, and (2) the corruption of the fine structure. [S0001-4966(97)04912-6]

PACS numbers: 43.71.Es, 43.66.Mk, 43.72.Dv [WS]

INTRODUCTION

Speech can be represented as a summation of frequency bands with amplitude-modulated signals. Each frequency band consists of a fine structure and a temporal envelope. The temporal envelope contains information that is essential for speech intelligibility. For common disturbances to speech, like noise and reverberation, the detrimental effects on intelligibility are generally understood as a consequence of reduced temporal envelope modulations. This is the basis of the concept of the modulation transfer function (MTF) and the speech transmission index (STI) for estimating intelligibility (Houtgast and Steeneken, 1985).

However, equal MTFs do not always lead to equal intelligibility; previous experiments (Drullman, 1995) showed that reduction of temporal modulations by several direct manipulations of the envelope is less detrimental to intelligibility than the same degree of reduction brought about by the addition of noise. A possible explanation was offered by Drullman (1995): first, the nonrelevant modulations introduced by the statistical nature of the noise leave the listener with a ''sorting problem'' (i.e., he/she is unable to separate the relevant speech modulations from the nonrelevant noise modulations) and second, noise affects the speech fine structure.

Direct manipulations of the speech envelope do not suffer from these side effects. The abovementioned study on direct manipulations of the temporal envelope was carried out to determine the relative contributions to intelligibility of peaks and troughs in the envelope, using very specific types of envelope processing. In the present study the effect on intelligibility of ''pure'' modulation reduction is investigated and compared to the effect of modulation reduction by addition of noise. The ''pure'' modulation reduction is performed by *proportionally* raising the troughs and lowering the peaks of the intensity envelope. This method of modulation reduction will be called ''deterministic'' to distinguish it from noise-induced reduction of modulations which has a stochastic character. The deterministic modulation reduction used here was applied to the entire range of modulation frequencies (uniform reduction). Similarly, the addition of continuous noise to a speech signal results in a uniform reduction.

I. METHOD

A. Materials and design

The speech materials consisted of 10 lists of 13 everyday Dutch sentences of 8 to 9 syllables read by a trained female speaker (Plomp and Mimpen, 1979). For masking in the SRT measurements, a Gaussian noise shaped according to the long-term average spectrum of the 130 sentences was used. Both the speech and the noise were digitized at a sampling rate of 15 625 Hz with 16-bit resolution.

For all sentences, deterministic modulation reduction of the temporal intensity envelope was performed in 24 1/4-

[a)]Current address: TNO Human Factors Research Institute, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands.

octave bands (linear-phase finite impulse response filter bank, slopes of at least 80 dB/oct), covering the range 0.1–6.4 kHz). Narrow bands were chosen, because after elimination of only wide-band modulations, part of the original narrow-band modulations can still be perceived by the auditory system (Drullman, 1995). The intensity envelope of each band was obtained by squaring the Hilbert (amplitude) envelope. The degree of reduction was controlled by multiplication of the band envelope relative to its average by the modulation-transfer factor[1] m, which can take any value between 0 (all modulations suppressed) and 1 (all modulations intact). To distinguish this deterministic modulation transfer from the effects by noise we will use a suffix (m_{det}). The modified envelope I_{mod} of each band at time t is given by

$$I_{\mathrm{mod}}(t)=\overline{I}+m_{\mathrm{det}}(I_{\mathrm{org}}(t)-\overline{I}), \quad 0\leqslant m_{\mathrm{det}}\leqslant 1, \tag{1}$$

where I_{org} symbolizes the original intensity envelope and $\overline{I}$ the long-term average intensity of that band.

Manipulations in the temporal domain will have an effect in the spectral domain. The spectral effect of modification of the temporal envelope according to Eq. (1) has been investigated by Drullman *et al.* (1996). It appears that uniform reduction of temporal modulations leads to uniform reduction of spectral modulations of (almost) equal magnitude. The spectral effect of temporal modulation reduction is thus a reduction toward the average bandfilter spectrum.

A preliminary test revealed that sentences presented in quiet were not completely intelligible when the modulation-transfer factor was set at 0.1. Therefore, sentences with transfer factors below 0.2 were presented in quiet and the number of correctly received sentences was scored. This will be referred to as the SIQ (Sentence Intelligibility in Quiet) experiment. For modulation-transfer factors from 0.2 to 1.0 the standard SRT procedure was carried out.

Ten processing conditions were investigated. In the SIQ experiment, the modulations in each frequency band were multiplied by the following factors: 0.00, 0.05, 0.10, and 0.15. In the SRT experiment the modulations were multiplied by factors running from 0.20 to 0.80 in steps of 0.15. Finally, the processing was performed with the modulation-transfer factor set at 1.0. This latter condition served as a reference without modification of the envelope.

The masking noise was processed separately for each condition in the same way as the sentences. There are two reasons for doing so. First, splitting up the wide-band speech signal into 1/4-octave bands and subsequently adding these bands resulted in slight high-frequency emphasis (less than 1 dB/oct). With the noise processed in the same way as the sentences, the (long-term) speech-to-noise ratio will remain constant over the entire spectral range in the SRT experiment. Second, as mentioned before, the statistical nature of the noise may introduce spurious modulations that disturb the perception of the speech modulations. Whether these noise modulations play a role or not, in any case they are reduced by the same amount as the relevant speech modulations.

B. Subjects

Subject were 12 normal-hearing students, whose ages ranged from 19 to 27. Their pure-tone air-conduction thresholds in the test ear did not exceed 15 dB HL at octave frequencies from 125 to 4000 Hz and at 6000 Hz.

C. Procedure

From the 10 lists of 13 sentences, 4 lists were used in the SIQ experiment and 6 in the SRT experiment. The lists were presented in a fixed order. The presentation order of the conditions was counter balanced over the listeners according to a diagram-balanced Latin square—4×4 for the SIQ experiment and 6×6 for the SRT experiment—to avoid order and lists effects. With 12 subjects, each sequence was presented to 3 subjects in the SIQ experiment and to 2 subjects in the SRT experiment. Six subjects started with the SIQ experiment and six with the SRT experiment.

In the SIQ experiment every sentence was presented once, at an A-weighted average level of 70 dB, after which the subject had to reproduce it. In the SRT experiment the level of the A-weighted masking noise was fixed at 70 dB; the level of the sentences was changed according to an adaptive up–down procedure. The first sentence in a list was repeated each time at a 4-dB higher level, until the listener could reproduce it correctly. The remaining 12 sentences were then presented only once, in an up–down procedure with a step size of 2 dB. The average signal-to-noise ratio for sentences 4–13 was adopted as the SRT for that condition.

The stimuli were presented monaurally through headphones (Sony MDR-CD999) at the test ear in a soundproof room. Before both the SRT and the SIQ experiment, a list of 13 sentences pronounced by a male speaker was presented, in order to familiarize the subjects with the procedure. For the SIQ experiment this list consisted of sentences with a modulation-transfer factor 0.20; for the SRT experiment another list was used with modulation-transfer factor 0.50.

II. RESULTS

Figure 1 shows the percentage correct score for sentences in quiet as a function of the imposed deterministic modulation-transfer factor m_{det}. Because the scores were not normally distributed over the listeners, medians and quartiles are displayed. By linear interpolation in Fig. 1, we find that the 50% correct score was reached for $m_{\mathrm{det}}=0.11$ (standard deviation=0.02). Wilcoxon tests for matched samples (Hays, 1988) showed that the scores differ significantly ($p<0.05$) for all but the 0.10 and 0.15 condition.

The mean SRT for sentences in noise as a function of m_{det} is shown in Fig. 2. Wilcoxon tests[2] for matched samples showed that the SRT is significantly elevated ($p<0.05$) for transfer factors from 0.65 downward. The curve drawn through the data will be described in Sec. III B, because we need the equations and assumptions of the next two sections to derive this curve.

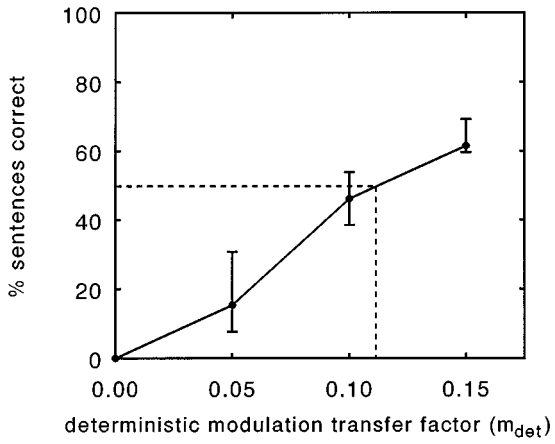


FIG. 1. Median percentage correct score for sentences in quiet as a function of the imposed deterministic modulation-transfer factor. Vertical bars represent the quartiles for each condition.

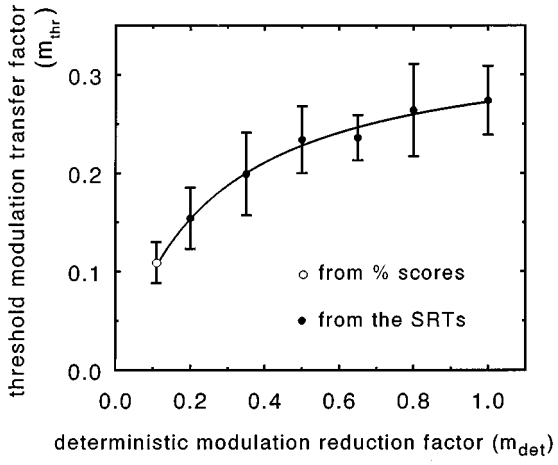


FIG. 3. Mean values of the threshold modulation-transfer factor at the speech reception threshold as a function of the imposed deterministic modulation-transfer factor. Vertical bars represent the standard deviation for each condition.

III. DISCUSSION

A. Combined effect of deterministic and noise-induced modulation reduction

In the SRT experiment, modulations in the temporal envelope were not only reduced by the imposed deterministic modulation reduction, but also by the addition of noise. The modulation-transfer factor brought about by the noise ($m_{\mathrm{noise,exp}}$) is determined by the signal-to-noise ratio at a given SRT (Houtgast and Steeneken, 1985)

$$m_{\mathrm{noise,exp}}=\frac{1}{1+10^{-\mathrm{SRT}/10}}. \tag{2}$$

The combined modulation-transfer factor at the speech reception threshold (m_{thr}) resulting from both modulation reductions can be calculated for each condition by multiplication of the deterministic and the noise factors:

$$m_{\mathrm{thr}}=m_{\mathrm{det}}\cdot m_{\mathrm{noise,exp}}. \tag{3}$$

When deterministic and noise-induced modulation reduction are equally detrimental to speech intelligibility, m_{thr} will have a constant value for different values of m_{det}.

Figure 3 shows the mean m_{thr} as a function of m_{det}. The filled dots are calculated from Eq. (3) based on the SRT data given in Fig. 2. The open dot represents the case $m_{\mathrm{thr}}=m_{\mathrm{det}}$, i.e., the m_{det} that corresponds to a 50% correct score for sentences in quiet. This data point has been computed from the data in Fig. 1 by linear interpolation. The curve fit to the data will be described in the next section, because we need the equations and assumptions of that section to derive the curve.

A one-way analysis of variance showed that the effect of deterministic modulation reduction on m_{thr} is highly significant [$F(5,50)=24$, $p<0.001$]. Pairwise comparisons with the Tukey HSD test (Hays, 1988) of the mean scores showed that m_{thr} for combined conditions of deterministic reduction and noise is significantly lower ($p<0.05$) for $m_{\mathrm{det}}\leqslant 0.65$ than for noise alone ($m_{\mathrm{det}}=1$). So, m_{thr} is not constant, which is caused by the fact that deterministic and noise-induced modulation reduction yield different effects on intelligibility.

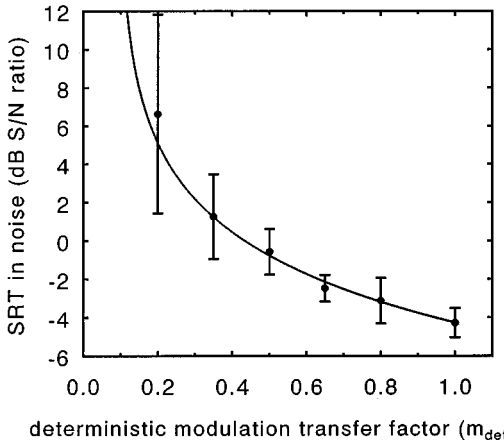


FIG. 2. Mean speech-reception threshold for sentences in noise as a function of the imposed deterministic modulation-transfer factor. Vertical bars represent the standard deviation for each condition.

B. Relationship between deterministic and noise-induced modulation reduction

The data in Fig. 3 represent different combinations of deterministic and noise-induced modulation reduction, that all lead to 50% intelligibility. It will be shown that these data can be described with a simple linear relationship between deterministic modulation reduction and the equivalent noise-induced modulation reduction that causes the same decrease in intelligibility.

Suppose the effectiveness of deterministic modulation reduction ($1-m_{\mathrm{det}}$) is a constant fraction of the effectiveness of modulation reduction introduced by noise ($1-m_{\mathrm{noise},i}$). To obtain the same intelligibility with both procedures this would imply:

$$1-m_{\mathrm{noise},i}=C(1-m_{\mathrm{det}}),\quad 0<C<1. \tag{4}$$

If in the experiment the imposed modulation reduction had been performed by addition of noise instead of direct

modulation reduction, the result would have been trivial, because m_{thr} would have been a constant $M_{thr,noise}$ for all values of the imposed modulation-transfer factor $m_{noise,i}$.

$$M_{thr,noise} = m_{noise,i} \cdot m_{noise,exp}. \quad (5)$$

With Eq. (4) $m_{noise,i}$ can be expressed as a function of an equivalent m_{det}, with the same effect on intelligibility. Using Eq. (5), the modulation-transfer factor associated with the noise that is added in the SRT experiment to obtain the threshold ($m_{noise,exp}$) can be expressed as a function of $m_{noise,i}$ and thus also as a function of m_{det}. Substitution of $m_{noise,exp}$ in Eq. (3) gives m_{thr} as a function of m_{det}.

$$m_{thr} = \frac{M_{thr,noise} \cdot m_{det}}{1 - C(1 - m_{det})}. \quad (6)$$

Fitting this equation by the least-squares method to the filled dots in Fig. 3, results in $C = 0.81$ and $M_{thr,noise} = 0.27$. The proportion of variance accounted for by the model is 57%. The lack-of-fit test on the pooled individual listener data (Walpole and Myers, 1989) revealed that the model accounts for a significant fraction of the variation [$F(4,66) = 0.41$, $p < 0.001$] and that the deviation from the model is not significant [$F(1,66) = 89$, $p = 0.80$]. Thus, given the experimental data, there is no need for a more complicated relationship between m_{det} and $m_{noise,i}$ than the one described by Eq. (4) with $C = 0.81$. By combining Eqs. (2), (4), and (5), we finally can derive the equation of the curve drawn in Fig. 2:

$$SRT = -10 \log\left(\frac{1 - C(1 - m_{det})}{M_{thr,noise}} - 1\right). \quad (7)$$

From $M_{thr,noise} = 0.27$ and Eq. (4), the threshold modulation-transfer factor when all modulations are reduced deterministically ($M_{thr,det}$) can be calculated. This value equals 0.10, which agrees well with the open dot in Fig. 3, that was obtained by linear interpolation in Fig. 1.

So, when the modulations are only reduced by addition of noise (e.g., $m_{det} = 1$), the speech reception threshold is reached when the modulation-transfer factor equals 0.27. In case of pure deterministic modulation reduction, the speech reception threshold is reached when the modulation-transfer factor equals 0.10. This confirms that deterministic modulation reduction is less detrimental to sentence intelligibility than the same degree of modulation reduction brought about by the addition of noise.

Another interpretation of Eq. (4) with $C = 0.81$ is that 81% of the effect of noise on intelligibility can be explained as the result of modulation reduction. To explain the remaining 19%, two other effects, specific for noise, have to be considered: (1) the introduction of nonrelevant modulations and (2) the corruption of the speech fine structure (see also Drullman, 1995).

C. Other modulation reduction schemes

We will now compare the threshold modulation-transfer factors for other schemes of modulation reduction with the two factors found in this study ($M_{thr,noise} = 0.27$ and $M_{thr,det} = 0.10$). Drullman (1995) investigated four other methods of modulation reduction. They were all performed on the amplitude envelopes of 24 1/4-octave bands and all left the speech fine structure intact. FT (flat troughs) set envelope parts below a certain target level (relative to the long-term rms level per band based on the total set of 130 sentences) to this target level; FP (flat peaks) set envelope parts above the target level to this level; and BLK (block pulse) set parts of the envelope below the target level to zero and above the target level to a fixed level. SN (speech and noise) replaced the original speech envelope by the speech+noise envelope. FT, FP, and BLK are deterministic modulation-reduction methods. SN is a stochastic modulation reduction method. Besides these processed signals (with intact speech fine structure) unprocessed speech and noise served as a reference (REF). In the case of REF, the target level simply was the actual relative noise level.

TABLE I. Overview of the effects on speech of different processing strategies together with the threshold modulation transfer factors. An asterisk indicates that the processing strategy mentioned in the left column produces the effect given in the upper row.

	Effects			
Processing	Modulation reduction	Nonrelevant modulations	Corruption of fine structure	M_{thr}
Addition of noise				
Present study	*	*	*	0.27
REF[a]	*	*	*	0.24
Envelope of speech+noise				
SN[a]	*	*		0.21
Deterministic				
Present study	*			0.10
FT, FP, BLK[a]	*			0.09, 0.11, 0.11

[a]From Drullman (1995).

The target levels at which the speech reception threshold (50% sentence intelligibility) was reached were 5.5, 6.5, 12, −25, and −31 dB *re*: rms for REF, SN, FT, FP, and BLK, respectively. For various target levels and the five modulation-reduction methods described, the corresponding modulation-transfer factors were measured (Drullman, 1995). It appears that the modulation-transfer factors at the speech reception threshold were 0.24, 0.21, 0.09, 0.11, and 0.11, for REF, SN, FT, FP, and BLK, respectively. Table I gives for different processing strategies an overview of the effects on speech that are involved together with the modulation-transfer factor at threshold (M_{thr}). Again, the stochastic modulation-reduction methods (SN and REF) are the most detrimental to intelligibility. SN (having the envelope of speech+noise, but the fine structure of speech) is somewhat less detrimental than REF (unprocessed speech+noise), which may be attributed to the speech fine structure being kept intact. For the deterministic modulation-reduction methods FT, FP, and BLK, essentially the same threshold modulation-transfer factors were found as for the present type of uniform, deterministic modulation reduction.

D. Results in relation to the STI

The MTF concept has led to the development of a speech transmission index (STI). The STI has proven to be

very successful in estimating the intelligibility of speech in the presence of common disturbances, like noise and reverberation. But previous experiments (Drullman, 1995) and the present study have demonstrated that its use cannot simply be extended to any manipulation of the temporal speech envelope. The origin of the modulation reduction is an important factor.

In calculating the STI, each modulation-transfer factor is transformed to a corresponding apparent signal-to-noise ratio $[(S/N)_{app}]$ irrespective of the actual sources of disturbance (Houtgast and Steeneken, 1985):

$$\left(\frac{S}{N}\right)_{app} = 10 \log\left(\frac{m}{1-m}\right). \quad (8)$$

This equation does not hold for deterministic modulation reduction. With Eq. (4) we can revise Eq. (8) in such a way that it also holds for the uniform deterministic modulation reduction performed in this study:

$$\left(\frac{S}{N}\right)_{app} = 10 \log\left(\frac{1-C(1-m)}{C(1-m)}\right), \quad (9)$$

where C is a constant that depends on the origin of the modulation reduction. C equals 0.81 in case of deterministic modulation reduction and C equals 1 in case of common disturbances, like noise or reverberation.

IV. CONCLUSIONS

Uniform, deterministic reduction of temporal modulations is less detrimental to sentence intelligibility than the same modulation reduction brought about by addition of noise. A linear relationship between deterministic and noise-induced modulation reduction causing the same decrease in intelligibility can describe the data well.

In case of deterministic modulation reduction, 50% intelligibility is reached when the modulation-transfer factor equals 0.10 and in case of noise-induced modulation reduction, when the modulation-transfer factor equals 0.27. Hence, the effect of additive noise on intelligibility cannot be understood completely as a result of modulation reduction per se. ''Pure'' modulation reduction can explain 81% of the detrimental effect of additive noise. To explain the remaining 19% two other factors have to be taken into account, as proposed by Drullman (1995): (1) the introduction of nonrelevant modulations and (2) the corruption of the speech fine structure.

ACKNOWLEDGMENTS

This research was supported by the Foundation ''Heinsius-Houbolt Fonds,'' The Netherlands. The authors wish to thank Tammo Houtgast, Joost Festen, and Niek Versfeld for their useful comments on the manuscript.

[1]Although the factor *m* has been introduced as ''modulation-*reduction* factor'' by Houtgast *et al.* (1980), we will denote this factor by the more consequent term ''modulation-*transfer* factor,'' because it quantifies the degree of transfer of the original intensity modulations.

[2]Because of the significant variations in the spread of the data among conditions, the significance of differences in the SRTs was not tested with parametric tests.

Drullman, R. (**1995**). ''Temporal envelope and fine structure cues for speech intelligibility,'' J. Acoust. Soc. Am. **97**, 585–592.

Drullman, R., Festen, J. M., and Houtgast, T. (**1996**). ''Effect of temporal modulation reduction on spectral contrasts in speech,'' J. Acoust. Soc. Am. **99**, 2358–2364.

Hays, W. L. (**1988**). *Statistics* (Holt, Rinehart, and Winston, New York), 4th ed.

Houtgast, T., Steeneken, H. J. M., and Plomp, R. (**1980**). ''Predicting speech intelligibility in rooms from the modulation transfer function. I. General room acoustics,'' Acustica **46**, 60–72.

Houtgast, T., and Steeneken, H. J. M. (**1985**). ''A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria,'' J. Acoust. Soc. Am. **77**, 1069–1077.

Plomp, R., and Mimpen, A. M. (**1979**). ''Improving the reliability of testing the speech reception threshold for sentences,'' Audiology **18**, 43–52.

Walpole, R. E., and Myers, R. H. (**1989**). *Probability and Statistics for Engineers and Scientists* (Macmillan, New York), 4th ed., pp. 381–385.

Spectral balance as a cue in the perception of linguistic stress

Agaath M. C. Sluijter,[a] Vincent J. van Heuven,[b] and Jos J. A. Pacilly[c]
Holland Institute of Generative Linguistics, Phonetics Laboratory, Leiden University, Cleveringaplaats 1, P.O. Box 9515, 2300 RA Leiden, The Netherlands

(Received 28 March 1995; revised 1 August 1996; accepted 2 August 1996)

In this study, the claim that intensity, as an acoustic operationalization of loudness, is a weak cue in the perception of linguistic stress is reconsidered. This claim is based on perception experiments in which loudness was varied in a naive way: All parts of the spectrum were amplified uniformly, i.e., loudness was implemented as intensity or gain. In an earlier study it was found that if a speaker produces stressed syllables in natural speech, higher frequencies increase more than lower frequencies. Varying loudness in this way would therefore be more realistic, and should bring its true cue value to the surface. Results of a perception experiment bear out that realistic intensity level manipulations (i.e., concentrated in the higher frequency bands) provide stronger stress cues than uniformly distributed intensity differences, and are close in strength to duration differences. [S0001-4966(97)00412-8]

PACS numbers: 43.71.Es, 43.70.Fq [RAF]

INTRODUCTION

Dutch and English are languages with word stress: one of the syllables of a word, especially when pronounced in citation form, is perceived as the most prominent one, the so-called lexical stress position of the word. The phonetic correlates of lexical stress in these languages are pitch, duration, loudness, and vowel quality (Lehiste, 1970; Beckman 1986, and references mentioned there). Of these, pitch and duration have been found the most important perceptual cues; intensity, as an acoustical operationalization of loudness, is generally claimed to be of lesser importance (among others: Fry, 1955, 1958; van Katwijk, 1974), while vowel quality is the least important cue (Fry, 1965; Rietveld and Koopmans-van Beinum, 1987). When words are spoken outside focus, i.e., without a pitch accent on the stressed syllable, the position of the stress has to be inferred from the remaining cues such as duration and intensity.

In the older linguistic and phonetic literature it was generally held that languages such as English and Dutch are characterized by so-called dynamic (rather than melodic) stress. That is to say, stressed syllables are produced with greater pulmonary and glottal effort, with greater loudness as the primary perceptual correlate (Sweet, 1906; Bloomfield, 1933). With the advent of speech synthesis techniques in the fifties this view was quickly discredited, when manipulating intensity (i.e., gain), as an operationalization of loudness variation, proved virtually inconsequential for stress perception (Fry, 1955, 1958 for English; Mol and Uhlenbeck, 1956 for Dutch; Issatchenko and Schädlich, 1966 for German).

In the present study, the claim that loudness is a weak cue in the perception of linguistic stress is reconsidered. Recently, Sluijter and van Heuven (1996) showed that intensity level differences between stressed and unstressed Dutch syllables are concentrated in the higher parts of the spectrum, whereas intensity differences in the lower part of the spectrum, i.e., below 500 Hz, were negligible. We assume that these differences in the higher parts of the spectrum are caused by a difference in the shape of the glottal waveform, due to an increase in vocal effort when producing stressed syllables, and are therefore a reflection of effort, and are perceived in terms of greater loudness.

The assumption that vocal effort is related to the perception of loudness was explored by Brandt *et al.* (1969). They independently varied vocal effort and intensity of continuous speech stimuli. In their experiments speech samples that were produced with greater effort, were estimated as louder than the same samples spoken with less effort, even when the mean intensity was adjusted so as to be constant. They considered the acoustic spectrum to be a special cue for the perception of vocal effort. Glave and Rietveld (1975) also examined the role of effort in speech loudness; their results confirmed that greater vocal effort is related to greater perceived loudness. Furthermore, they showed that the spectra of vowels spoken with greater effort have more intensity in the higher-frequency region, which they assumed to be caused by the changes in the source spectrum due to a more pulse-like shape of the glottal waveform.

This operationalization of loudness variation, i.e., increasing intensity in the higher frequency bands only, differs substantially from implementing loudness in terms of changing the gain factor uniformly across the spectrum as was done in the perceptual experiments above. Therefore, varying the acoustical correlate of loudness in a more realistic way, i.e., by varying the spectral balance,[1] should bring out the true cue value of loudness for stress perception.

If, indeed, varying intensity level in the higher frequency bands only is a perceptually more effective stress cue than applying uniform intensity level increments, a second question arises: What is the importance of the loudness cue relative to other stress cues? In order to keep this second question within manageable proportions, we will examine

[a]Now at KPN Research, P.O. Box 421, 2260 AD Leidschendam, The Netherlands. Electronic mail: a.m.c.sluijter@research.kpn.com
[b]Electronic mail: heuven@rullet.leidenuniv.nl
[c]Electronic mail: pacilly@rullet.leidenuniv.nl

the importance of intensity level manipulations relative to that of duration manipulation, i.e., the cue that has been advanced as the most reliable stress cue so far.

It is not the intention of the present study to question the primacy of the $F0$ cue in stress perception, since we regard $F0$ movement as a cue for sentence accent rather than for linguistic word stress. There is ample evidence, e.g., in Dutch, that an $F0$ movement with the appropriate excursion size ($\geqslant$4 semitones) and time alignment (cf. 't Hart *et al.*, 1990; Hermes and Rump, 1993) is a sufficient cue for accent, and *a fortiori* for stress, since accents are normally associated with the lexically stressed syllable of a word. In fact, when the accent is shifted to a nonstressed syllable so as to signal a metalinguistic contrast as in *I said SUGgest not DIgest,*[2] the original stress cues in the second syllable of *suggest* are almost completely obliterated and transferred to the initial syllable, cf. Sluijter and van Heuven (1995). However, the $F0$ cues are not invariant stress cues, since they disappear at the sentence level when the word is deaccented through focus manipulation (cf. van Heuven, 1987; Sluijter and van Heuven, 1996). Formant changes, finally, have consistently been reported as the least important cue for word stress (and sentence accent).

We will therefore examine the relative strength of the two implementations of loudness and duration in unaccented, i.e., nonfocused, targets.

In the experiment described below we studied the perception of stress position in the disyllabic Dutch nonsense word *nana* by manipulating vowel duration, spectral balance (intensity level increments in the higher frequency bands only) and intensity (uniformly distributed gain increments) in accordance with our production data (Sluijter and van Heuven, 1996). The hypothesis to be tested is that spectral balance is a stronger stress cue than overall intensity, and that the importance of spectral balance as a stress cue will approximate (or even surpass) that of duration. The possible finding that more realistic loudness manipulations provide a stronger stress cue than the traditional operationalization of loudness as gain/intensity should then, at least in part, rehabilitate the claim of the above mentioned older literature by Sweet (1906) and Bloomfield (1933).

I. PERCEPTION EXPERIMENT I

A. Methods

1. Material

We used the reiterant nonsense word pair /ˈnaːnaː/-/naːˈnaː/. This type of speech allows us to vary duration, spectral balance and intensity without taking into account segmental differences between both syllables, e.g., differences in intrinsic duration (Peterson and Lehiste, 1960) and intrinsic intensity (Lehiste and Peterson, 1959) of vowels, and possible perceptual compensation for these features. Reiterant speech was also used by Morton and Jassem (1965), van Katwijk (1974), Berinstein (1979) and many others in similar experiments and is assumed to be like nonreiterant speech in all aspects which are important in the study of prosody (Larkey, 1982).

TABLE I. Overview of the duration manipulations yielding seven duration steps. Durations are given (in ms) for first ($\sigma1$) and second syllable ($\sigma2$) separately, as well as total word duration ($\sigma1+\sigma2$).

Duration	$\sigma1$	$\sigma2$	$\sigma1+\sigma2$
1 *NAna*	250	185	435
2	230	200	430
3	210	215	425
4 *neutral stimulus*	190	230	420
5	170	245	415
6	150	260	410
7 *naNA*	130	275	405

We used the unstressed syllable *na* of the sentence *Wil je naˈna zeggen* /vɪ| jə naːnaː zɛɣə/ 'Will you [naˈna] say,' uttered by a male speaker with a pitch movement on *zeggen*, taken from the production study. This speaker was chosen out of a set of ten because the quality of his voice was preserved best in LPC resynthesis in comparison with the other male and female speakers.

We concatenated two syllables *na* to form the disyllabic nonsense word *nana*. The duration of the syllables was varied in seven steps from *ˈnana* to *naˈna* in accordance with our production data (Sluijter and van Heuven, 1996). We took a representative duration range for reiterant speech averaged over the speakers. This led to the following experimental values: the initial syllable was varied in seven steps of 20 ms from 250 to 130 ms, the second syllable was varied in seven steps of 15 ms from 185 to 275 ms. Note that an increase of the duration of the first syllable covaries with a decrease of the duration of the second syllable. The stimulus with an initial syllable of 190 ms and a final syllable of 230 ms (number 4) was meant to be temporally ambiguous for stress perception. The longer average duration of the second syllable was copied from actual speech production so as to reflect the influence of word-final lengthening (Wightman *et al.*, 1992; Sluijter and van Heuven, 1996). Table I gives an overview of the resulting stimuli.

In order to reduce the dimensionality of the stimulus space, we implemented spectral balance in terms of variable intensity levels below and above 0.5 kHz. It appeared from our production data (Sluijter and van Heuven, 1996) that the intensity levels in the three octave bands (B2–B4) were correlated (r^2 between 0.45 and 0.57), whereas there was no correlation between the base band B1, and any of the higher octaves (r^2 between 0.04 and 0.23). The spectral balance of the syllables was therefore varied by increasing the levels of the frequency components above 500 Hz by 3, 6, or 9 dB, in either the initial or the final syllable. We used the digital filtering facilities of the speech and signal processing package XAudlab (Lagendijk, 1992) implemented on a Silicon Graphics Indigo/Irix computer. The filtering and filter design algorithms implemented in this package use the standard FIR structure and DFT approach. The spectral balance steps were a straightforward quantization of the differences between the stressed and unstressed realizations of the syllables *na* in our production study. We applied uniform intensity level increments to all the frequencies above 500 Hz, although strictly speaking the intensity differences in the third filter band

TABLE II. In the left-hand part of the table the intensity level manipulations per step are presented. Levels were increased for components above 500 Hz. These manipulations caused overall intensity level increases of the syllables, which are presented in the right part of the table. These values were used to vary intensity level uniformly in all bands.

	Increased levels above 500 Hz		Increase in overall intensity level (incl. baseband)	
Step	$\sigma 1$	$\sigma 2$	$\sigma 1$	$\sigma 2$
1	+9 dB	⋯	+3 dB	⋯
2	+6 dB	⋯	+2 dB	⋯
3	+3 dB	⋯	+1 dB	⋯
4	⋯	⋯	⋯	⋯
5	⋯	+3 dB	⋯	+1 dB
6	⋯	+6 dB	⋯	+2 dB
7	⋯	+9 dB	⋯	+3 dB

(1.0–2.0 kHz) should be a little larger than those in the second (0.5–1.0 kHz) and fourth (2.0–4.0 kHz) filter bands. Crucially, however, we did not add any intensity to the base band.

Larger differences than the 9-dB increase in the higher bands occur occasionally in our production data, but this value was chosen as the maximum increment as stimuli with larger intensity level differences in the higher bands sounded less than acceptable.

These vocal effort/spectral balance manipulations yielded overall intensity level changes of approximately 1, 2, or 3 dB, respectively. Consequently, these steps were used to vary overall intensity level. Overall intensity level was varied by simply multiplying the sample values of either the initial or the final syllable by 1.12, 1.26, and 1.41, respectively. Table II gives an overview of the manipulations.

As can be seen in Table II, the overall intensity level differences in both stimulus sets are identical. There are seven duration levels, seven intensity levels, and two implementation methods. This nominally yields 98 stimuli but there were only 91 in practice since stimuli with the neutral intensity level (i.e., step 4) are identical for the two methods. The first part *Wil je, nana*, and the last part of the sentence *zeggen* were concatenated and resynthesized using straightforward LPC synthesis. As a consequence spectral discontinuities were smoothed over a window length of 25 ms. A sample frequency of 10-kHz, 4.5-kHz low-pass filter and 12-bit amplitude resolution were used for both analysis and resynthesis (18 reflection coefficients, Hamming window length 25.6 ms, window shift 10 ms).

Stimuli were presented without a pitch movement on the target in a fixed carrier phrase *Wil je [target] zeggen* (Will you [target] say). The carrier sentence was synthesized with a declining pitch contour, modeled after the pitch contour of the original sentence, such that the target was part of a falling declination line. An accent-lending pitch movement was realized on the first syllable of *zeggen*. The targets were presented in their original context since presenting stimuli out of their original context induces strong perceptual bias to perceive the stress on the first syllable (van Heuven and Menert, 1996). The prefinal position in the sentence was originally chosen to avoid preboundary lengthening in the targets; in the present experiment it is therefore necessary to avoid perceptual compensation for preboundary lengthening by maintaining this position.

2. Subjects and procedure

One stimulus tape was prepared containing the 91 stimuli in two different random orders. The 182 stimuli were presented in blocks of 13 utterances with 2-s intervals between utterances, offset to onset, and a larger interval and a 500-ms tone of 1000 Hz separating the blocks. This was done to prevent subjects from losing their way on the answer sheet, and to give them time to turn the pages of their answering booklet. The tape started with five practice utterances to familiarize the subjects with their task. Forty-six listeners participated in the test. Twenty-four subjects (phonetically trained staff and students of the Faculty of Arts) were tested in two groups in a language laboratory at Leiden University. They listened to the tapes over headphones. Twenty-two (phonetically naive) subjects participated in the test as part of a phonetics class taught by the second author, and were tested in a classroom at Leiden University. They listened to the tape over loudspeakers. Subjects were instructed to determine the stress position of *nana* in each utterance (with binary forced choice) and to note their responses on the response sheets provided. The experiment lasted approximately 30 min.

B. Statistical analysis

We determined the number of judgments favoring initial stress for each stimulus and expressed this as a percentage, henceforth $p(init)$.

There were three goals for the statistical analysis. The primary goals were to establish the relative strengths of duration and intensity level manipulations as stress cues, and to determine to what extent the way of varying intensity (overall versus above 500 Hz only) interacts with the effects of duration and intensity level. An additional goal was to determine to what extent the way of presentation interacts with the above effects. A four-way analysis of variance was performed, with $p(init)$ as the dependent variable, and with *presentation* (headphones versus loudspeakers), *method* of varying intensity (intensity level increments in all bands versus spectral balance, i.e., increasing intensity above 0.5 kHz only), *duration* (seven steps) and *intensity level* (seven steps) as fixed effects and with repetition as repeated measure.[3] The effects of *duration* and *intensity level* variations will show up as main effects in the ANOVA. The importance of *method* and *presentation* will be visible in their interactions with *duration* and *intensity* level. The main effects of *presentation* and *method* are irrelevant in this research, since they will merely reflect a difference in overall bias favoring one stress position over the other.

C. Results

1. Global presentation

We computed the consistency of each subject by comparing their answers on the first and the second presentation of the stimuli. Subjects who were not consistent in more than

TABLE III. Main effects and interactions of duration, intensity level, presentation (headphones versus loudspeakers), and method (of varying intensity: overall versus high frequency bands only) on *p(init)*. *F* ratio, significance of *F* and percentage of explained variance (η^2) are given.

Effects	*F*	sign.	η^2
Main effects			
Duration	761.3	<0.001	68
Intensity level	129.4	<0.001	12
Presentation	2.8	NS	0
Method of variation	3.8	NS	0
Two-way interactions			
Duration * intensity level	7.8	<0.001	4
Duration * presentation	53.5	<0.001	5
Duration * method	5.4	<0.001	0
Intensity level * presentation	7.8	<0.001	1
Intensity level * method	46.1	<0.001	4
Presentation * method	<1	NS	0
Three-way interactions			
Duration * int. level * presentation	1.9	0.003	1
Duration * int. level * method	2.4	<0.001	1
Duration * presentation * method	1.6	NS	0
Int. level * presentation * method	4.1	0.001	0

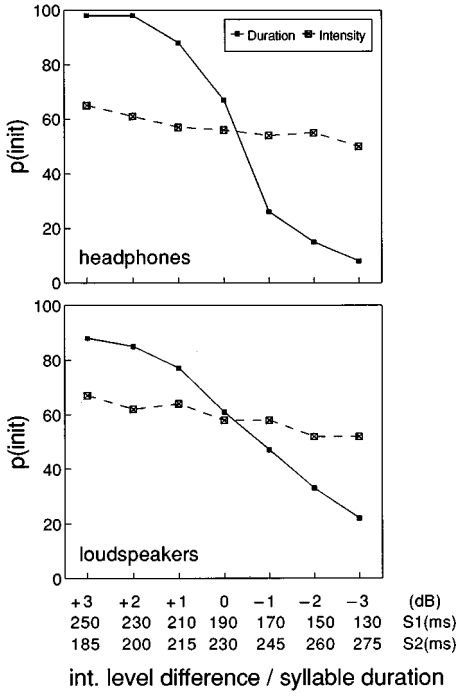


FIG. 1. Percentage of listeners ''initial stress'' judgments, $p(init)$, for the 91 stimuli *nana* as a function of syllable duration (solid lines) and overall intensity (dashed lines). The differences in *intensity level* ($IL_{\sigma 1}-IL_{\sigma 2}$ in dB), obtained by spectrally uniform amplification, are given along the *x* axis, top line. *Duration* values (in ms) are given on the middle and bottom lines for the first and second syllable, respectively. The results are presented for each presentation condition separately: headphones (upper panel) and loudspeakers (lower panel).

60% of the cases were omitted from further analysis. The 60% consistency cutoff point was chosen as there was a clear discontinuity between the six poorest subjects and the 40 individuals who remained in the analysis. Twenty-one subjects who listened to the tape over headphones and 19 subjects who listened to the tape over loudspeakers were used for further analysis.

The listening test yielded a total of 7280 responses (91 stimuli * two repetitions * 40 subjects). Overall, 57% of the responses favored initial stress, which indicates that there is a slight bias for initial stress. This bias is above chance, as determined by a binomial test ($p<0.001$).

In Table III the main effects and interactions of *duration*, *intensity level*, *presentation*, and *method* are given.

There is a large effect of both *duration* and *intensity level* on $p(init)$. In answer to our question if varying intensity level in a more realistic way, i.e., by varying the spectral balance, has an effect on stress perception, we can provisionally conclude from the highly significant interaction of *intensity level* with *method*, that the method of variation has at least a considerable influence on the effect of *intensity level* on $p(init)$. Furthermore, the significance of the two- and three-way interactions with *presentation* means that the way of presentation has an influence on both the effect of *duration* and *intensity level* on $p(init)$. Given the significant two- and three-way interactions we decided to study the main effects of duration and intensity level separately for each presentation condition (headphones versus loudspeakers) and for each method of varying intensity (overall level versus manipulating spectral balance). Therefore, we ran two separate two-way analyses of variance with *duration* and *intensity* (uniformly distributed gain increments, henceforth *intensity*) as fixed effects and with repetition as repeated measure and two more analyses with *duration* and *spectral balance* as fixed effects. The results are described below in separate subsections for each way of varying intensity level.

2. Intensity (uniformly distributed gain increments)

In this subsection, the effect of *duration* and *intensity*, the latter varied by spectrally uniform amplification, on $p(init)$ is examined. Figure 1 shows the decrease of the percentage perceived initial stress as a function of duration and intensity level difference. The duration of the first syllable decreases from left to right, while at the same time the duration of the second syllable increases. The intensity scale gives the difference in overall intensity level (IL) between the initial syllable and the final syllable ($IL_{\sigma 1}-IL_{\sigma 2}$). The upper panel displays the results for the stimuli presented over headphones, the lower panel those for the stimuli presented over loudspeakers. This way of presenting the data does in no way mean that we assume the duration and the intensity range to be absolutely identical. However, the similarity of both ranges is that they are both a representative reflection of ranges found in our production data (see Sec. I A 1).

When stimuli are presented over headphones, the whole range of intensity change produces only a slight decrease of *p(init)*: from 65% to 50%. The range of duration change produces a much larger decrease of *p(init)*: from 98% to 8%. *Duration*, *intensity*, and their interaction together explain 97% of the variance. Although the contribution of *intensity* is statistically significant [$F(6,91)=4.9$, $p<0.001$], it is only small compared to that of *duration* [$F(6,91)=315.5$, $p<0.001$]. *Intensity* alone explains a mere 2% of the variance. *Duration* on the other hand, explains as much as 93% of the variance. There is a significant interaction between *duration* and *intensity* [$F(36,49)=1.8$, $p=0.26$], which ex-

plains 2% of the variance. This interaction is due to the fact that overall intensity level variations have little or no influence at the extremes of the duration scale, where judgments are mainly guided by duration differences, whereas they have a larger influence on *p(init)* in the temporally more ambiguous stimuli.

As can be seen in the lower panel of Fig. 1, presenting the stimuli over loudspeakers mainly affects the effectiveness of duration as a stress cue and hardly influences the perceptual contribution of intensity level differences. In this case duration produces a less steeply sloping decrease, from 88% to 22%, whereas intensity again produces a decrease of 15%. Again, the effects of both *duration* and *intensity* are significant [$F(6,91)=90.5$, $p<0.001$ and $F(6,91)=5.1$, $p=0.001$, respectively]. *Duration* explains 80% of the variance and *intensity* 5%. Together with their interaction, they explain 93% of the variance, although the interaction was not significant in this condition [$F(36,49)=1.4$, NS].

Our intermediate conclusion is that intensity level variation, as used in this experiment, implemented by spectrally uniform amplification, is only a minor stress cue, whether stimuli are presented over headphones or over loudspeakers.

3. Spectral balance (intensity level variation by increments in the higher frequency bands only)

Figure 2 shows the decrease of *p(init)* as a function of duration ratio and difference in spectral balance. The duration range is the same as in Fig. 1, but now the intensity level differences are obtained by increasing the levels in the higher frequency bands only. The intensity level scale gives the difference in spectral balance between the initial syllable and the final syllable ($B_{\sigma 1}-B_{\sigma 2}$). Again, the upper panel presents the data of the stimuli presented over headphones, the lower panel of the stimuli presented over loudspeakers.

The whole range of spectral balance produces a decrease of 41%: from 77% to 36% when stimuli are presented over headphones. The duration range produces a decrease of 86%: from 95% to 9%. *Duration, spectral balance* and their interaction together explain 99% of the variance. Both *duration* and *spectral balance* have a significant effect on *p(init)* [*duration*: $F(6,91)=420.5$, $p<0.001$; *spectral balance*: $F(6,91)=73.7$, $p<0.001$]. *Duration* alone explains 76% of the variance, whereas *spectral balance* explains 13% of the variance. The significant interaction between *duration* and *spectral balance* [$F(36,49)=9.0$, $p<0.001$] is again due to the fact that variations in spectral balance have less influence on stress judgments at the extremes of the duration range.

When stimuli are presented over loudspeakers, the effect of *duration* on *p(init)* decreases. However, while intensity (Sec. I C 2) proves equally ineffective through headphones as over loudspeakers, *presentation* strongly influences the relative strength of effort and duration as stress cues. *Duration* and *spectral balance* produce an almost equal decrease of *p(init)*: 80% to 24% for *duration* versus 86% to 20% for *spectral balance*. This, in fact, means that subjects rely more heavily on differences in spectral balance than on duration differences when stimuli are presented over loudspeakers. Both *duration* and *spectral balance* have a highly significant effect on *p(init)* [*duration*: $F(6,91)=77.2$, $p<0.001$; *spectral balance*: $F(6,91)=115.5$, $p<0.001$]. Together with their interaction they explain 96% of the variance. *Duration* alone explains ''only'' 35%, whereas *spectral balance* explains as much as 53%. The significant interaction of *duration* and *spectral balance* [$F(36,49)=3.1$, $p<0.001$] is due to the fact that the more extreme values of one parameter add disproportionally more weight as the other parameter is more ambiguous.

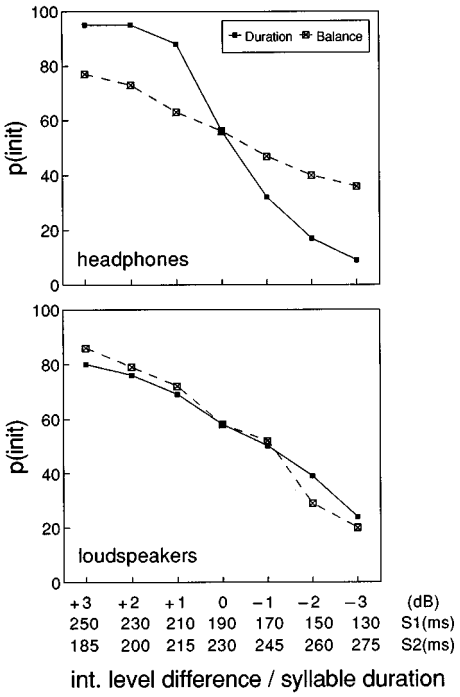


FIG. 2. Percentage of listeners ''initial stress'' judgments, *p*(*init*), for the 91 stimuli *nana* as a function of syllable duration (solid lines) and overall intensity (dashed lines). The differences in *spectral balance* ($B_{\sigma 1}-B_{\sigma 2}$ in dB), obtained by amplification of frequency components above 500 Hz only, are given along the *x* axis, top line. *Duration* values (in ms) are given on the middle and bottom lines for the first and second syllable, respectively. The results are presented for each presentation condition separately: headphones (upper panel) and loudspeakers (lower panel).

We conclude from these results that realistic intensity level manipulations (i.e., mimicking speech production effort by incrementing intensity level in the higher frequency bands only) provide a relatively strong stress cue, and in fact approximate the cue value of duration differences, whereas overall intensity level differences do not provide a substantial stress cue.

Since the reliability of duration as a cue is degraded when the stimuli are presented over loudspeakers, the relative cue value of spectral balance in this situation becomes more important. One explanation could be that subject differences (phonetically trained versus phonetically naive) were responsible for the difference in effectiveness of the duration cue. Of course, an alternative explanation of this interaction is that accurate perception of duration differences suffers from reverberation of the acoustic signal in the room in which the subjects were tested. Locating syllable boundaries in reverberant speech is more difficult since their exact locations are obscured by energy reflections of preceding segments. As a result, the variation in vocal effort became

relatively more important as a stress cue since its acoustical correlate (spectral balance) is not easily affected by reverberation. The experiment reported on in the next section was specifically set up to allow us to choose between the two alternative explanations suggested above.

II. PERCEPTION EXPERIMENT II

A. Effect of ''reverberation'' on the perception of differences in duration and spectral balance

In a room, the acoustic signal produced by either a talker or a loudspeaker may reach a listener by many individual soundpaths. The original speech at the talker's (or loudspeaker's) position and the resulting sound at the listener's position are not identical. Comparing the specific distribution of sound intensity over frequency and time of the original speech with that of the transmitted speech, a certain degree of smearing of the finer details is found: the temporal intensity distribution will be blurred by the combined effects of the many individual soundpaths with various time delays (Houtgast and Steeneken, 1973, 1985; Duquesnoy and Plomp, 1980).

We assume that reverberation, which is a result of myriad reflected sound waves, and is mainly a distortion in the temporal domain, is responsible for the fact that the relative importance of duration as a cue in stress perception decreased when the stimuli were presented over loudspeakers. It has been amply demonstrated that reverberation has a considerable effect on speech intelligibility. These effects appear to be due to the reflections that arrive at the subjects' ear(s) later than about 30 ms after the direct signal, while earlier reflections are integrated with the direct sound (Gelfand and Silman, 1979 and references mentioned there).

In order to rule out alternative explanations for the reverberation effect based on subject differences (see above), we ran a control experiment. We presented both nonreverberant and reverberant stimuli over headphones with the same duration and intensity level manipulations as in the previous experiment and asked subjects *in a within-subjects design* to determine the stress position of each stimulus.

B. Methods

1. Stimulus material

The reverberant stimuli were produced by processing the master test recordings through a Yamaha SPX 90II digital multi-effect processor. The SPX 90II creates a highly natural sounding reverberation. Reverberation time for this particular processor is defined as the length of the time it takes for the level of reverberation at 1 kHz to decrease by 60 dB. Usually natural reverberation varies according to the frequency of the sound: the higher the frequency the more the sound tends to be absorbed by walls, furnishings and even air. We decided not to alter the reverberation time of the high frequencies in proportion to the mid-frequency reverberation time.

We decided to use a reverberation time of 0.6 s for our stimuli. This value was chosen so that an impulse recorded in a sound insulated booth but processed through the SPX 90II sounded and looked more or less identical to an impulse recorded in the reverberant room in which the stimuli were presented in the previous experiment. Figure 3 presents an example of a test item (*Wil je* ˈ*nana zeggen*) with and without artificial reverberation.

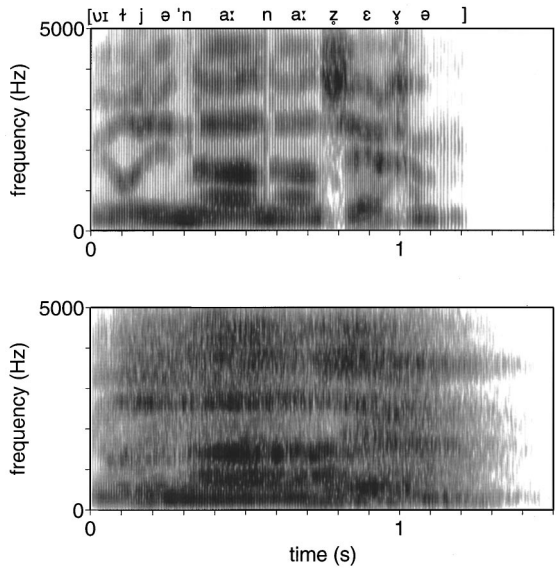


FIG. 3. Example of a test item (*Wil je* ˈ*nana zeggen*) without (upper panel) and with (lower panel) artificial reverberation.

3. Subjects and procedures

A stimulus set was prepared containing the 182 stimuli (91 with and 91 without reverberation) in four different random orders. The third and fourth orders were identical to the first and second, the only difference being that they were recorded in reverse sequence. The 182 stimuli were presented on-line in blocks of 13 utterances with 2-s intervals between utterances and a larger interval between blocks. The procedure was similar to that in the first experiment. Forty-four subjects (staff and students of the Faculty of Arts) participated in the experiment. Seven subjects were phonetically trained and 37 were phonetically naive. The latter subjects were paid for their service. They were tested in four groups in a language laboratory at Leiden University. Each group listened to one of the four different orders. They listened to the stimuli over good quality stereo headphones.

C. Results

1. Global presentation

The reliability of the subjects was determined by relating their individual scores to the composite group score. In order to know how each of them affected the reliability of the group, Cronbach's α was calculated when each of the subjects was removed from the group in turn. We wanted to use the same number of subjects as in the first experiment. We therefore eliminated the four subjects whose exclusions yielded the largest increase of α. Consequently, 40 subjects were used for further analysis.

We determined the number of judgments favoring initial stress for each stimulus and calculated the percentage, *p(init)*. The listening test yielded a total of 7280 responses (182 stimuli * 40 subjects). Overall 56% of the responses

TABLE IV. Main effects and interactions of duration, intensity level, presentation (nonreverberant versus reverberant stimuli), and method (of varying intensity: overall versus high-frequency bands only) on *p(init)*. *F* ratio, significance of *F* and percentage of explained variance (η^2) are given.

Effects	*F*	sign.	η^2
Main effects			
Duration	338.0	<0.001	60
Intensity level	78.8	<0.001	14
Presentation	28.8	<0.001	1
Method of variation	24.5	<0.001	1
Two-way interactions			
Duration * intensity level	2.5	0.004	3
Duration * presentation	53.5	<0.001	9
Duration * method	3.4	0.009	1
Intensity level * presentation	5.4	<0.001	1
Intensity level * method	27.0	<0.001	5
Presentation * method	<1	NS	0
Three-way interactions			
Duration * int. level * presentation	2.9	0.001	3
Duration * intensity level * method	1.4	NS	1
Duration * presentation * method	4.3	0.002	1
Int. level * presentation * method	1.7	NS	0

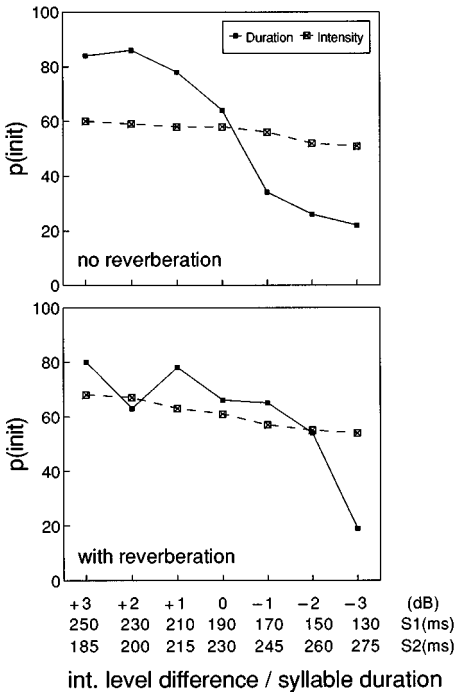


FIG. 4. Percentage of listeners ''initial stress'' judgments, *p*(*init*), for the 91 stimuli *nana* as a function of syllable duration (solid lines) and overall intensity (dashed lines). The differences in *intensity level* ($IL_{\sigma 1}$–$IL_{\sigma 2}$ in dB), obtained by spectrally uniform amplification, are given along the *x* axis, top line. *Duration* values (in ms) are given on the middle and bottom lines for the first and second syllable, respectively. The results are presented for each reverberation condition separately: no reverberation (upper panel) and with artificial reverberation (lower panel).

favored initial stress, which indicates that there is a slight bias for initial stress. This bias is above chance, as determined by a binomial test ($p<0.001$).

As in the previous experiment, we ran a four-way analysis of variance, with *p(init)* as the dependent variable, and with *presentation* (reverberant versus nonreverberant), *method* (adding intensity in all bands versus adding intensity in higher bands only), *duration* (seven steps) and *intensity level* (seven steps) as fixed effects. There were no repeated measures. Since there is no residual variance, the variance caused by the fourth-order interaction was used as the error term. In Table IV the main and interaction effects are given. As can be seen in Table IV, the crucial main effects and interactions are quite similar to those in the previous experiment. There are large effects of both *duration* and *intensity level* on *p(init)*, although the effect of *duration* on *p(init)* is smaller than in the first experiment. The significant main effect of *presentation* indicates that there was a difference in stress bias between reverberant stimuli and nonreverberant stimuli: 59% versus 54%, respectively, which we attribute to the fact that the end of the second syllable of *nana* is more strongly demarcated by the unvoiced fricative [z̥], than the initial syllable, which is succeeded by an identical syllable. Therefore, the perceived length of the initial syllable is possibly more strongly influenced by reverberation than the second syllable.

The significance of the two- and three-way interactions with *presentation* means that reverberation has an influence on both the effect of *duration* and *intensity level* on *p(init)*. Crucially, significant two- and three-way interactions with *presentation* are found similar to the interactions in the first experiment. This indicates that the effect of reverberation is highly comparable to the effect of the way of presentation in the first experiment. This is an indication that reverberation was indeed (at least for the greater part) responsible for the difference in relative importance of *duration* and *spectral balance* between the two presentation conditions. As in Secs. I C 2 and 3, we will now study the main effects of *duration* and *intensity level* in more detail separately for *presentation* (reverberant versus nonreverberant) and *method* (uniform intensity level versus spectral balance). Results are presented in the next subsection.

2. Reverberant versus nonreverberant speech

We ran two separate two-way analyses of variance with *duration* and *intensity* as fixed effects and two more analyses with *duration* and *spectral balance* as fixed effects. There were no repeated measures: only percentages of explained variance but no *F* ratios could be computed.[4] Figure 4 shows the decrease of the percentage perceived initial stress, *p(init)*, as a function of duration ratio and intensity presented as in Fig. 1 with uniform intensity level differences. The upper panel shows the data for the nonreverberant stimuli, the lower panel shows the data for the reverberant stimuli. Figure 5 shows similar data, but now with differences in spectral balance as in Fig. 2.

Figures 4 and 5 show that the effectiveness of duration deteriorates considerably for the reverberant stimuli.[5] As can be seen in Fig. 4, *intensity* does not serve as a stress cue at all for the nonreverberant stimuli. The effectiveness of this cue slightly increases for the reverberant stimuli. This tendency was also observed in the previous experiment.

The results for *spectral balance* (Fig. 5) are comparable to those in the previous experiment: again a considerable

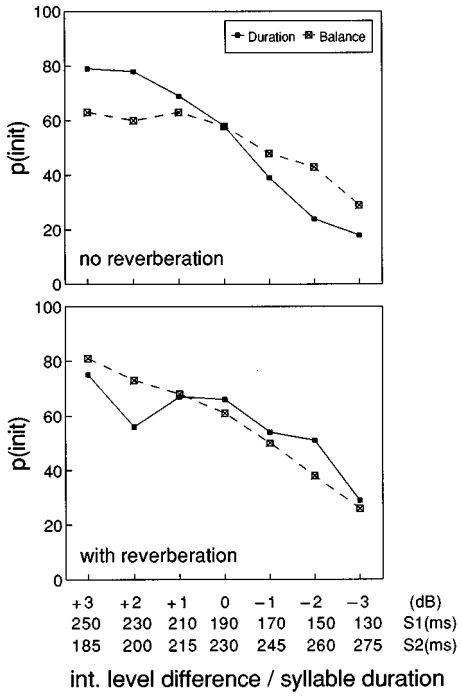


FIG. 5. Percentage of listeners ''initial stress'' judgments, $p(init)$, for the 91 stimuli *nana* as a function of syllable duration (solid lines) and overall intensity (dashed lines). The differences in *spectral balance* ($B_{\sigma 1}-B_{\sigma 2}$ in dB), obtained by amplification of frequency components above 500 Hz only, are given along the *x* axis, top line. *Duration* values (in ms) are given on the middle and bottom lines for the first and second syllable, respectively. The results are presented for each reverberation condition separately: no reverberation (upper panel) and with artificial reverberation (lower panel).

increase in effectiveness of *spectral balance* is found for the reverberant stimuli.

In the next section we will compare the results of both experiments in more detail.

III. COMPARISON OF EXPERIMENTS 1 AND 2

In Table V, we present an overview of the percentages explained variance for duration, intensity and spectral balance in both experiments to compare the relative strength of the stress cues in both experiments. The left-hand part of the table presents the data for experiment 1, in which stimuli were presented to half of the subjects over headphones and to half of the subjects over loudspeakers (*separate conditions*). The right-hand part of the table presents the data of the present experiment (2), in which both reverberant and nonreverberant stimuli were presented in a *within-subjects design* over headphones (*mixed condition*).

TABLE V. Relative strength of stress cues (in % explained variance η^2) in reverberant and nonreverberant stimuli, presented in separate and mixed conditions (experiment 1 and experiment 2, respectively).

	Experiment 1 separate conditions		Experiment 2 mixed condition	
	Overall int. (Fig. 1)	Spectral balance (Fig. 2)	Overall int. (Fig. 4)	Spectral balance (Fig. 5)
No reverb				
Duration	93	76	94	73
Intensity level	2	13	0	18
Dur. * int.	2	10	6	9
Residue	3	1	…	…
Reverb				
Duration	80	35	84	35
Intensity level	5	53	6	57
Dur. * int.	8	8	10	8
Residue	7	4	…	…

As can be seen in Table V the percentages explained variance in both experiments are almost identical. We conclude on the basis of these results that duration indeed suffered from reverberation and that reverberation was therefore responsible for the relative increase in effectiveness of spectral balance when stimuli were presented over loudspeakers.

In the present experiment, variations in duration did not lead to an equally large change in *p(init)* as in the previous experiment. In the nonreverberant speech condition, *p(init)* decreased with roughly 60% from about 80% to 20%, whereas in the previous experiment in this condition a range was covered between 98% and 8%. This could possibly be due to the fact that reverberant and nonreverberant stimuli were presented in random succession, which might have prevented our listeners from tuning in to one specific speech type.[6]

In summary, the importance of duration as a cue to stress perception decreased under reverberation ($T=0.6$ s), whereas the relative contribution of spectral balance manipulations increased strongly. The magnitude of the effects in both experiments were in the same range. The effectiveness of overall intensity, however, was hardly affected by reverberation and was equally poor in both experiments. On the basis of these results we conclude that the use of duration as a cue for stress suffers from reverberation. As a result, loudness (as a reflection of vocal effort) becomes relatively more important as a stress cue showing that its acoustical correlate (spectral balance) is not easily affected by reverberation.

IV. GENERAL DISCUSSION AND CONCLUSIONS

In this study we reconsidered the general claim that loudness is a weak cue in the perception of stress. This traditional claim was based on perception experiments in which loudness was varied in a naive way: All parts of the spectrum were amplified uniformly. We hypothesized that varying loudness more realistically will make it a stronger stress cue, and that we could possibly rehabilitate the traditional claim that languages such as Dutch and English have dynamic (rather than melodic or temporal) stress.

From the results of both experiments, we conclude that loudness implemented as a difference in overall intensity level (i.e., manipulating gain without changing spectral balance) provides only a marginal stress cue. Of course, we need not be surprised that intensity level variations turn out to provide only a marginal stress cue. In fact, it seems to us that intensity level variation will never have communicative significance, for the simple reason that intensity level is too susceptible to noise. If the speaker accidentally turns his head, or passes a hand across his mouth, intensity level drops of greater magnitude than those caused by the difference between stressed and unstressed syllables will easily occur. For this reason, manipulating intensity in stress perception

experiments seemed ill-advised. The reason why it was used in the classical studies by Fry (1955) and Mol and Uhlenbeck (1956) must have been that there were simply no alternatives available for investigating the role of loudness in stress perception.

In contrast, loudness realistically implemented as the acoustical reflection of greater vocal effort, is a reliable stress cue, close in strength to duration. Moreover, the differences in spectral balance provide an even stronger stress cue than duration when accurate perception of syllable and segment boundaries is hampered, for instance in a reverberant environment. Examples of such reverberant listening conditions in daily life abound. In fact, studying speech communication in rooms, halls etc. is probably more realistic than in sound-insulated booths and free-field situations. Therefore, it seems that listeners have different cooperating cues at their disposal to determine linguistic stress position. The effectiveness of the different cues depends on environmental circumstances in which speech is perceived.

Results of a perception experiment carried out by Beckman (1986) for English and Japanese showed that these two languages differed greatly as to the relative importance of $F0$, duration and loudness as perceptual cues to stress. Both Japanese and English listeners were presented with disyllabic words in which all these parameters were varied according to production data. Japanese is an archetypal nonstress-accent language, a so-called pitch-accent language, with $F0$ as the most consistent acoustical correlate of stress/accent. English is an archetypal stress-accent language with the same acoustical correlates of stress and accent as Dutch. The comparison between English and Japanese listeners showed that Japanese listeners seemed to rely heavily on differences in $F0$ and they hardly used any of the other cues. English listeners also relied heavily on $F0$, although to a much lesser extent. Loudness, however, was also found to be a very effective cue for English listeners in stress perception. Loudness in this experiment was operationalized as ''total amplitude,'' a measure of power integrated over the entire duration of the vocalic nucleus (i.e., energy), rather than as peak intensity. Beckman assumes this measure to be closely related to loudness and she attributes the success of this cue to this relation:

> Thus the total amplitude may be a better correlate of stress than is either duration or intensity alone and it may be a more consistent perceptual cue simply because it is a better measure of loudness,... (Beckman, 1986, p. 197).

In our view this measure of loudness is equally unrealistic as overall intensity level manipulations are. Beckman in fact measured the combined effect of peak intensity and duration. It is therefore no surprise that this measure yields considerably better results than either duration or peak intensity alone. It has only been established for pure tones of a relatively short duration that differences in duration are responsible for differences in the perception of loudness. Although the literature agrees about the fact that there is a certain threshold value above which duration changes no longer influence loudness, the literature largely disagrees as to determining the exact value of this threshold. However, despite the great variability of results regarding the threshold value among the various studies, they largely agree on the fact that temporal integration of energy occurs at very short durations (Beckman, 1986 and references mentioned there). Therefore, although this measure may have some relevance for plosives (i.e., the longer a noise burst, the louder it is perceived), it has no relevance for vowels and sonorants, since these sounds are no short acoustic events. Therefore, in our view, this operationalization of loudness has no relevance in vocalic nuclei.[7]

In our view, the ultimate test to investigate whether English listeners are more sensitive to loudness than Japanese listeners, would be to synthesize similar stimuli as used in Beckman (1986) while separating focused and nonfocused material and varying loudness in the way described in the present article. If it is indeed true that languages such as Dutch and English have dynamic accent as opposed to pitch accent in languages such as Japanese, Japanese listeners will be insensitive to these more realistic loudness manipulations as well, whereas English listeners would make considerable use of these differences.

In addition to the above mentioned, more linguistically oriented implications, the findings of the present study have some more practical, application-based implications as well. The results can probably be used to improve the quality of speech synthesis. In future research, experiments should be executed investigating if stress and focus domains could be more optimally synthesized if we take the present results into account. There are elaborate rule-sets in Dutch text-to-speech systems to predict whether or not a word should be accented (Quené and Kager, 1993; Dirksen and Quené, 1993). If a word is accented, all its syllables, stressed as well as unstressed, should be lengthened, at least in Dutch, relative to syllables of a word that remains unaccented (Eefting, 1991; van Heuven, 1993). The stressed syllables of both accented and unaccented words should be marked by a combination of (extra) longer duration[8] and greater loudness. The present experiments showed that the relative importance of these cues depends on the listening circumstances; it is therefore necessary to represent both cues optimally in synthetic speech to guarantee adequate stress perception independent of listening circumstances especially because for unaccented words these cues are the only remaining cues to stress. Furthermore, in our experiment stress was varied so as to reflect production data. However, intensity level, spectral balance and duration could be combined in a more extreme way, for instance by both adding and shifting intensity levels. Listeners could probably prefer more strongly marked stress positions when listening to synthetic speech, because of the fact that there is not always a one-to-one mapping of what speakers do and what listeners want. The intelligibility of synthesized speech in text-to-speech systems could possibly improve by a more accurate marking of stress and accent since the former facilitates the recognition of words in continuous speech (cf. van Heuven, 1988), while the latter prompts the listener to give priority to bottom-up processing exactly there where it matters (cf. Terken and Nooteboom, 1987; van Donselaar, 1995).

We assumed the differences in spectral balance to be caused by a more pulse-like shape of the glottal waveform while producing stressed syllables. Future manipulations could be made even more realistically by manipulating the glottal pulse separately instead of using digital filtering of the oral output.

To conclude this paper, the most important finding of this study is that listeners are more susceptible to intensity level variations when detecting stress position than hitherto has been assumed. This is due to the fact that intensity level differences in our experiments were implemented in a more realistic way, i.e., by amplification in the higher frequency bands only, as the acoustical reflection of an increase in vocal effort used to produce stressed syllables. The results can be viewed as a first step to rehabilitation of the claim that languages such as Dutch and English have dynamic stress, with perceived loudness as its most reliable cue.

ACKNOWLEDGMENTS

Portions of this research have been presented at the ESCA workshop on Prosody, Lund (September 1993) and at the 127th meeting of the Acoustical Society of America, Cambridge, MA (June 1994). The authors would like to thank H. Traunmüller, J. W. de Vries, S. G. Nooteboom, and one anonymous reviewer for comments on earlier versions of this paper, ideas, and discussion.

[1]It was pointed out to us by Hartmut Traunmüller that ''spectral emphasis'' might be a better term. We agree, but stick to the term ''spectral balance'' to insure terminological uniformity with our earlier publications.

[2]Note, however, that we used Dutch words. There is no guarantee that English will behave like Dutch. As a case in point, a Dutch word spoken without a pitch accent is pronounced some 15% faster than its accented counterpart (linear time compression, cf. Eefting, 1991; Sluijter and van Heuven, 1995). A similar experiment showed that only the accented foot, but not the entire word, is time-expanded in American English (Turk and Sawush, 1995).

[3]A similar analysis was performed on the arcsine transformed percentages (cf. Studebaker, 1985). There were no crucial differences, so we decided to use the nontransformed percentages in all the analyses performed on the data in this paper.

[4]In this type of situation it is not uncommon to adopt the highest interaction as the numerator term. The second-order interaction is the only interaction in this analysis and it is inherent to this type of experiment that this interaction plays a systematic role: When one cue is ambiguous the other one becomes more important; consequently, the interaction is not a suitable numerator term. Since the primary goal of this analysis is to quantify the relative magnitude of the effects (the significance of which has been shown in earlier experiments), rather than to determine the significance of the effects, we decided to refrain from any significance testing at all.

[5]Unexpectedly, in this condition subjects hardly used duration as a cue in duration step 2 (230–200), whereas they heavily relied on duration in step 3. We do not have an explanation for this effect and we assume that there is some unknown acoustic interference of reverberation and duration in some of the stimuli.

[6]Besides, the subjects were mainly students who had never participated in listening experiments before. The results of the 20 most reliable subjects, as determined with Cronbach's α, cover a much larger range, comparable to the range covered in the first experiment. The seven phoneticians who participated in the present experiment all belonged to this group. This means that subject differences could partly be held responsible for the distortion of the duration results.

[7]Beckman (1986) did not consistently separate focused and nonfocused material. The relative strength of $F0$ may therefore be overestimated, in any case in English and probably also in Japanese.

[8]The relative importance of stress cues may differ from language to language. Specifically stress cues such as duration (Berinstein, 1979) and pitch (Potisuk *et al.*, 1996) assume a lower position in the rank order of cues as these parameters are simultaneously exploited in other linguistic contrasts (vowel quantity and lexical tone, respectively).

Beckman, M. E. (**1986**). *Stress and Non-Stress Accent* (Foris, Dordrecht).

Berinstein, A. E. (**1979**). ''A Cross-linguistic study on the perception and production of stress,'' *Working Papers in Phonetics* (University of California, Los Angeles), No. 47.

Bloomfield, L. (**1933**). *Language* (Holt, Rinehart and Winston, New York).

Brandt, J. F., Ruder, K. P., and Shipp, Jr., I. (**1969**). ''Vocal Loudness and Effort in Continuous Speech,'' J. Acoust. Soc. Am. **46**, 1543–1548.

Dirksen, A., and Quené, H. (**1993**). ''Prosodic analysis: The next generation,'' in *Analysis and Synthesis of Speech: Strategic Research Towards High-Quality Text-To-Speech Generation*, edited by V. J. van Heuven and L. C. W. Pols (Mouton de Gruyter, Berlin), pp. 131–144.

Donselaar, W. van (**1995**). ''Effects of accentuation and given/new information on word processing,'' doctoral dissertation, Utrecht University, Utrecht.

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. **68**, 537–544.

Eefting, W. Z. F. (**1991**). ''The effect of information value and accentuation on the duration of Dutch words, syllables, and segments,'' J. Acoust. Soc. Am. **89**, 412–424.

Fry, D. B. (**1955**). ''Duration and intensity as physical correlates of linguistic stress,'' J. Acoust. Soc. Am. **27**, 765–768.

Fry, D. B. (**1958**). ''Experiments in the perception of stress,'' Lang. Speech **1**, 126–152.

Fry, D. B. (**1965**). ''The dependence of stress judgments on vowel formant structure,'' in *Proceedings of the 5th International Congress on Phonon Science, Münster 1964* (Karger, Basel), pp. 306–311.

Gelfand, S. A., and Silman, S. (**1979**). ''Effects of small room reverberation upon the recognition of some consonant features,'' J. Acoust. Soc. Am. **66**, 22–29.

Glave, R. D., and Rietveld, A. C. M. (**1975**). ''Is the effort dependence of speech loudness explicable on the basis of acoustical cues?,'' J. Acoust. Soc. Am. **58**, 875–879.

Hart, J. 't, Collier, R., and Cohen, A. (**1990**). *A Perceptual Study of Intonation; An Experimental-Phonetic Approach to Speech Melody* (Cambridge U.P., Cambridge).

Hermes, D. J., and Rump, H. H. (**1993**). ''The role of pitch in lending prominence to syllables,'' in *Proceedings of an ESCA Workshop on Prosody*, Working papers **41**, Department of Linguistics, Lund University, edited by D. House and P. Touati, pp. 28–31.

Heuven, V. J. van (**1987**). ''Stress Patterns in Dutch (Compound) Adjectives: Acoustic Measurements and Perception Data,'' Phonetica **44**, 1–12.

Heuven, V. J. van (**1988**). ''Effects of stress and accent on the human recognition of work fragments in spoken context: gating and shadowing,'' in *Proceedings of the 7th FASE/Speech-88 Symposium*, edited by W. A. Ainsworth and J. N. Holmes (The Institute of Acoustics, Edinburgh), pp. 811–818.

Heuven, V. J. van (**1993**). ''On the temporal domain of focal accent,'' in *Proceedings of an ESCA Workshop on Prosody*, Working papers **41**, Department of Linguistics, Lund University, edited by D. House and P. Touati, pp. 132–135.

Heuven, V. J. van, and Menert, L. (**1996**). ''Why stress position bias?,'' J. Acoust. Soc. Am. **100**, 2439–2451.

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.

Houtgast, T., and Steeneken, H. J. M. (**1985**). ''A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria,'' J. Acoust. Soc. Am. **77**, 1069–1077.

Issatchenko, A. V., and Schädlich, H. J. (**1966**). ''Untersuchungen über die deutsche Satzintonation,'' Stud. Grammatica **7**, 7–64.

Katwijk, A. van (**1974**). *Accentuation in Dutch; An Experimental Linguistic Study* (Van Gorcum, Amsterdam).

Lagendijk, M. (**1992**). *The XAudlab User Manual* (Speech Processing Expertise Centre, Leidschendam, The Netherlands).

Larkey, L. S. (**1982**). ''Reiterant speech: An acoustic and perceptual validation,'' J. Acoust. Soc. Am. **73**, 1337–1345.

Lehiste, I. (**1970**). *Suprasegmentals* (MIT, Cambridge, MA).

Lehiste, I., and Peterson, G. E. (**1959**). ''Vowel amplitude and phonemic stress in American English,'' J. Acoust. Soc. Am. **31**, 428–435.
Mol, H. G., and Uhlenbeck, G. M. (**1956**). ''The linguistic relevance of intensity in stress,'' Lingua **5**, 205–213.
Morton, J., and Jassem, W. (**1965**). ''Acoustic correlates of stress,'' Lang. Speech **8**, 148–158.
Peterson, G. E., and Lehiste, I. (**1960**). ''Duration of syllable nuclei in English,'' J. Acoust. Soc. Am. **32**, 693–703.
Potisuk, S., Gandour, J., and Harper, M. P. (**1996**). ''Acoustic correlates of stress in Thai,'' Phonetica **53**, 200–220.
Quené, H., and Kager, R. (**1993**). ''Prosodic sentence analysis without parsing,'' in *Analysis and Synthesis of Speech: Strategic Research Towards High-Quality Text-to-Speech Generation*, edited by V. J. van Heuven and L. C. W. Pols (Mouton de Gruyter, Berlin), pp. 115–130.
Rietveld, A. C. M., and Koopmans-van Beinum, F. J. (**1987**). ''Vowel reduction and stress,'' Speech Commun. **6**, 217–229.
Sluijter, A. M. C., and Heuven, V. J. van (**1995**). ''Effects of Focus Distribution, Pitch Accent and Lexical Stress on the Temporal Organization of Syllables in Dutch,'' Phonetica **52**, 71–89.
Sluijter, A. M. C., and Heuven, V. J. van (**1996**). ''Spectral balance as an acoustic correlate of linguistic stress,'' J. Acoust. Soc. Am. **100**, 2471–2485.
Studebaker, G. A. (**1985**). ''A ''rationalized'' arcsine transform,'' J. Speech Hear. Res. **28**, 455–462.
Sweet, H. (**1906**). *A Primer of Phonetics* (Clarendon, Oxford).
Terken, J. M. B., and Nooteboom, S. G. (**1987**). ''Opposite effects of accentuation and deaccentuation on verification latencies for 'given' and 'new' information,'' Language Cogn. Process. **2**, 145–163.
Turk, A. E., and Sawush J. (**1995**). ''The domain of the durational effects of accent,'' MIT Speech Comm. Group Working Papers **X**, also J. Phon (to appear).
Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P. J. (**1992**). ''Segmental durations in the vicinity of prosodic phrase boundaries,'' J. Acoust. Soc. Am. **91**, 1707–1717.

Prosodic features at discourse boundaries of different strength

Marc Swerts[a)]
Institute for Perception Research (IPO), P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands

(Received 22 November 1995; revised 17 June 1996; accepted 27 August 1996)

This paper presents the design and the evaluation of a method to study prosodic features of discourse structure in unrestricted spontaneous speech. Past work has indicated that one of the major difficulties that discourse prosody analysts have to overcome is finding an independent specification of hierarchical discourse structure as to avoid circularity. Previous studies have tried to solve this problem by constraining the discourse or by basing segmentations on a specific discourse theory. The current investigation first explores the possibility of experimentally determining discourse boundaries in unrestricted speech. In a next stage, it is investigated to what extent boundaries obtained in this way correlate with specific prosodic variables: the features pause, pitch range, and type of boundary tone are studied as a function of discourse structure. © *1997 Acoustical Society of America.* [S0001-4966(97)01701-3]

PACS numbers: 43.71.Es, 43.70.Fq, 43.70.Hs, 43.70.Bk [RAF]

INTRODUCTION

It is intuitively clear that a coherent discourse exhibits a certain syntax, comparable to that at the sentence level. This implies that spoken or written texts consist of larger-scale information units, or discourse segments, which can be viewed as building a hierarchical structure. That is, segments may contain or may be embedded in others: for instance, someone may talk about his holidays, with subtopics on hotel, food, etc. The structure is generally reflected in the distribution of particular linguistic phenomena such as anaphora (Grosz and Sidner, 1986; Reichman, 1985; Webber, 1988; Geluykens, in press) or discourse particles (Shiffrin, 1987; Passonneau and Litman, 1993). It has also its consequences for prosody, which is investigated in this paper, since such features as speech melody, pause and tempo are likely candidates to highlight the make-up of a spoken text.

However, investigators who study linguistic, for instance, prosodic, correlates of discourse structure are confronted with a serious methodological problem. Ideally, one would like to obtain an ''independent'' specification of junctures in the information flow and of the mutual relationships between segments. It is of course known that prosody may influence subjects' perceptions of discourse structure, so that it is legitimate that investigators use it as one source of information to determine discourse structure. But it should be guaranteed that the identification of boundaries does not solely depend on prosodic considerations, to minimize the risk of circularity. In order to make them useful for prosody research, the criteria developed to mark discourse also need to be explicit and reproducible, and need to be more than one individual's intuitive notion of information structure.

In practice, however, it is unclear to what extent such autonomous and reproducible macro-labeling can be achieved on the basis of existing discourse theories. Although these may start from legitimate assumptions, taking, for instance, coherence relations (Mann and Thompson, 1986) or topicality (van Kuppevelt, 1995; Brown and Yule, 1983) as a basic organizing principle of discourse, they often lack a degree of explicitness to enable (naive) subjects to reliably label a given text in terms of structural features. Coding schemes to annotate dialogue structure in terms of utterance functions have been developed and tested in the HCRC Map Task Corpus (Carletta *et al.*, 1996a, 1996b) and the VERBMOBIL project (Reithinger, 1995; Jekat *et al.*, 1995). However, there are virtually no empirical investigations, except for a few in the Grosz and Sidner framework (see below), to test to what extent subjects are able to consistently identify ''discourse segments'' using theory-based guidelines. On the other hand, story grammars such as the ones by Thorndyke (1977) and Mandler and Johnson (1977) do present generative grammars, with explicit rewrite rules defining units and relationships, but those schemata are very much genre specific, i.e., typical for a particular type of narrative prose, so that their usefulness as a general system for discourse labeling remains questionable.

There have been a few attempts reported in the literature to overcome the problem of independently specifying discourse structure. In a first line of research, the whole issue is somewhat circumvented by looking at constructed speech materials. One group of researchers has looked at read-aloud texts with predetermined paragraph boundaries (for instance, Thorsen, 1985, 1986; Lehiste, 1975; Sluijter and Terken, 1994; Brubaker, 1972; Bruce, 1982); similarly, others have focused on tightly constrained types of spontaneous speech, by experimentally eliciting discourse in such a way that it becomes easily segmentable in consecutive information units (Terken, 1984; Swerts and Collier, 1992; Swerts *et al.*, 1994; Swerts and Geluykens, 1994; Venditti and Swerts, 1996). In this way, prosodic features of discourse segments can be adequately investigated. It was indeed found that larger-scale information units have an impact on the distribution of accents, and prosodic characteristics such as pause, globally declining melodic patterns, local boundary tones, and durational variation. These studies are limited, though, in that the structures investigated, while being controlled, are overly simple. An intrinsic danger, therefore, is that the role of prosody as a means to signal information structure is exaggerated, since there are no ''disturbing'' factors of interac-

[a)]Electronic mail: swerts@natlab.research.philips.com

tive or attitudinal nature. It is thus an open question to what extent findings from these experimental studies can be generalized to more complex types of discourse.

The second approach is theory-based, in an attempt to motivate segmentations on the basis of an explicit model of discourse structure, i.e., the one proposed by Grosz and Sidner (1986). In studies by Grosz and Hirschberg (1992; see also Hirschberg and Grosz, 1992) and Passonneau and Litman (1993), seven subjects were instructed to segment a set of monologues, using speaker intention as a criterion. It turns out that there is some variation between labelers as ''no two segmentations were identical'' (Hirschberg and Grosz, 1992, p. 443). In particular the specification of hierarchical relationships between segments appears to be difficult. It is decided in these studies either to concentrate on only those structural features agreed upon by all labelers (Grosz and Hirschberg, 1992), or retain those boundaries assigned by at least four out of seven labelers (Passonneau and Litman, 1993). The studies report that the resulting discourse segments have clear prosodic features in terms of pitch range (Grosz and Hirschberg, 1992) and pause (Grosz and Hirschberg, 1992; Passonneau and Litman, 1993). However, from these investigations it cannot be decided whether prosody can also be exploited to directly signal the level of embeddedness of a given segment. Passonneau and Litman (1993) have limited their study to a purely linear intention-based segmentation task, since more complex segmentation was thought to be too cumbersome, given the average length of their narratives. Grosz and Hirschberg (1992), on the other hand, do try to conduct an empirical study of hierarchies, but do not really use these hierarchical representations in the phonetic analyses afterward, since boundaries at different levels of embeddedness are collapsed to represent one category.

There have been a few more efforts in the mean time to explore the possibilities of determining discourse structure within the framework of Grosz and Sidner (Hirschberg *et al.*, 1995; Nakatani *et al.*, 1995). Even a manual with instructions for segmenting discourse has been produced, which is based on this particular theory and which is meant for ''naive'' subjects, but this guide still needs to be further evaluated (Nakatani *et al.*, 1995). Despite all these previous attempts, the current paper for two reasons introduces yet another method to establish discourse structure. First, although newer versions of the labeling manual will probably lead to more consistent results, it will probably not yield complete agreement on discourse structure, certainly when subjects are confronted with difficult, incoherent texts (Condon and Čech, 1996). Second, there are currently no generally accepted techniques available to assess the reliability of different discourse labelings. Carletta (1996) criticizes the measures used in many other discourse studies to evaluate the agreement between labelers, and proposes a so-called kappa statistic as an alternative.

Therefore, to arrive at an independent discourse labeling in view of prosody research, this paper addresses another approach, which—as will become clear in the next section—exploits the fact that subjects vary in the way they agree on discourse segmentations. It is partly inspired by Rotondo (1984), although he only considered labelings of texts, and not speech or transcribed speech. The goals of our study are twofold: first, it will be explored to what extent the experimentally oriented method offers a useful alternative to already existing procedures; second, it is investigated to what extent boundaries obtained in this way correlate with specific prosodic variables. More specifically, the paper studies whether prosody may play a role in signaling hierarchical discourse structure.

I. METHOD

Rotondo (1984) starts from the idea that from group segmentation data one can derive a hierarchical discourse structure. When subjects, for instance, are asked to mark where one ''complete thought'' has ended and another one begins, they will expectedly have some uncertainty about what constitutes such a unit. Consequently, such a task will lead to inconsistencies among subjects, since some will put a lot of boundaries, whereas others only a few, or in different places. It is logical to assume, however, that there will be less disagreement about stronger breaks since they present clearer transitions in the flow of information. It can thus be expected that more people will feel inclined to mark these as a boundary. Therefore, instead of taking the variance between labelers as a disadvantage, one can rather exploit it to specify hierarchically different discourse boundaries. To arrive at this goal, the segmentations of relatively many labelers are needed. Basically, boundary strength is then computed as the proportion of subjects agreeing on a given break, assuming that the degree of certainty about a given discourse boundary reflects its level of embeddedness. At the same time, the task of simply having to mark the boundary between two consecutive units is relatively easy and less time consuming, compared to a technique where subjects are asked to express boundary strength on, for instance, a ten-point scale, which was previously used to study prosodic phrasing at the sentence level (de Pijper and Sanderman, 1994).

The speech materials used in the present study were taken from a larger corpus of 30 spontaneous Dutch monologues (Beun, 1991). These were originally collected in an experiment in which subjects had to describe a set of paintings. It is important to notice that the resulting narratives were not controlled in terms of discourse structure, since, in principle, speakers had the freedom to organize their descriptions as they liked. Twelve monologues, i.e., six descriptions produced by two speakers (MM and LK), amounting to 46.5 min of speech in total, were selected for further analysis as they were comparable in length (3 to 5 min), and the speakers were both female, using about the same pitch range (which could facilitate prosodic comparisons afterward).

Those selected descriptions were used in a task that was individually performed by 38 subjects, who were instructed to mark paragraph boundaries in transcriptions of the monologues. These were presented without punctuation or specific layout to indicate paragraph structure. Subjects were told to draw a line between the word that ended one paragraph and the one that started the next paragraph. No explicit definition of a paragraph was given.

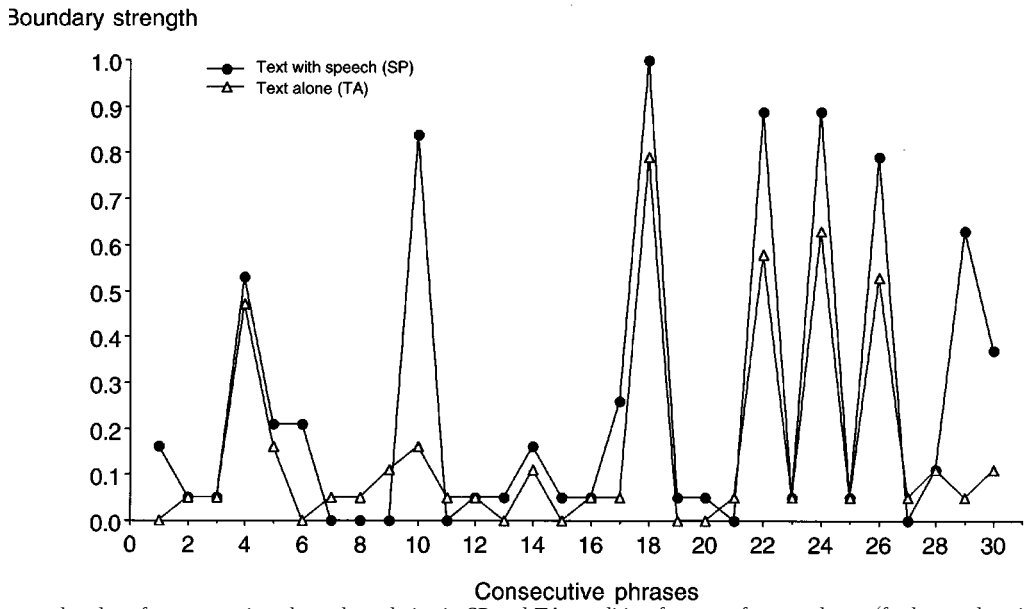


FIG. 1. Boundary strength values for consecutive phrase boundaries in SP and TA condition for part of a monologue (further explanations in text).

There were two conditions: half of the subjects could listen to the actual speech (SP condition), whereas the other half only saw the written version (TA condition). The reason to have both these conditions, a procedure also used by Grosz and Hirschberg (1992), was to gain insight into the added value of prosody. The underlying hypothesis is that prosody helps the listener to better understand the discourse so that speech is comparably less ambiguous than text alone.

A typical example of part of a text is given below, followed by a literal translation in English. The two digits in round brackets represent the boundary strength estimates, computed as the proportion of subjects indicating that there was a break, for the SP and TA condition, respectively. For sake of presentation, boundaries of strength 0 in either of the two conditions are only given when there is a stronger break in the other condition.

> het is echt een paard dat [uh] over iets heen springt heel heel snel (0.26; 0.11) de man die d'r opzit die zit ook helemaal in zo'n gebogen [uh] [uh] ruitershouding met zijn billen omhoog en zijn [uh] hoofd (0;0.05) in de manen van het paard (0.95;0.16) het paard is wit (0.11;0) [uh] ruiter is is [uh] rozig rood (0.53;0.79)
>
> (it is really a horse that [uh] jumps across something very very fast the man who sits on it he really sits also in such a bent over [uh] [uh] rider's position with his bum in the air and his [uh] head stuck in the mane of the horse the horse is white [uh] rider is is [uh] pinkish red)

As can be seen, the breaks between word clusters may vary from relatively weak ones (for instance, 0.05) to relatively strong ones (0.95).

In principle, as subjects were not given any constraints on how they should segment the discourse, every word boundary was a potential candidate for a breaking point between two paragraphs. But from the previous example, it is clear that one can distinguish some ''minimal units,'' i.e., sequences of words not separated by any of the labelers. It appeared that almost all the boundaries of those units coincided with prosodic phrase boundaries, although minimal units could contain more than one phrase; therefore, in the rest of this paper, the prosodic phrase will be taken as a unit of analysis. To this end, prosodic phrases were marked by one experienced phonetician (other than the author), who did not have access to the segmentation results of the 38 labelers. She was asked to assign both the boundaries of phrases and the kind of boundary tone (high, mid, or low) at their respective ends. No explicit definition was given of either prosodic phrase or boundary tone. Her labelings resulted in a total of 889 phrases for the two monologues (535 for LK and 354 for MM). In taking the prosodic phrase as an unit of analysis, the study becomes similar to that of Passonneau and Litman (1993) and Hirschberg and Grosz (1992) who restricted subjects to placing boundaries between prosodic phrases.

To check whether the proportion of subjects agreeing on a break can truly be seen as a measure of boundary strength, one monologue of speaker MM was used as input for a small comparison test. The monologue, split into consecutive phrases, was presented to 20 subjects, none of whom had participated in the previous labeling experiment. They were asked to assign a number to each transition between two consecutive phrases using a ten-point scale, with ''1'' meaning ''very weak boundary'' (word boundary) and ''10'' meaning ''very strong boundary'' (paragraph boundary). As in the previous test, half of the subjects could listen to the speech (SP condition), whereas the other half had but the text to label the discourse (TA condition). It appears that this method of establishing the depth of a discourse boundary gives results which are very comparable to the one based on proportions of subjects. The correlations between the ratings on the ten-point scale (averages of ten labelers) and the proportions, are highly significant, both in the SP ($r=0.75$, $n=30$, $p<0.0001$) and in the TA condition ($r=0.70$, $n=30$, $p<0.0001$). In other words, the assumption that the degree of

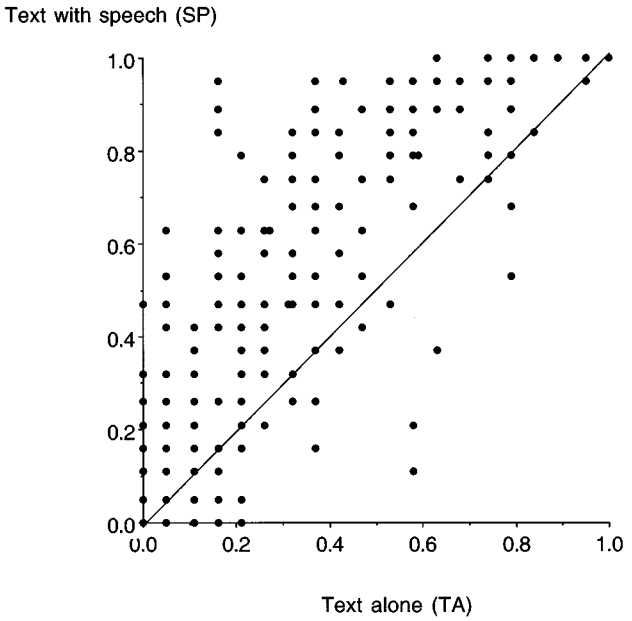


FIG. 2. Scattergram of boundary strength values in the SP and TA condition.

certainty about a given discourse boundary reflects its level of embeddedness appears to be valid.

II. RESULTS

A. Comparing text-with-speech (SP) and text-alone (TA)

A first idea of the segmentation results and of the differences for the two conditions can be derived from Fig. 1, which shows the boundary strength values for part of one typical monologue by speaker MM. The figure reveals that the two experimental conditions produce results which have both similarities and differences. First inspection reveals that the two conditions are comparable in the sense that they both have major breaks in similar positions; they are different in that segmentation is clearer in the text-with-speech case: stronger breaks are more clearly visible in the SP condition as proportionally more subjects agree on a paragraph transition. Also, focusing on the passage around phrase number 10 in Fig. 1 shows that in the text-alone condition, different subjects may share the impression that within a ''window'' of a few phrases, there is a break, but they disagree on the exact location of this boundary. Such sections which receive different segmentations in the TA condition (indicating structural ambiguity) often appear to be unambiguous when subjects have access to speech.

This is supported quantitatively in Fig. 2, which gives a scattergram of the values obtained in the SP and the TA condition. The Pearson correlation coefficient between the responses is relatively high: $r=0.89$ ($n=889$, $p<0.001$) which shows that there is a strong association between the two conditions. This is even true when the 0 boundaries are left out of the comparison, as they form an overwhelming class: $r=0.82$ ($n=252$, $p<0.001$). The data points in this figure that fall on the diagonal represent the scores that were identical in the two conditions. As a general picture, however, one can observe that most of the scores are located above this line, meaning that in the SP condition systematically more subjects agree on a paragraph transition.

As a particular data point in Fig. 2 may contain more than one value, it is not visible from this plot how many values fall within a certain range. This becomes obvious, however, when looking at Fig. 3 which, for the two conditions, gives a distribution of the strength estimates. These are clustered into five groups, i.e., one cluster containing values for phrase boundaries which up to 25% of the labelers marked as a paragraph transition, the next cluster having agreements between 25% and 50%, etc. Phrases within minimal units, i.e., units not subdivided by any of the labelers, were taken as a separate category, since they formed a relatively large group. First, the figure reveals—as expected—that in both conditions you get relatively few ''strong'' boundaries and relatively many ''weak'' ones. At the same time, it appears that the SP condition compared with the TA condition has comparatively more values in the extreme clusters, i.e., the 0 boundaries and the ones of strength >0.75. Apparently, when subjects have access to speech, they are surer about both strong and zero boundaries, whereas in the TA condition, the scores are more diffuse. In other words, this indicates that people have clearer segmentations in the SP condition. This is supported with a chi square test, which shows that the two distributions are statistically different ($\chi^2=40.36$, $p<0.001$, $df=4$): the effect is mainly due to the differences in the <0.25 and the >0.75 classes.

B. Phonetic analyses

The previous section has brought to light that segmentation is clearer when subjects have access to speech. To explain this, one need to learn more about how particular utterances are actually spoken, because this is information subjects lack when they have but the text to label the discourse. In this section, the prosodic structure of the monologues is studied, in order to find out whether it can (partly) account for the differences between the two conditions.

The speech was phonetically analyzed to establish potential relationships between boundaries obtained by the Rotondo method and prosodic features. Only the phonetic correlates of the boundaries in the SP condition are studied, since prosody obviously could not have played a role in the other condition. Measurements include pitch range and pause, taking F_0 maximum[1] and silent interval as the respective acoustic correlates. The distribution of different boundary tones was also investigated, as transcribed by an independent labeler (see above). The choice of these prosodic variables was inspired by the literature, since these were mentioned as being important phonetic structuring devices. Potentially interesting prosodic markers of discourse structure, such as final lengthening (Lehiste, 1979) and variation in speech tempo (Brubaker, 1972), were not investigated, because phonetically segmented and annotated versions of the monologues were not available.

1. Pause

A first major device reported in the literature to mark boundaries of different kinds is pause (for instance, Swerts

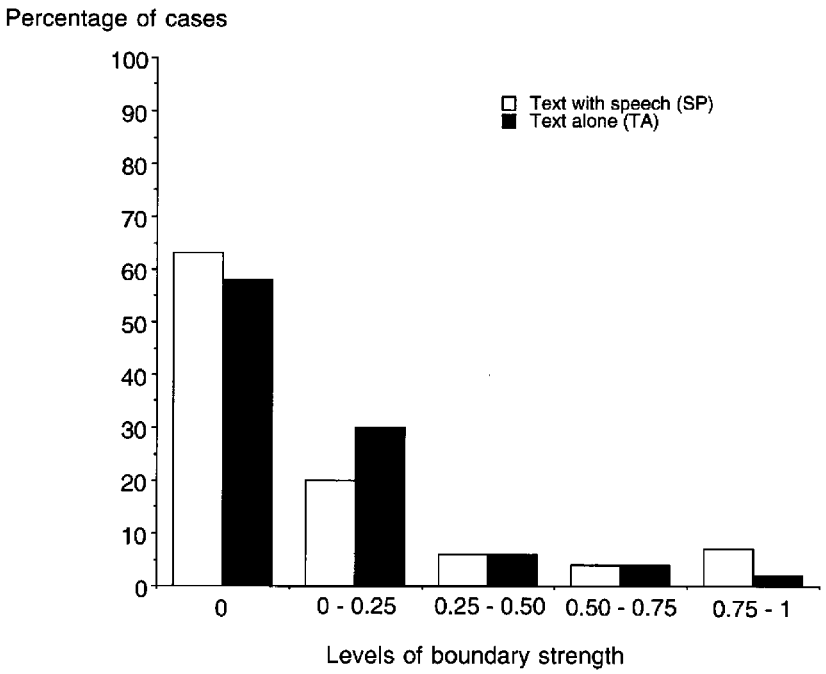


FIG. 3. Percentage of responses for different degrees of boundary strength in SP and TA condition.

and Geluykens, 1994; Grosz and Hirschberg, 1992; de Pijper and Sanderman, 1994; Passonneau and Litman, 1993). Structural breaks are generally accompanied by a significant silent interval. It can therefore be expected that the breaks given by the current method are also signalled by pause. More specifically, it is explored whether differences in the strength of a break are marked by differences in the length of the pause. Strangert (1993) found in Swedish spoken news bulletins that the difference between phrases, sentences, and paragraphs is reflected in the length of the silent interval at the boundary.

The place and length of pauses were taken from the original database (Beun, 1991), in which only those pauses were retained of at least 1 s long. Additionally, shorter silent intervals were measured manually by the author with the minimum value set to 0.25 s.

There is a moderate, but very significant correlation between the boundary strength values and the pause durations: $r=0.63$ ($n=889$, $p<0.0001$), showing that there is trend for longer pauses to be associated with stronger breaks. Table I gives the average boundary strength values for each of six pause duration categories, for both speakers pooled and separately. An analysis of variance reveals a significant overall difference between the six pause categories ($F_{(5,883)}=184.35$, $p<0.0001$). *Post-hoc* Scheffé tests showed that all the conditions differed from each other significantly ($p<0.05$).

2. Pitch reset

Previous studies, at various levels of linguistic analysis (for instance, de Pijper and Sanderman, 1994; Swerts and Geluykens, 1994), have reported that breaks between information units, can be highlighted by means of melodic discontinuities. Pitch tends to decline in the course of an utterance, but is generally reset at junctures in the information flow. Therefore, it was hypothesized that there will be a strong tendency for a declination reset at major breaks given by the Rotondo method. It was also explored whether the depth of a break is reflected in the amount of resetting. Swerts *et al.* (1996) found for Swedish spontaneous and read-aloud monologues a systematic difference in pitch reset between units differing in depth, i.e., phrases and utterances, with a weaker reset in the former than in the latter.

The pitch resets were measured in two steps. First, in any given phrase, the highest F_0 peak in an accented syllable at the vowel's amplitude maximum was taken as a measure

TABLE I. Mean boundary strength values (and respective standard deviations and number of observations) associated with different degrees of pause length (in seconds) for speakers MM and LK, pooled and separately.

	Both			MM			LK		
Pause length	$\bar{x}$	s.d.	n	$\bar{x}$	s.d.	n	$\bar{x}$	s.d.	n
No pause	0.01	0.04	447	0.01	0.04	188	0.01	0.05	259
<1 s	0.08	0.16	198	0.08	0.18	76	0.08	0.15	122
1–2 s	0.26	0.32	159	0.23	0.29	48	0.28	0.33	111
2–3 s	0.37	0.28	24	0.24	0.32	8	0.44	0.25	16
3–4 s	0.58	0.29	30	0.44	0.30	14	0.69	0.24	16
>4 s	0.78	0.24	31	0.78	0.22	20	0.77	0.30	11

TABLE II. Mean boundary strength values (and respective standard deviations and number of observations) associated with different degrees of pitch reset (in semitones) for speakers MM and LK, pooled and separately.

	Both			MM			LK		
Pitch reset	$\bar{x}$	s.d.	n	$\bar{x}$	s.d.	n	$\bar{x}$	s.d.	n
<-4.11	0.02	0.05	19	0.02	0.04	9	0.02	0.07	10
$-4.11--2.07$	0.07	0.19	95	0.06	0.18	40	0.07	0.19	55
$-2.07--0.03$	0.07	0.19	310	0.08	0.19	118	0.07	0.18	192
-0.03–2.01	0.11	0.23	277	0.09	0.23	111	0.12	0.23	166
2.01–4.05	0.37	0.38	101	0.32	0.39	39	0.40	0.36	62
>4.05	0.40	0.38	25	0.39	0.39	12	0.41	0.39	13

of pitch range. This measure was taken rather than the actual F_0 peak, since the latter sometimes constituted a somewhat exaggerated value, certainly in cases where a prominence-lending pitch rise combined with a melodic boundary marker (Grosz and Hirschberg, 1992; Beckman and Ayers, 1994; Menn and Boyce, 1982). Second, the distance (in terms of semitones) was measured between the pitch range values before and after a given boundary of a particular strength. If there was an upward reset, this measure gives a positive number, and vice versa.

The correlation between the boundary strength values and the pitch reset data is low, but significant ($r=0.35$, $n=827$, $p<0.0001$). Table II lists the average boundary strength values for different degrees of pitch reset, again for both speakers pooled and separately. To create six categories, the reset data were clustered taking steps of 1 s.d. An analysis of variance showed a significant overall effect of pitch reset on boundary strength ($F_{(5,821)}=33.13$, $p<0.0001$). *Post-hoc* Sheffé comparisons showed that the resets are clustered in two groups, the four weakest versus the two strongest degrees of reset.

3. Boundary tone

The last prosodic variable investigated was a type of boundary tone. Several studies (Brown *et al.*, 1980; Swerts and Geluykens, 1994; Swerts *et al.*, 1994; Swerts *et al.*, 1994) have shown that speakers can signal information structure by means of different boundary tones. That is, phonologically low tones tend to occur at the end of discourse segments, whereas the high ones are generally found inside such units. Accordingly, the hypothesis here is that low tones are more typical in the final position of a major discourse segment, and conversely, high ones are more typical of more shallow breaks.

Originally, the transcriber who marked phrase boundaries (see above) was instructed to also mark them as ending in a low, mid, or high boundary tone. No specific framework was given. In this paper, the latter two categories are collapsed into one, i.e., nonlow boundary tones, because the transcriber complained about the difficulty to reliably distinguish between the high and mid tones. In this way, the prosodic labeling becomes identical to the distinction made by Brown *et al.* (1980).

Table III gives the average boundary strength values for the two types of low boundary tone, again for the two speakers pooled and separately. An analysis of variance shows that the two types are significantly different ($F_{(1,887)}=87.03$, $p<0.0001$).

C. Prediction of discourse structure

The previous section has shown that there are indeed systematic correspondences between the break indices obtained by the Rotondo method and prosodic variables. It is now explored to what extent the discourse structure can really be predicted by means of these prosodic variables. It is not expected that one can fully explain the segmentation results on the basis of the pausal and melodic cues alone, since subjects to a large extent appear to score similarly in the SP and TA conditions (see r values given above). The task given to the labelers essentially was a semantic one, as they actually were asked to mark information units for which basically only the content is needed. The hypothesis is that prosody may help the job of segmenting the text as it supports particular segmentation hypotheses on the part of a listener.

In order to find out more about the significance of prosodic variables for the segmentation task, a general linear model was fit to the scores from the SP condition using a method of least squares regression. The following additive model, not including interactions,

$$\text{boundary}=\text{pause}+\text{reset}+\text{tone}, \tag{1}$$

with pause and range clustered into six levels and with two types of boundary tone (low and nonlow) can account for 58% of the variance, with significant effects for each of the prosodic parameters (see Table IV). If the two-way and

TABLE III. Mean boundary strength values (and respective standard deviations and number of observations) associated with type of preceding boundary tone for speakers MM and LK, pooled and separately.

	Both			MM			LK		
Boundary tone	$\bar{x}$	s.d.	n	$\bar{x}$	s.d.	n	$\bar{x}$	s.d.	n
nonlow	0.09	0.22	761	0.08	0.21	290	0.10	0.22	471
low	0.31	0.36	128	0.27	0.35	64	0.36	0.37	64

TABLE IV. Results of a general linear model fit on the segmentation data in the SP condition.

Pause	$F_{(5,815)}=195.82, \quad p<0.0001$
Reset	$F_{(5,815)}=20.52, \quad p<0.0001$
Boundary	$F_{(1,815)}=30.98, \quad p<0.0001$

three-way interactions are added to the model, the explained variance increases to 65%. If the model is simplified to each of the prosodic factors separately, it turns out that pause can explain 51% of the variance, pitch reset 17% and boundary tone 9%.[2]

From the above, it can be concluded that prosody may have functioned as a cue to the labelers, but that it is clearly not the only factor. Obviously, subjects have probably relied extensively on the content of the texts presented. Since this was the only information for subjects in the TA condition, their scores, clustered into five categories (see above), were added as a parameter to the additive model in (1). The reason to include these responses is to gain insight into the relative importance of the semantic cues. In doing so, the explained variance increases to 85%. Including two-way interactions results in an explained variance of 90%. A fit on the text cues separately gives a value of 77%. So one may conclude that the textual cues are clearly predominant for the labelers, but that the contribution of the prosodic variables studied here is certainly not negligible.

III. DISCUSSION

The contribution of this paper is primarily methodological in that it presents a technique to analyze hierarchical discourse structure and its potential phonetic correlates in unrestricted discourse. It is a useful alternative to existing methods, as it is general and reproducible. A minor disadvantage, however, is that the boundary strength measure (ideally) requires a large amount of subjects. Of course, as Rotondo (1984) already remarked before, this method gives a somewhat temporary solution since ultimately one would like to have an accurate theory of discourse processing to get an adequate representation of text structures. Consequently, a logical follow-up to the current work might be to try and determine in what ways the experimentally based discourse boundaries correspond to junctures predicted by discourse theories.

There are, however, a few drawbacks which need to be solved in the future, although they are probably not typical for this approach alone. First, some of the assumptions somewhat implicitly underlying this technique is that topic changes are abrupt in the sense that they occur locally between a phrase ending one information unit and another starting the next one. The technique so far does not capture more gradual and smooth transitions between discourse segments. Similarly, the method does not account for the fact that labelers may be confused about the exact spot at which a boundary occurs, but they may agree that within a certain ''region'' there is change of information unit. Both problems may, however, be solved by applying some sort of window on the data.

As for the prosodic results, it is interesting to see that prosodic variables such as pitch reset, pause length, and proportion of low boundary tones increase continuously with boundary strength at the discourse level. This is similar to prosodic phrasing results below the level of the sentence: Wightman *et al.* (1992) found that preboundary lengthening appears ''to occur in several gradations corresponding to some of the perceptually distinct break indices'' (p. 1714); de Pijper and Sanderman (1994) also report that the strength of prosodic boundaries in spoken utterances perceived by listeners is not simply a binary feature as their continuous perceptual boundary measure covaries with such features as pause, pitch and declination reset in a seemingly gradient manner. From the current study, a similar picture emerges for phrasing at the macrolevel, because prosody has been shown to mark boundaries of segments at multiple levels of embeddedness.

Of course, the features studied in this paper are not the only potentially interesting ones. In particular, preliminary observations suggest that transitions between major information units are accompanied by hesitation phenomena, such as filled pauses. This is in agreement with the claim by Chafe (1980) that significant breaks in the flow of information lead to conspicuous speech disfluencies, indicating that these are points where a speaker is planning a lot to verbalize his ideas. Similarly, Brubaker (1972) found that initial utterances of a paragraph tend to be uttered at a slower speaking rate than other utterances, being consistent with a ''reduction of uncertainty hypothesis,'' i.e., that speaking rate increases and pause duration decreases as uncertainty about the remaining content of the paragraph decreases (see also Swerts *et al.*, 1994). An interesting area for future research, therefore, is to explore whether discourse structure achieved by the Rotondo method has some psycholinguistic relevance. Given the claims by Chafe (1980), a speaker-oriented hypothesis is that hesitations will tend to occur at major breaks in the flow of information. First analyses on the monologues investigated here (Swerts *et al.*, 1996) indeed show that at least one type of disfluencies, i.e., filled pauses, may carry information about discourse structure, in that stronger breaks in the flow of information are more likely to co-occur with filled pauses than weaker ones. Moreover, filled pauses at stronger breaks tend to be segmentally and prosodically different from other ones, and more often have preceding and following silent pauses.

The Introduction remarked on the difficulty of obtaining independent specification of junctures in the information flow and of the mutual relationships between them. If such an independent specification could be achieved, the contribution of prosodic features could then be analyzed by studying correlations between the features and the independently established junctures. This article has introduced a new method to achieve such independent specification; this is based on equating boundary strength with the frequency of junctural judgments by subjects. The strong correlations found between the strength of junctures thus obtained and the strength of prosodic marking confirms the efficacy of the methodology and opens up a new path of research.

ACKNOWLEDGMENTS

M. Swerts is also affiliated with the University of Antwerp (UIA) and with the Flemish Fund for Scientific Research (F.W.O). Thanks are due to R.-J. Beun for providing the speech materials, to E. Blaauw for transcribing them in terms of prosodic phrases, and to R. Collier and A. A. Sanderman for commenting upon an earlier version of this paper.

[1]The fundamental frequency (F_0) is taken as the acoustic correlate of pitch and was determined by the method of subharmonic summation (Hermes, 1986).

[2]Note that these percentages, given our general linear model fitting procedure, do not have to add up to the value of the model with the prosodic parameters combined.

Beckman, M. E., and Ayers, G. (**1994**). ''Guidelines for ToBI labelling. Manuscript and accompanying speech materials,'' Ohio State University. [Obtain by writing to tobi@ling.ohio-state.edu]

Beun, R. J. (**1991**). Transcripties spontaan gesproken monologen [Transcriptions of spontaneously produced monologues], IPO report 792.

Brown, G., and Yule, G. (**1983**). *Discourse Analysis* (Cambridge U.P., Cambridge).

Brown, G., Currie, K., and Kenworthy, J. (**1980**). *Questions of Intonation* (Croom Helm, London).

Brubaker, R. S. (**1972**). ''Rate and pause characteristics of oral reading,'' J. Psycholinguist. Res. **1**, 141–147.

Bruce, G. (**1982**). ''Textual aspects of prosody in Swedish,'' Phonetica **39**, 274–287.

Carletta, J. (**1996**). ''Assessing agreement on classification tasks: the kappa statistic,'' Computational Linguist. **22** (2), 249–254.

Carletta, J., Isard, A., Isard, S., Kowtko, J., Newlands, A., Doherty-Sneddon, G., and Anderson, A. (**1996a**). ''HCRC Dialogue Structure Coding Manual,'' HCRC/TR-82, Human Communication Research Centre, University of Edinburgh.

Carletta, J., Isard, A., Isard, S., Kowtko, J., Newlands, A., Doherty-Sneddon, G., and Anderson, A. (**1996b**). ''Dialogue Structure Coding and Its Uses in the Map Task'' (in preparation).

Chafe, W. L. (**1980**). *Pear Stories: Cognitive, Cultural and Linguistics Aspects of Narrative Production* (Ablex, Norwood, NJ).

Condon, S. L., and Čech, C. G. (**1996**). ''Discourse Coding Reliability: Problems and Possibilities'' (in preparation).

Geluykens, R. (**in press**). *The Pragmatics of Discourse Anaphora in English: Evidence from Conversational Repair* (Mouton de Gruyter, Berlin).

Grosz, B., and Hirschberg, J. (**1992**). ''Some intonational characteristics of discourse structure,'' Proc. International Conference on Spoken Language Processing, Banff, Canada, October 1992, pp. 492–432.

Grosz, B., and Sidner, C. L. (**1986**). ''Attentions, intentions, and the structure of discourse,'' Computational Linguist. **85**, 363–394.

Hermes, D. J. (**1986**). ''Measurement of pitch by subharmonic summation,''J. Acoust. Soc. Am. **83,** 257–264.

Hirschberg, J., and Grosz, B. (**1992**). ''Intonational Features of Local and Global Discourse Structure,'' *Proc. of the Speech and Natural Language Workshop* (DARPA, Harriman, NY), pp. 441–446.

Hirschberg, J., Nakatani, C., and Grosz, B. (**1995**). ''Conveying discourse structure through intonation variation,'' Proc. ESCA Workshop on Spoken Dialogue Systems: Theories and Applications, pp. 189–192, Vigsø, Denmark, May/June 1995.

Jekat, S., Klein, A., Maier, E., Maleck, I., Mast, M., and Quantz, J. J. (**1995**). ''Dialogue Acts in VERBMOBIL,'' Universität Hamburg, DFKI Saarbrücken, Universität Erlangen, TU Berlin, Verbmobil Report 65.

van Kuppevelt, J. (**1995**). ''Discourse structure, topicality and questioning,'' J. Linguist. **31**, 109–147.

Lehiste, I. (**1975**). ''The phonetic structure of paragraphs,'' in *Structure and Process in Speech Perception*, edited by A. Cohen and S. Nooteboom (Springer-Verlag, Berlin), pp. 195–206.

Lehiste, I. (**1979**). ''Perception of sentence and paragraph boundaries,'' in *Frontiers of Speech Communication Research*, edited by B. Lindblom and S. Öhman (Academic, London), pp. 191–201.

Mandler, J. M., and Johnson, N. S. (**1977**). ''Remembrance of Things Parsed: Story Structure and Recall,'' Cognitive Psychol. **9**, 111–151.

Mann, W., and Thompson, S. (**1986**). ''Relational Propositions in Discourse,'' Discourse Processes **9**, 57–90.

Menn, L., and Boyce, S. (**1982**). ''Fundamental frequency and discourse structure,'' Language Speech **25**, 341–383.

Nakatani, C., Grosz, B., and Hirschberg, J. (**1995**). ''Discourse structure in spoken language: Studies on speech corpora,'' Proc. AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation. Stanford, March.

Nakatani, C., Grosz, B. J., Ahn, D. D., and Hirschberg, J. (**1995**). ''Instructions for Annotating Discourses,'' Technical Report Number TR-21-95. Center for Research in Computing Technology (Harvard University, Cambridge, MA).

Passonneau, R. J., and Litman, D. J. (**1993**). ''Intention-based segmentation: human reliability and correlation with linguistic cues,'' Proc. ACL-93, Ohio State University: Association for Computational Linguistics.

de Pijper, J. R., and Sanderman, A. A. (**1994**). ''On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues,'' J. Acoust. Soc. Am. **96,** 2037–2047.

Reichman, R. (**1985**). *Getting Computers to Talk like You and Me: Discourse Context, Focus, and Semantics* (Bradford, Cambridge).

Reithlinger, N. (**1995**). ''Some experiments in Speech Act Prediction,'' Proc. AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation, Stanford, March.

Rotondo, J. A. (**1984**). ''Clustering analyses of subjective partitions of text,'' Discourse Processes **7**, 69–88.

Shiffrin, D. (**1987**). *Discourse Markers* (Cambridge U.P., Cambridge).

Sluijter, A., and Terken, J. (**1994**). ''Beyond sentence prosody: Paragraph intonation in Dutch,'' Phonetica **50**, 180–188.

Strangert, E. (**1993**). ''Speaking style and pausing,'' PHONUM **2**, 121–137.

Swerts, M., and Collier, R. (**1992**). ''On the controlled elicitation of spontaneous speech,'' Speech Commun. **11**, 463–468.

Swerts, M., and Geluykens, R. (**1994**). ''Prosody as a marker of information flow in spoken discourse,'' Language Speech **37**, 21–43.

Swerts, M., Collier, R., and Terken, J. (**1994**). ''Prosodic predictors of discourse finality in spontaneous monologues,'' Speech Commun. **15**, 79–90.

Swerts, M., Bouwhuis, D. G., and Collier, R. (**1994**). ''Melodic cues to the perceived ''finality'' of utterances,'' J. Acoust. Soc. Am. **96**, 2064–2075.

Swerts, M., Strangert, E., and Heldner, M. (**1996**). ''F_0 declination in spontaneous and read-aloud speech,'' Proc. International Conference on Spoken Language Processing, Philadelphia, October 1996, pp. 1501–1504.

Swerts, M., Wichmann, A., and Beun, R.-J. (**1996**). ''Filled pauses as markers of discourse structure,'' Proc. International Conference on Spoken Language Processing, Philadelphia, October 1996, pp. 1033–1036

Terken, J. (**1984**). ''The distribution of pitch accents in instructions as a function of discourse structure,'' Language Speech **27**, 269–289.

Thorndyke, P. W. (**1977**). ''Cognitive Structures in Comprehension and Memory of Narrative Discourse,'' Cognitive Psychol. **9**, 77–110.

Thorsen, N. G. (**1985**). ''Intonation and text in Standard Danish,'' J. Acoust. Soc. Am. **77**, 1205–1216.

Thorsen, N. G. (**1986**). ''Sentence intonation in textual context—supplementary data,'' J. Acoust. Soc. Am. **80**, 1041–1047.

Venditti, J., and Swerts, M. (**1996**). ''Prosodic cues to discourse structure in Japanese,'' Proc. International Conference on Spoken Language Processing, Philadelphia, October 1996, pp. 725–728.

Webber, B. (**1988**). ''Discourse deixis: Reference to discourse segments,'' Proc. of the 26th Annual Meeting, pp. 113–122, Buffalo, Association for Computational Linguistics.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P. J. (**1992**). ''Segmental durations in the vicinity of prosodic phrase boundaries,'' J. Acoust. Soc. Am. **91**, 1707–1717.

On the perceptual distance between speech segments

Oded Ghitza and M. Mohan Sondhi
Acoustics and Audio Communication Research, Bell Laboratories, Murray Hill, New Jersey 07974

(Received 5 June 1996; accepted for publication 2 August 1996)

For many tasks in speech signal processing it is of interest to develop an objective measure that correlates well with the perceptual distance between speech segments. (Speech segments are defined as pieces of a speech signal of duration 50–150 ms. For concreteness, a segment is considered to mean a diphone, i.e., a segment from the midpoint of one phoneme to the midpoint of the adjacent phoneme.) Such a distance metric would be useful for speech coding at low bit rates. Saving bits in those systems relies on a perceptual tolerance to acoustic perturbations from the original speech—perturbations whose effects typically last for several tens of milliseconds. Such a distance metric would also be useful for automatic speech recognition on the assumption that perceptual invariance to adverse signal conditions (e.g., noise, microphone, and channel distortions, room reverberation, etc.) and to phonemic variability (due to nonuniqueness of articulatory gestures) may provide a basis for robust performance. In this paper, attempts at defining such a metric will be described. The approach in addressing this question is twofold. First psychoacoustical experiments relevant to the perception of speech are conducted to measure the relative importance of various time-frequency ''tiles'' (one at a time) *when all other time-frequency information is present*. The psychophysical data are then used to derive rules for integrating the output of a model of auditory-nerve activity over time and frequency. [S0001-4966(97)03901-5]

PACS numbers: 43.72.Ar, 43.71.An, 43.71.Cq, 43.66.Ba [JS]

INTRODUCTION

This paper is concerned with the derivation of a quantitative measure of the perceptual distance between speech segments, where by ''speech segment'' we mean a piece of a speech signal of duration 50–150 ms—in particular, a diphone, i.e., a segment from the midpoint of one phoneme to the midpoint of the adjacent phoneme. In deriving such a measure, we will present a model of how the auditory system integrates auditory nerve activity over time and frequency. A measure of the perceptual distance would be of interest in its own right. It would also have practical applications. For instance, the influence of perturbations introduced by low bit rate coders may extend, in general, over segment length intervals. The design and evaluation of such coders should therefore benefit from the derivation of a perceptual distance of the type considered here. Also, we believe that such a perceptual distance would provide a robust measure for automatic speech recognition. This belief is based on the following reasoning. Human beings perform far better than any existing automatic speech recognizer, especially when the speech signal has been degraded due to variations in the transmission path, the presence of noise, articulatory changes induced by the noise (i.e., the Lombard effect), etc. Therefore, use of a distance measure based on perceptual dissimilarity may be expected to improve automatic speech recognition.

We approach the derivation of such a measure of distance in two main steps. First of all we conduct a psychoacoustical experiment relevant to the perception of speech. In traditional psychophysical experiments speech is rarely used as a test stimulus. Typically these experiments are concerned with (a) masking of steady-state signals by other steady-state signals (e.g., masking of tones by noise, noise by a tone, etc.); or (b) measurement of the just-noticeable difference (jnd) of some steady-state property (e.g., jnd for amplitude or frequency of a tone, jnd for formant frequencies or pitch, etc.). Speech, however, is a highly nonstationary signal. For processing speech signals, this nonstationarity is dealt with by partitioning the signal into contiguous ''frames'' (i.e., short time windows of about 20- to 30-ms duration). Within each frame the signal may be regarded as stationary. However, from frame to frame there is considerable nonstationarity. It is not clear how masking properties, jnds, etc., change due to this nonstationarity. Therefore the traditional studies cited above are of limited application to problems such as speech coding at low bit rates and automatic speech recognition. Not surprisingly, almost all progress in these areas has come from application of signal processing techniques, with little help from psychophysics.

This paper is aimed at improving this situation. In contrast to experiments of the type mentioned above, our experiment involves ''segment level'' properties, i.e., properties of the whole segment rather than those of individual frames. For concreteness we will consider diphones, although longer segments could be studied by similar methods. Our experiment is aimed at quantifying the relative importance of various time-frequency regions (which we call ''**tiles**'') for the perception of a given segment. To achieve this, we study the perceptual effects of modifying a selected tile **while at the same time leaving the information in all other time-frequency regions unchanged**. This experiment, which we call the ''tiling'' experiment is described in the next section.[1]

The second important step of our approach is to **simulate** the tiling experiment. The simulation depends upon a definition of distance between observation vectors based on the ensemble interval histogram (EIH). As discussed in

(Ghitza, 1994), the EIH is a functional model of how auditory-nerve firings are analyzed in the auditory periphery. The desired distance measure is derived by driving the performance of the simulated tiling experiment to mimic that of the human subjects.

The detailed description of these procedures given in the next few sections may be summarized as follows: The psychophysical paradigm used is the diagnostic rhyme test (DRT). The word pairs in the DRT are modified by interchanging judiciously selected time-frequency regions (tiles). This modified database is used in the standard DRT, and the error patterns induced by these changes are recorded. The same DRT is then simulated by an array of speech recognizers. These recognizers use a parametrized distance between EIH vector sequences derived from the speech waveform. The parameters of the distance metric are jointly optimized over relevant tiling conditions so as to mimic the error patterns of the human subjects. The optimal set of parameters then defines the desired perceptual distance metric. (Note that such an optimization can be performed with any choice of observation vectors. We chose EIH on the assumption that the optimized parameters will have relevance to human speech perception because, as mentioned above, the EIH is a functional model of the auditory periphery (Ghitza, 1994).

I. PSYCHOPHYSICS

The experiment used in our search for the perceptual distance is what we call the ''tiling'' experiment. It has been described in a recent publication (Ghitza, 1993a), so we will give only a brief description of it here. In this experiment we measured the relative importance of various time-frequency ''tiles'' by studying the perceptual effects of modifying these tiles one at a time, or by simultaneously modifying various combinations of these tiles. It is important to note that when a particular tile (or a combination of tiles) is modified, the information in the rest of the time-frequency plane is left unaltered. In this way we measure the perceptual importance of that tile (or combination) in the presence of all other time-frequency information.

For the psychophysical paradigm we have chosen the DRT, which was first suggested by Voiers (1983), and which has been in extensive use for evaluating speech coders. In the DRT, Voiers uses 96 pairs of confusable words spoken by several male and female speakers. All the words are of the consonant-vowel-consonant (CVC) type, and the words in each pair differ only in the initial consonant. In an attempt to uniformly cover the speech subspace associated with initial diphones, the DRT database was designed such that the target diphones are equally distributed among six phonemic distinctive features (16 word pairs per feature) and among eight vowels. The feature classification follows the binary system suggested by Jakobson *et al.* (1952)[2] and the target consonants in each pair differ in the presence or absence of one of these dimensions. An explanation of these attributes, as well as the complete list of words, may be found in Ghitza (1993a).

The database is used in a very carefully controlled psychophysical procedure. The listeners are well trained and quite familiar with the database, including the voice quality

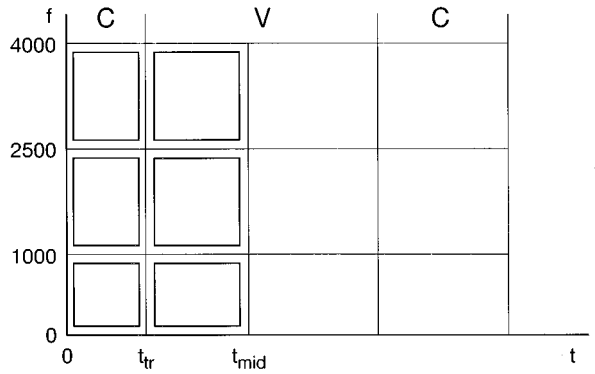


FIG. 1. Configuration of the six time-frequency tiles chosen for the tiling experiment.

of the individual speakers. As for the psychophysical procedure, a one-interval two-alternative forced-choice paradigm is used. A word pair is selected at random and displayed as text on a screen. One of the words in the pair (selected at random) is next presented aurally, and the subject is required to indicate which of the two words was heard. The procedure is repeated until all the words in the database have been presented. The errors made by the subjects are recorded.

For the tiling experiment the DRT was conducted on several distorted versions of Voiers' standard database. The details of the signal processing involved in creating those distortions may be found in Ghitza (1993a). Briefly, we divided the time-frequency plane into nonoverlapping regions called ''tiles'' that cover the target diphone in each pair of words in the DRT. Ideally, one should use many small tiles, but the experiments become increasingly time consuming and expensive with increasing number of tiles. From considerations of feasibility, we decided to use six tiles with the configuration shown in Fig. 1. The six regions were chosen on the basis of the following rough reasoning: On the time axis a break at the boundary between the C and V portions of the target diphone is an obvious choice. This boundary as well as the midpoints of the C and the V were hand labeled by a trained phonetician (see Ghitza, 1993a). On the frequency axis two breaks were selected. A break at 1 kHz is suggested by the known change in the properties of nerve firings at approximately this frequency (e.g., loss of synchrony beyond 1 kHz). A break at 2.5 kHz corresponds roughly to the upper limit of the second formant frequency (Peterson and Barney, 1952). We will call the resulting frequency regions as band-1 (0–1 kHz), band-2 (1–2.5 kHz), and band-3 (2.5–4 kHz).

Each distorted database was generated by interchanging a particular tile (or a combination of tiles) between the target diphones of each of the 96 pairs of words in the database. Such an interchange is illustrated in Fig. 2, in which the tile selected is the consonant part of the target diphone between 1 and 2.5 kHz. In a similar manner, a total of 14 distorted versions of the database were created. As described in Ghitza (1993a), special care was taken to minimize artifacts in the speech signals due to the interchange operation.

A DRT test was performed on the original database as well as on each of these distorted versions. The error for each

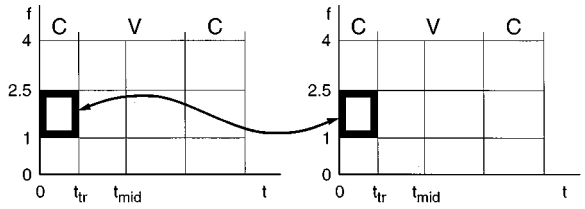


FIG. 2. Illustrating the interchange of tiles for a pair of words in the DRT database.

word pair, for each of these distortion conditions was recorded for each of three speakers and each of eight listeners. As described in Ghitza (1993a), these experiments demonstrated that perceptually, the interchange of the entire diphone in each band is far more dominant than the interchange of the consonant part or the vowel part alone. Therefore, to derive the parameters of the perceptual distance we used only the original, undistorted database and the distorted versions corresponding to the interchange of the entire diphone in band-1, band-2, and band-3, respectively.

These error patterns constitute the psychophysical data which we would like to mimic with a simulated DRT test. The simulated DRT described in the next section uses the same speech waveforms as were used in the psychophysical test. The optimization procedure used to drive the error patterns of the simulated test toward those of the human subjects is discussed in Sec. III.

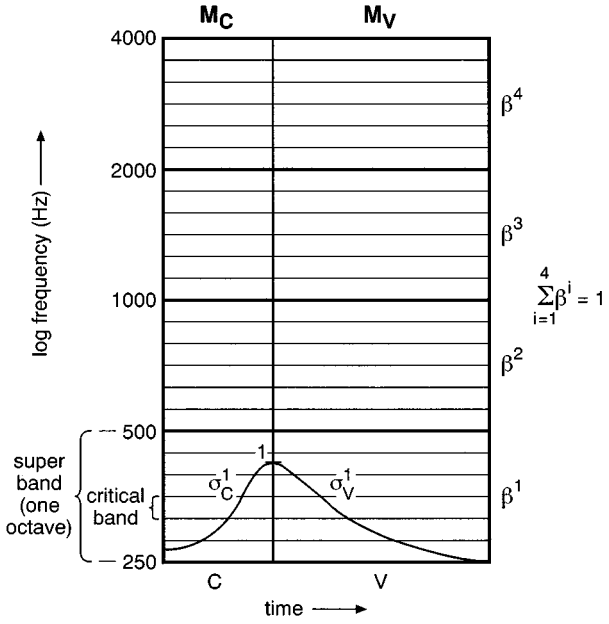


FIG. 3. A schematic diagram describing the parameters used in the perceptual distance metric. The matrices $\mathbf{M}_C$ and $\mathbf{M}_V$ capture the intuitive notion of ''lateral interaction'' between neighboring critical bands. The γ_n^i's weight the vector distance near the transition point more heavily than vector distances further away from it. The β^i's reflect the relative importance of the ith superband, in the presence of normal activity in all other superbands, and are subject to the constraint $\Sigma_{i=1}^4 \beta^i = 1$.

II. SIMULATION

The method of simulating the DRT has been described in Ghitza (1993b), and the reader is referred to that article for details. Recall that the DRT is a one-interval, two-alternative forced-choice experiment. At each stage the subject knows which pair of words has been selected and that one of them will be presented at random. The subject must make a decision and indicate which word was heard. We therefore postulate that for each of these binary decisions, the subject is able to retrieve from memory a recognizer optimized for that pair of words.

In view of this postulate, we **simulate** the DRT by replacing the human subject by an array of recognizers (one for each pair of words in the database). The particular type of speech recognizer that we use in the simulation has also been described in a recent article (Ghitza and Sondhi, 1993), so we will not describe it in detail here. Suffice it to mention that the recognizers utilize hidden Markov models with **nonstationary** states, where each state is a template of a diphone. When used in the DRT, each recognizer in the array reduces to a binary recognizer for a pair of initial diphones, since the second diphone of the CVC is identical for the two words in each pair.[3] Thus correct recognition occurs if and only if the initial diphone of the test utterance is closer to the initial diphone of the correct word model than to the initial diphone of the other word model of the pair. The word models were derived as follows: Every speaker in the DRT database provides two repetitions of each word. We assign one of these repetitions to be the ''training'' database and the other to be the test database. The set of word models is obtained from the training database by hand segmentation. If more data were available for training, more accurate word models could be derived. We believe, however, that by restricting our experiment to a speaker-dependent mode we reduce acoustic variability, so that word models derived from just one repetition are accurate enough.

The errors made in this simulation are entirely governed by the definition of distance between the test diphone and the model diphone. The parametric form chosen for this distance is therefore of crucial importance for the successful derivation of the perceptual metric. Let us discuss briefly the parametrization that we have chosen, and some of the considerations which led us to this choice. For this discussion, it will be helpful to refer to Fig. 3.

We begin by defining a diphone as a sequence of feature vectors—one for each frame of the speech signal—roughly 100 frames per second. Our choice of feature vector is a 24-dimensional EIH vector, with the histogram bins allocated on the ERB scale. (ERB stands for equivalent rectangular bandwidth, which is the bandwidth of a hypothetical rectangular filter that approximates the critical band of the human auditory filters. See detailed definition in Ghitza, 1994.) As mentioned earlier, the EIH is a functional model of how auditory-nerve firings are analyzed in the auditory periphery (Ghitza, 1994).

Let $\mathbf{x}_k$, $k=1,2,\ldots,K$ be the EIH vectors of some given test sequence $\mathbf{X}$. We need to define the distance of $\mathbf{X}$ from a template (or state) sequence $\mathbf{S}$ with EIH vectors $\mathbf{s}_n$, $n=1,2,\ldots,N$. The length N of the template sequence is in general, different for different templates. The length K of the test sequence is arbitrary within some broad range of values.

Note that each component of the EIH vector functionally represents the activity in a local region of the frequency spectrum (roughly a critical band). We postulate that these components are processed in groups—or **superbands**—to arrive at a distance within each group, and then the individual distances are combined to give the overall distance between the EIH vector sequences **X** and **S**. In the auditory periphery, presumably, there is a continuum of overlapping superbands covering the frequency spectrum. In our functional model we replace this continuum by four nonoverlapping superbands—roughly one octave wide. Each superband consists of six ERB bins of the whole EIH vector.

In order to define a distance between **X** and **S** we first time align the two sequences by warping **X** onto **S**. For this we use the usual dynamic time warp (DTW) algorithm,

$$\Delta^2(\mathbf{S},\mathbf{X}) = \frac{1}{N} \min_{k(n)} \sum_{n=1}^{N} d^2(\mathbf{s}_n, \mathbf{x}_{k(n)}), \tag{1}$$

where $k(n)$ is the time warp [with $k(1)=1$ and $k(N)=K$], and d is the Euclidean distance between $\mathbf{s}_n$ and $\mathbf{x}_{k(n)}$. Let $\tilde{\mathbf{x}}_n$, $n=1,2,\ldots,N$, be the EIH vectors of **X** after this alignment. We next define the distance between individual vectors $\tilde{\mathbf{x}}_n$ and $\mathbf{s}_n$. Let $\mathbf{s}_n^i$ and $\tilde{\mathbf{x}}_n^i$ be the ith subvectors of $\mathbf{s}_n$ and $\tilde{\mathbf{x}}_n$, respectively, representing the ith superband.[4] Let $\mathbf{m}_n^i$ be a 6×6 matrix defined for the ith superband, for the time index n of the template sequence **S**. Define $\hat{\mathbf{s}}_n^i$ and $\hat{\mathbf{x}}_n^i$ as the unit length vectors

$$\hat{\mathbf{x}}_n^i = \frac{\mathbf{m}_n^i \tilde{\mathbf{x}}_n^i}{\|\mathbf{m}_n^i \tilde{\mathbf{x}}_n^i\|}, \tag{2a}$$

$$\hat{\mathbf{s}}_n^i = \frac{\mathbf{m}_n^i \mathbf{s}_n^i}{\|\mathbf{m}_n^i \mathbf{s}_n^i\|}, \tag{2b}$$

where $\|\cdot\|$ denotes the Euclidean norm, or length, of a vector. We next define the distance between the subvectors $\tilde{\mathbf{x}}_n^i$ and $\mathbf{s}_n^i$ by the relation

$$d(\mathbf{s}_n^i, \tilde{\mathbf{x}}_n^i) = \|\hat{\mathbf{s}}_n^i - \hat{\mathbf{x}}_n^i\|. \tag{3}$$

With this definition of distance between vectors, the distance between the sequences within a superband, $\mathbf{X}^i$ and $\mathbf{S}^i$ is defined as

$$D^2(\mathbf{S}^i,\mathbf{X}^i) = \frac{1}{N} \sum_{n=1}^{N} \gamma_n^i d^2(\mathbf{s}_n^i, \tilde{\mathbf{x}}_n^i). \tag{4}$$

Here, $d^2(\mathbf{s}_n^i, \tilde{\mathbf{x}}_n^i)$ is weighted by the factor γ_n^i depending upon position along the template sequence.

In general, the distance D between the EIH sequences **S** and **X** can be any function of the $D^2(\mathbf{S}^i,\mathbf{X}^i)$'s. For the present we assume that they are linearly combined. Thus

$$D^2(\mathbf{S},\mathbf{X}) = \sum_{i=1}^{4} \beta^i D^2(\mathbf{S}^i,\mathbf{X}^i), \tag{5}$$

where $D^2(\mathbf{S}^i,\mathbf{X}^i)$ is as defined in Eq. (4), and the β^i are subject to the constraint $\Sigma_{i=1}^4 \beta^i = 1$. The β^i's reflect the relative importance of the ith superband, in the presence of normal activity in all other superbands. (See Fig. 3.)

With vector distance defined as in Eqs. (2) and (3), the matrices $\mathbf{m}_n^i$ may be regarded as the submatrices of a 4×4 block diagonal matrix $\mathbf{M}_n$, in which each block has dimension 6×6. The entries in the matrices $\mathbf{M}_n$, $n=1,2,\ldots,N$, are part of the set of parameters to be determined by optimization as discussed in the next section. We have allowed the matrices $\mathbf{M}_n$ here to depend arbitrarily on the time index n of the template. With this generality, however, the number of parameters to be optimized becomes too large. We therefore restrict $\mathbf{M}_n$ to two possibilities: $\mathbf{M}_n = \mathbf{M}_C$ if the index n is in the consonant portion of the template, and $\mathbf{M}_n = \mathbf{M}_V$ if it is in the vowel portion. If the matrices $\mathbf{m}_n^i$ are chosen to be diagonal, then they serve to specify the relative importance of different components of the EIH vector (or, essentially, different critical bands). With a more general structure, they can capture the intuitive notion of ''lateral interaction.'' That is, the notion that the output of a channel might be influenced by the activity in neighboring channels. In our study we chose the $\mathbf{m}_n^i$ matrices to be tridiagonal.

The γ_n^i are introduced because we believe that the vectors near the transition point are more important for recognizing the diphone than vectors further away from it. With this in mind, we specify the function γ_n^i with just two parameters, $\sigma^i \equiv (\sigma_C^i, \sigma_V^i)$. These are the variances of two Gaussian curves with peaks at the transition point—one for the consonant part and one for the vowel part. (See Fig. 3.)

In principle, the set of parameters that define the distance D in Eq. (5) (e.g., the matrices $\mathbf{M}_C$ and $\mathbf{M}_V$ and the parameters σ^i and β^i) should be allowed to be different for different diphones. This is again not feasible because of the number of parameters involved. Note that the total number of diphones is on the order of 2000 in English. In the DRT database alone, the number of diphones is 192. Unique matrices for each diphone would require an enormous number of parameters. We therefore restrict the number of parameters by using the same sets for diphones with ''similar'' properties. At present we group together consonants into seven categories according to **manner** of articulation (voiced and unvoiced stop, voiced and unvoiced fricative, nasal, glide, and affricate). The vowels are grouped into four categories according to the location of the constriction (low back, high back, low front, and high front). This gives us 28 classes of diphones, and we assign a parameter set to each such class.

In summary, the distance of a test segment **X** from a diphone template **S** is derived as follows: Depending on the template **S**, choose the appropriate parameter set $\mathbf{M}_C$, $\mathbf{M}_V$, σ^i, and β^i. Then compute the distance according to Eqs. (1)–(5). For a given specification of all the parameters, the definition of D gives us a parametrized distance which depends on the template (or state). The entire set of parameters is optimized to best mimic human performance as described in the next section.

Finally, let us note that the parametrization described above is not necessarily optimal. Indeed we believe it can be improved in several ways. Allowing a greater range of choices for the matrices $\mathbf{M}_C$ and $\mathbf{M}_V$ as well as allowing the submatrices $\mathbf{m}_n^i$ to be full matrices (rather than tridiagonal)

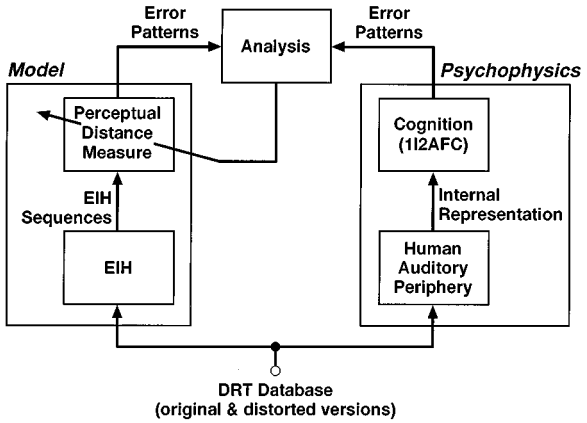


FIG. 4. A schematic diagram describing the optimization procedure. The parameters of the perceptual distance measure are iteratively adjusted to match the error patterns produced by the machine to those of the human subjects, jointly over several tiling conditions. In the box marked ''cognition,'' the abbreviation 1I2AFC stands for the ''one-interval two-alternative forced-choice'' paradigm.

are obvious possible improvements in the definition of the vector distance d. A more promising improvement is a generalization of the form of the segment distance D in the following manner: Let $\mathbf{V}^s$ be the $(24N)$-dimensional vector obtained by concatenating the N vectors $\mathbf{s}_n$ in the template sequence. Similarly let $\tilde{\mathbf{V}}^x_\kappa$ be the test sequence after warping on to the template sequence with a mapping $\kappa \equiv k(n)$. Then we can define the distance between $\mathbf{S}$ and $\mathbf{X}$ to be given by

$$D^2(\mathbf{S},\mathbf{X}) = \min_\kappa \mathbf{V}^{s\prime}\mathbf{\Phi}\tilde{\mathbf{V}}^x_\kappa, \tag{6}$$

where $\mathbf{\Phi}$ is a positive definite matrix and $'$ denotes matrix transpose. Here $\mathbf{\Phi}$ can be regarded as a block matrix with N^2 blocks, each block being a 24×24 matrix. Then the D of Eq. (5) is a special case of the D of Eq. (6) in which $\mathbf{\Phi}$ is a block diagonal matrix. With a full matrix $\mathbf{\Phi}$, we again have the problem of a large number of parameters to be estimated. As a first step, therefore, we might just generalize to a block tridiagonal matrix.

III. OPTIMIZATION

A schematic description of the optimization procedure is shown in Fig. 4. The right-hand side depicts the psychophysical data collected in the tiling experiment. The left-hand side shows the outputs of the simulated tiling experiment. The parameters of the simulation are iteratively adjusted to best mimic the psychophysical data.

Let $\boldsymbol{\theta}$ denote the parameter set which goes into the definition of D in Eq. (5), i.e., the $\mathbf{M}_C$'s, $\mathbf{M}_V$'s, σ^j's and β^j's. These are the adjustable parameters. In addition we have the template sequences $\mathbf{S}_j$, $j=1,192$—one for each initial diphone in the (undistorted) training database. These are kept fixed throughout the optimization procedure. For a given set of values for the parameters $\boldsymbol{\theta}$ (and the fixed templates) we define a cost function C which quantifies how badly the simulation performs when compared to the psychophysical data. Once C is defined, we use an optimization program to iteratively adjust the parameters $\boldsymbol{\theta}$ in order to minimize C. The program we use is a variant of Newton's method (Gay, 1983).

To complete the description of our optimization procedure, let us now indicate the definition of the cost function C. The data we are attempting to mimic are the responses of each of eight listeners to presentations of each of the words in the database spoken by each of three speakers, and distorted by each of K tiling conditions. (As mentioned in Sec. II, for optimization we chose only the distortions corresponding to the interchange of the entire diphone in band-1, band-2, and band-3, respectively. This gives $K=3$.) To give a concise definition for C it is convenient to define a set of indices.

Let:

p denote the word *p*air in the database $(1\leqslant p\leqslant 96)$,

w denote the *w*ord within a pair $(1\leqslant w\leqslant 2)$,

k denote the tiling condition $(0\leqslant k\leqslant K)$, with 0 for the undistorted database,

s denote the *s*peaker $(1\leqslant s\leqslant 3)$,

l denote the *l*istener $(1\leqslant l\leqslant 8)$.

Let x_{skpw} denote the speech signal corresponding to a chosen set of indices. Then the human response is a binary number, $h_l(x_{skpw})$, for each selection of the indices l,s,k,p,w in the range given above for each index. That is, h is 0 if listener l identified x_{skpw} correctly, and 1 otherwise. For a given set of parameters, $\boldsymbol{\theta}$, let $m(x_{skpw},\boldsymbol{\theta})$ denote the machine's response.

With these definitions, the cost function that we minimize is defined as[5]

$$C=\sum_{s=1}^{3}\sum_{k=0}^{K}\sum_{p=1}^{96}\sum_{w=1}^{2}\sum_{l=1}^{8}[h_l(x_{skpw})-m(x_{skpw},\boldsymbol{\theta})]^2. \tag{7}$$

The optimal solution $\boldsymbol{\theta}^*$ represents the parameters of the perceptual distance which provide the best mimic, jointly over all K tiling conditions. The accuracy of $\boldsymbol{\theta}^*$ depends upon the amount of data used for the optimization. In our database the number of tokens per **manner** class ranged between 36 and 165; the number of tokens per vowel category was 144.

IV. RESULTS

In terms of evaluating the validity of our approach, two questions come to mind. First, how closely can the machine error patterns be made to match human error patterns? And second, how does the performance of the ''optimal'' metric—derived by optimizing on ''tiling'' type of distortions—generalize to other kinds of distortions?

Figures 5(a)–(d) and 6(a)–(d) present results in an attempt to answer the first question. In all these figures we present error patterns by plotting the error rates for each of the Jakobson–Fant-Halle dimensions.[6] For each of these dimensions we plot two error rates.[7] At the abscissa marked ''+'' we plot the error rates for the subset of words in which the attribute is present, and at the abscissa marked ''−'' we plot the error rates for the words in which the attribute is absent. In the top panel of every figure the error rates for human subjects (solid line) are compared to those for the

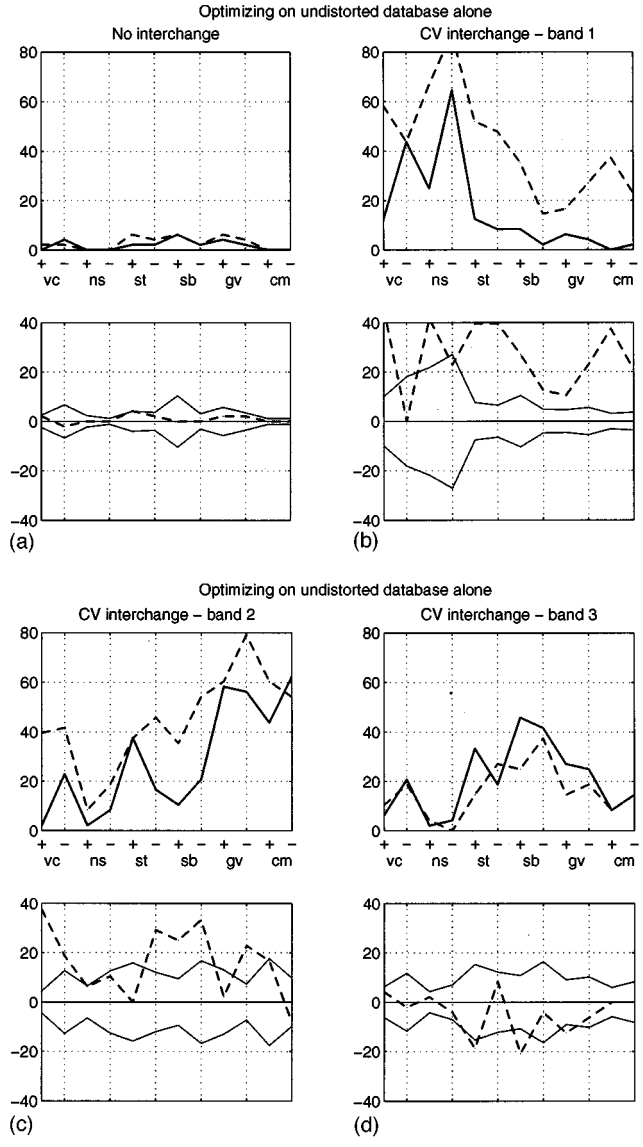


FIG. 5. Optimizing on undistorted database alone. *Top panels:* Mean human (solid line) and machine (dashed line) performance on the DRT database. The mean human performance is derived across three speakers and eight subjects. The abscissa of every plot indicates the six phonemic categories: ''**vc**'' is for voicing, ''**ns**'' for nasality, ''**st**'' for sustention, ''**sb**'' for sibilation, ''**gv**'' for graveness and ''**cm**'' for compactness''. The ''+'' sign stands for attribute present and the ''−'' sign for attribute absent. The ordinate represents the number of words in the category that, when played to the listener, were judged to be the opposite word in the word pair (i.e., the listener ''switched'' to the opposite category). The switch is represented as a percentage (relative to 16 which is the total number of words per phonemic category). *Bottom panels:* Difference between mean human performance and the machine performance (dashed line) compared to the human standard deviation (solid lines). The plots are for the original database (a) and for the three tiled versions obtained by interchanging bands 1, 2, and 3 of the entire diphone [(b), (c), and (d), respectively].

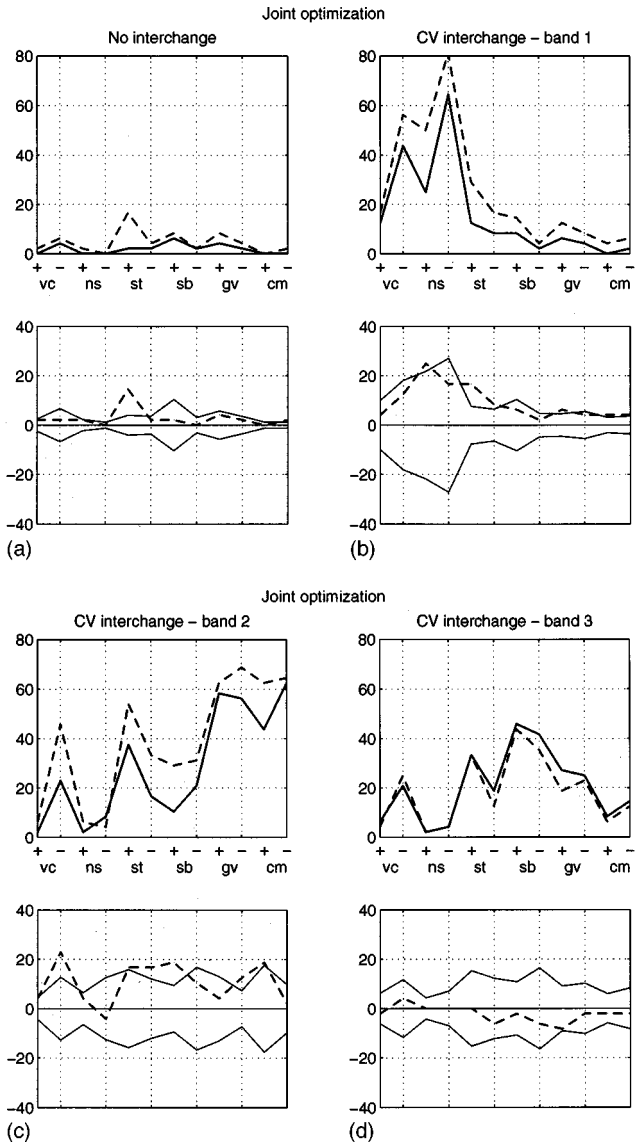


FIG. 6. Same format as in Fig. 5, but for joint optimization.

machine (dashed line). In the bottom panel the dashed line shows the error rate for the machine minus the error rate for the human subjects. Also plotted for comparison are two solid lines representing ± one standard deviation of the error rate for human subjects.

We first tested if the model structure is flexible enough for the purpose of mimicking the human performance for one tiling condition alone—say, the undistorted, original, DRT database. For this test we optimized the parameters of the distance metric, $\boldsymbol{\theta}$, on just the tiling condition 0. As it is seen in Fig. 5(a), the model can mimic the performance of the human subjects quite well—the difference between machine and human performance (bottom panel) is within one standard deviation for all the dimensions. However, this model (whose parameters were derived by optimizing on the undistorted database alone) fails to mimic human performance for other tiling conditions. Figure 5(b) and (c) shows that the machine makes significantly more errors than human subjects for the tiling conditions in which band-1 or band-2 is interchanged for the entire diphone. To arrive at a single model that is able to mimic human performance under different tile interchanges, $\boldsymbol{\theta}$ should be **jointly optimized** over several tiling conditions.

In Fig. 6(a)–(d) we show the same comparisons as in Fig. 5(a)–(d), except that now the parameters are optimized jointly over four tiling conditions: undistorted database and the tilings in which bands 1, 2, 3, respectively, are inter-

TABLE I. Experiment with the DRT database degraded by additive Gaussian white noise. The entries are the errors, summed over all the Jakobson–Fant–Halle dimensions, in percent.

	Clean	30 dB	20 dB	10 dB
Human	3	2	3	7
EIH (perceptual)	5	8	13	24
EIH (L_2)	18	17	21	27
MEL-CEP (L_2)	11	16	25	38

changed over the entire target diphone. It is seen from Fig. 6(a) that machine performance on the undistorted database is slightly worse than that in Fig. 5(a), and the machine performance for sustention is more than a standard deviation away from human performance. However, in exchange for this small deterioration, the performance for the other tiling conditions is now much closer to human.

As for the question of how the optimal metric generalizes, we ran a simulated DRT experiment on the DRT database degraded by additive Gaussian white noise. We used three different definitions of $d(\cdot,\cdot)$ (a) the observation vectors were 13th-order Mel-Cepstrum (MEL-CEP) and d was the L_2 (i.e., Euclidean) distance; (b) the observation vectors were EIH and d was the L_2 distance; and (c) the observation vectors were EIH and the distance metric was the optimized perceptual metric derived above. Table I shows the results for those three DRT simulations and for the human subjects, as a function of SNR. The entries are the errors, summed over all the Jakobson–Fant–Halle dimensions, in percent. From Table I we conclude that although the machine performance using EIH with perceptual metric does not match human performance, it is superior to the performance using EIH with L_2 metric (and also to the performance with L_2 norm between MEL-CEP vectors). Figure 7 shows the detailed distribution of these errors along the Jakobson–Fant–Halle dimensions. The detailed distributions are shown for the human subjects, the EIH with the perceptual distance and the MEL-CEP with the L_2 distance. The results for the EIH with L_2 distance were omitted in order not to clutter the figure. The figure demonstrates that the pattern of error distribution for the perceptual distance generally follows the pattern for human subjects.

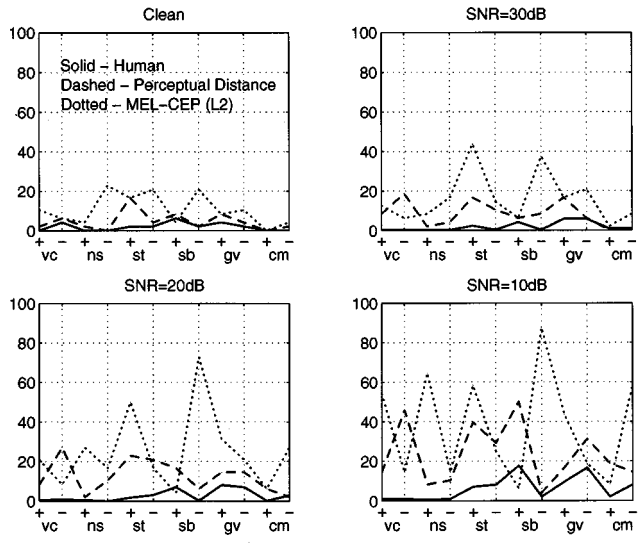


FIG. 7. Mean human performance (solid line), machine performance with the perceptual distance (dashed line) and machine performance with 13th order Mel-Cepstrum and L_2 distance (dotted line) for the DRT database in the presence of additive Gaussian white noise. The axes are as described in Fig. 5, for various SNR. (a) Clean speech, (b) SNR=30 dB, (c) SNR=20 dB, (d) SNR=10 dB.

V. DISCUSSION

In the preceding sections we have presented a method for deriving a perception-based measure of distance between speech segments. The segments we chose to investigate are diphones, although longer segments could be studied in a similar manner.

In our model, the template of a diphone is represented as a high-resolution sequence of EIH vectors (one vector every 10 ms). This template represents the articulatory gesture while moving from C to V, in terms of the time course of the EIH vectors. It may be thought of as the pattern of the diphone that is stored in memory during the early stages of language acquisition. Note that the template implicitly contains information about the **place** of articulation of the consonant. An unknown ''input'' diphone is compared to a template by first time warping it to the template and then computing a distance between the aligned sequences. This distance is expressed in terms of a set of parameters $\boldsymbol{\theta}$ which are allowed to depend upon the template. These parameters quantify the perceptual deviation from the diphone template. In order to keep the number of parameters manageable, we group the consonants into seven groups and the vowels into four groups, and assign the same parameters to the consonants and vowels within the same group. For consonants, we postulate that the parameters depend upon the **manner** of production (voiced and unvoiced stop, voiced and unvoiced fricative, nasal, glide, and affricate). The **vowels** are grouped according to the location of the constriction (front high, front low, back high, back low). In this way all C–V diphones are grouped into 28 different classes. Note that this grouping is only for the parameters $\boldsymbol{\theta}$ (that weight different time-frequency regions according to their relative perceptual importance). The templates themselves are not grouped.[8] Note also that two diphones, say /**ba**/ and /**da**/, whose consonants belong to the same **manner** class are assigned the same $\boldsymbol{\theta}$. The information about their different places of articulation is implicitly contained in the templates of the two diphones.

In deciding upon a structure for the distance we postulate that the auditory periphery processes the input in parallel frequency ''superbands'' (about an octave wide) and produces a distance in each such band. In our functional model we take four contiguous superbands, although in the auditory periphery there is presumably a continuum of overlapping bands. The distances from all superbands can be combined in many ways, providing the overall distance between the diphones. Here, we combine them linearly.

Throughout this study, we used the Jakobson–Fant–Halle feature space. These dimensions were used by Voiers to structure the DRT database, and we present our results along the same dimensions. It is worth noting, however, that

our aim is to derive a distance metric that is completely independent of the distinctive feature set. For this reason, $\boldsymbol{\theta}^*$ was computed by optimizing the cost function defined in Eq. (7). This cost function accumulates the contributions of the individual words in the database without regard to their distinctive features. However, the exact pairwise comparisons made will have some influence on the values of the optimal parameters.

The most important, and we believe novel, aspect of our work is the fact that we derive the distance measure on the basis of **perceptual** dissimilarity. We do that by mimicking human performance in the DRT framework, using tiling type of distortions. In this restricted task, at least, the metric performs significantly better than others that we have tried.

As a final note, we speculate that the perceptual distance derived here may be used to define a jnd for diphones (or phonemes). This jnd may be defined as a change for which the perceptual distance attains a threshold value.

ACKNOWLEDGMENTS

We would like to express our thanks to Ben Gold and the anonymous reviewers for many helpful suggestions to improve the manuscript.

[1]As mentioned above, psychophysical experiments dealing with speech perception are rare. Some experiments reported in the literature that have some relevance to the present paper are those of Fletcher (1953), Miller and Nicely (1955), Houtgast and Steeneken (1985), and Drullman *et al.* (1994). In Fletcher's experiments (Fletcher, 1953) subjects had to respond to stimuli that contained only parts of the speech signal (e.g., low-pass or high-pass filtered speech). Miller and Nicely (1955) studied the effect of filtering and additive noise on the confusion matrices for various phonemes. The experiments of Houtgast and Steeneken (1985) and Drullman *et al.* (1994) are concerned with the effects of filtering the speech envelope in contiguous frequency bands. Our experiments differ from all these in that we study the effects of modifying selected time-frequency regions of a speech signal while leaving the rest of the signal unchanged.

[2]The six Jakobson–Fant–Halle dimensions are *voicing*, *nasality*, *sustention*, *sibilation*, *graveness*, and *compactness*. The *voicing* (**vc**) feature characterizes the nature of the source, being periodic or nonperiodic. The *nasality* (**ns**) feature indicates the existence of a parallel resonator representing the nasal cavity. The terms *sustention* (**st**) and *sibilation* (**sb**) are due to Voiers. They correspond, respectively, to the continuant-interrupted and strident-mellow contrasts of Jakobson *et al.* (1952). Finally *graveness* (**gv**) and *compactness* (**cm**) represent broad resonance features of the speech sound, related to place of articulation.

[3]We are assuming that the effects of coarticulation due to the initial consonant do not extend beyond the midpoint of the vowel. This appears to be an accurate assumption, at least for the DRT database.

[4]Throughout the paper we used subscripts to indicate time index of a template, and superscripts to indicate superbands.

[5]If m in Eq. (7) is chosen to be a binary number, like the human responses, then C would be a discontinuous function of m which could be difficult to optimize. We therefore make $m(x_{skpw}, \boldsymbol{\theta})$ a real number between 0 and 1, whose value depends upon the distances of x_{skpw} from the two templates for the pair of words p. If d_{cor} and d_{inc} are the distances from the correct and incorrect templates, respectively, then we choose $m(x_{skpw}, \boldsymbol{\theta}) = [1+\arctan \alpha(d_{\text{inc}} - d_{\text{cor}})]/2$. The exact value of α is not very critical. The important property is that if $d_{\text{cor}} \gg d_{\text{inc}}$ then m goes to 0 and if $d_{\text{cor}} \ll d_{\text{inc}}$ then it goes to 1.

[6]The errors could be presented in other ways, e.g., along place-manner dimensions, or in the form of a confusion matrix. Note, however, that a confusion matrix format is inappropriate here because the psychophysical paradigm is a two-alternative forced-choice, and also because many binary comparisons are missing in the database. As to the choice of distinctive features, we chose the Jakobson–Fant–Halle dimensions because (a) Voiers' DRT database is organized along those dimensions and, (b) because those dimensions reflect acoustic properties in time and frequency (Jakobson *et al.* 1952).

[7]Note that here, we use the notion of ''error rate'' in the context of DRT, i.e., a binary decision paradigm: an occurrence of an error means that the listener ''switched'' to the opposite category.

[8]It may be argued that the articulatory gestures are quite similar for C–V diphones in which the vowel is the same and the place of articulation of the consonant is the same—e.g., /**ma**/ and /**ba**/. However, the corresponding spectra and EIH vectors are still quite distinct. Hence grouping of the templates themselves is not justified.

Drullman, R., Festen, F. M., and Plomp, P. (**1994**), ''Effect of temporal envelope smearing on speech reception,'' J. Acoust. Soc. Am. **95**, 1053–1064.

Fletcher, H. (**1953**). *Speech and Hearing in Communication* (Krieger, Huntington, NY).

Gay, D. M. (**1983**). ''Algorithm 611—Subroutines for unconstrained minimization using a model/trust-region approach,'' ACM Trans. Math. Software **9**, 503–524.

Ghitza, O. (**1994**). ''Auditory models and human performance in tasks related to speech recognition and speech coding,'' IEEE Trans. Speech Audio **2**(1), 115–132.

Ghitza, O. (**1993a**). ''Processing of spoken CVCs in the auditory periphery. I. Psychophysics,'' J. Acoust. Soc. Am. **94**, 2507–2516.

Ghitza, O. (**1993b**). ''Adequacy of auditory models to predict internal human representation of speech sounds.'' J. Acoust. Soc. Am. **93**, 2160–2171.

Ghitza, O., and Sondhi, M. M. (**1993**). ''Hidden Markov Models with Templates as Nonsationary States: An Application to Speech Recognition,'' Comput. Speech Lang. **7**(2), 101–119.

Houtgast, T., and Steeneken, H. J. M. (**1985**). ''A review of the MTF concept in room acoustics and its use for estimating speech intelligibility,'' J. Acoust. Soc. Am. **77**, 1069–1077.

Jakobson, R., Fant, C. G. M., and Halle, M. (**1952**). ''Preliminaries to speech analysis: the distinctive features and their correlates,'' Technical Report No. 13, Acoustic Laboratory, Massachusetts Institute of Technology, Cambridge, MA.

Miller, G. A., and Nicely, P. E. (**1955**). ''An analysis of perceptual confusions among some English consonants,'' J. Acoust. Soc. Am. **27**, 338–352.

Peterson, G. E., and Barney, H. L. (**1952**). ''Control methods used in a study of the vowels,'' J. Acoust. Soc. Am. **24**, 175–184.

Voiers, W. D. (**1983**). ''Evaluating processed speech using the Diagnostic Rhyme Test,'' Speech Technol. **1**(4), 30–39.

Accurate frequency tracking of timpani spectral lines

Donald L. Sullivan
349 Nashua Road, Sharon, New Hampshire 03458

(Received 17 April 1996; revised 1 August 1996; accepted 14 August 1996)

A procedure is described for determining the frequency of timpani sound spectral lines to accuracies on the order of 0.05 Hz. A peak tracking algorithm uses a fourth-degree polynomial fitted to a Hamming window Fourier transform as the model function for a least-squares fit to observed spectral peaks. Accuracies of the tracking process are estimated for single sinusoids with added Gaussian noise and for pairs of closely spaced sinusoids. Tracker results are presented for sounds from a good quality 26-in. drum with emphasis on the harmonicity, or lack thereof, of the various partials. © *1997 Acoustical Society of America.* [S0001-4966(97)03312-2]

PACS numbers: 43.75.Hi, 43.60.Gk [WJS]

INTRODUCTION

Timpani sound spectra show a fairly close approximation of their musically relevant partial frequencies to a harmonic series with missing fundamental. Figure 1 shows the evolution in time of the sound spectrum from a single mezzopiano stroke on a 26-in. Ludwig Professional Series drum tuned to C3 (130.8 Hz). Rossing (1982) shows that the musically relevant vibration modes of the membrane have 1,2,3,... diametral nodal lines and no nodal circles. These modes, with frequency ratios very close to 1,1.5,2..., are labeled (11),(21),(31)... in Fig. 1 and correspond to the principal tone, fifth, octave, etc. The modal frequencies differ from those expected for a circular membrane primarily because of air loading on the membrane (Rossing, 1982; Christian *et al.*, 1984; Fletcher and Rossing, 1990, Chap. 18) which coerces the frequencies into a near-harmonic relationship. Membrane stiffness and interaction with air vibration modes inside the kettle are suggested by these authors as responsible for additional small perturbations in modal frequencies.

The author, an amateur timpanist, is interested in knowing why some timpani sound better than others. This question is no more susceptible to a final answer for timpani than for violins but, as with violin researches, a study of those features which can be quantified may lead to simplified fine tuning procedures and improved drum designs. Professional timpanists emphasize the importance of a drum's design and construction details as a prerequisite for obtaining high quality sound. Of particular concern are the roundness and regularity of the bowl rim and the roundness and flatness of the counterhoop. The lore pertaining to head materials, construction, and longevity is nearly as extensive as that for violin bows. However, even the best drum will not sound good if the differential tension adjustments around the counterhoop are not done with skill. Last, but by no means least, is the manner in which the drum is struck, including strike position and strength, contact duration, and mallet construction.

This paper describes measurement and data analysis methods for examining in considerable detail the time evolution of timpani sound spectra with emphasis on the harmonicity of the musically relevant partials, i.e., those which are nearly harmonic. The methods are then applied to sounds from a good orchestral drum and selected results are presented.

I. DATA ACQUISITION AND PRELIMINARY PROCESSING

Microphone signals are preamplified, passed through an antialiasing filter, and digitized with 12-bit resolution at a sample rate of 5120.3 samples/s. The odd sample rate comes about from the limited choice of integers available as divisors for the 10-MHz crystal clock on the data acquisition board. A divisor of 1953 gives a sample rate of 5120.3 Hz and the smallest difference (0.006%≈0.01 cents) from the desired rate of 5120 Hz. Since the human ear's ability to detect pitch changes is on the order of a few cents at timpani partial frequencies (Rossing, 1990, Fig. 7.2), the sampling rate error is negligible. Data records are stored in 16 bit integer format files on the 386 20-MHz PC used for data acquisition. Acquisition software is written in C. Data processing is typically done on a 90-MHz Pentium PC using software written in C and also using Axum, a graphics package with a C-like programming language.

For timpani data collection the drum is mounted on a turntable. The unidirectional microphone is typically placed 3 in. above the drum counterhoop (the peripheral metal ring through which tension is applied to the head), and pointing toward the center of the head. Other microphone positions have been used, particularly a position 22 in. above the counterhoop, about where the player's head would be. The chosen position gives better signal to noise ratios and still deals with the directional characteristics of the radiated sound (Fletcher and Rossing, 1990, p. 510).

A typical data set consists of one 30 000-sample record from a single mezzopiano stroke at each of 6/8 azimuths, mostly chosen at the tuning screws. Radial strike position is a few inches in from the bowl edge, the usual spot chosen to produce a good sound. Mallets of medium hardness were used for the data presented here. To date, no effects from room acoustics have been observed. Such effects would, in any case, affect only the relative amplitudes of spectral components and not their frequencies.

Power spectral density (PSD) estimates are made for successive frames of 1024 data samples, each frame offset

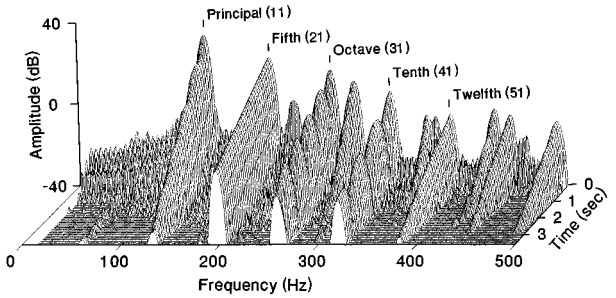


FIG. 1. Power spectral density plot of sound from a single stroke on a 26-in. Ludwig Professional Series kettledrum tuned to C3 (130.8 Hz).

512 samples (≈0.1 s) from its predecessor. This time resolution is adequate to follow the evolution of timpani partials. A Hamming window is applied to the frame which is then padded with zeros to a length of 4096 samples. The padded frame is passed through a standard fast Fourier transform (FFT) routine and the logarithm of the squared magnitude of the result is stored as the PSD estimate. The 5120.3/s sampling rate and the 4096 sample padded frame size yield samples of the PSD every 1.25 Hz (within 0.006%), although the 1024 sample length of the actual data frame limits the ability of the process to resolve spectral components closer together than 15–20 Hz, the width of the Hamming window FFT main peak.

II. PEAK TRACKING

Peak tracking schemes with frequency resolution finer than that inherent in the FFT process are summarized in Brown (1996) and their accuracies are discussed in Brown and Puckette (1993). The two broad categories of frequency domain tracking methods used to date have been quadratic fit and various phase methods. The tracking scheme described in this paper is an extension of the quadratic fit method to use a fourth degree polynomial fit to observed spectral peaks. This polynomial fit method has time resolution and frequency accuracy adequate for tracking timpani partials. Phases calculated from a single FFT frame (Brown, 1996) can be used if improved time resolution is required.

A. Peak tracking algorithm

Once the approximate location of a spectral peak has been established, an accurate estimate of its frequency and amplitude is desired. The polynomial fit method, like its predecessors, corrects for the location of the peak at a frequency other than one of the quantized FFT frequencies. The implementation described here takes advantage of the fact that the vibration modes of interest are very lightly damped ($Q > 100$). In this case it is not unreasonable to assume that the Fourier transform of the peak signal is a delta function. The observed spectral peak will therefore be the transform of the windowing function, at least in the case where the peak is isolated. Complications arising from the common case where the peak is a doublet are discussed below.

A least-squares fit is made between a fourth-degree polynomial and the window transform PSD. This polynomial, whose parameters are peak frequency (Hz) and amplitude (dB), is used as the model function for a Marquardt–Levenberg (M–L) nonlinear least-squares ''best fit'' parameter estimation routine (Press *et al.*, 1992). One begins with a rough starting estimate of peak frequency and amplitude derived from the first FFT frame. The two FFT frequencies which surround this initial frequency estimate are determined. If the initial frequency lies in the center half of the interval between these two FFT frequencies, the six FFT data points surrounding the initial frequency are chosen for best fit to the model function. If the initial frequency is in the lower or upper quarter of the interval, five FFT data points centered at the lower or upper FFT frequency are chosen. With this choice of data points the M–L process finds the model peak frequency and amplitude that minimizes the sum of squared residual differences (RSS) between the model and the selected 5/6 data points. Once the M–L routine returns an updated estimate of the parameters, the new frequency estimate is examined to see if it has crossed the subinterval boundaries described above. If not, the current parameter estimates are taken as final. If so, those values are used as initial estimates for another iteration. Iterations are limited to five to prevent limit cycles. The final fit is then inspected for acceptability. If the RSS from the M–L process is greater than the arbitrarily chosen value of 3 (a very poor fit), the fit is rejected for this peak and time frame. The previous starting estimate is then passed to the M–L routine for processing the same peak in the next time frame. This process was found empirically to lead to a reasonably good set of final peak frequency estimates for general timpani partial tracking. The choice of RSS rejection threshold is discussed in Sec. II B 2.

A fourth-degree polynomial offers, at negligible computational cost, a modestly improved fit to the window transform PSD compared with a quadratic fit. Polynomial coefficients are computed, for a polynomial of specified degree, by passing a synthesized sinusoid through the Hamming-windowed FFT process and doing a least-squares fit to the seven data points (dB versus frequency) surrounding the sinusoid frequency. Residuals for the fit are on the order of 0.05 dB for quadratic fit and 0.001 dB for fourth degree fit. The RSS can be as large as 0.006 for a second-degree polynomial and 0.000 002 for a fourth-degree polynomial, depending on the location of the sinusoid frequency between two FFT quantized frequencies. While none of these errors is large, RSS values from fits to real timpani spectral peaks are often smaller than 0.001, so the fourth-degree polynomial was chosen in order to eliminate as many uncertainties as possible in the computations.

Figure 2(a) shows the final fit to a synthesized 125.4-Hz sinusoid with added Gaussian noise whose standard deviation is 5% of the sinusoid peak amplitude. The noise values are generated by routine gasdev( ) (Press *et al.*, p. 289), which returns a normally distributed deviate with zero mean and unit variance. These noise values are multiplied by 5% of the sinusoid amplitude and added to the sinusoid sample value.

This amount of noise is clearly audible and much larger than the background noise in real timpani recordings. 125.4 Hz is near the lower boundary of the central subinterval be-

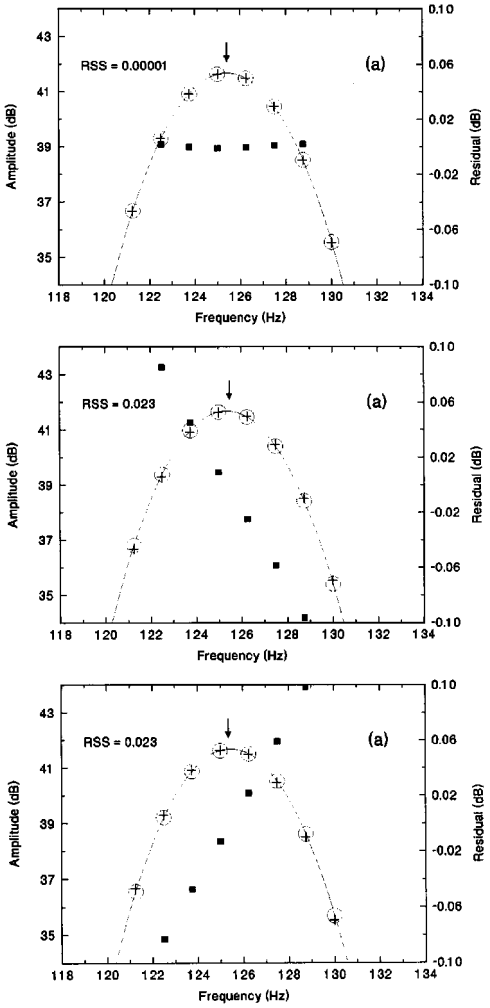


FIG. 2. Polynomial model fit to a 125.400-Hz synthesized sinusoid +5% Gaussian noise. +=sinusoid FFT; line=model prediction; ○=model prediction at FFT frequencies; ■=residual (sinusoid FFT-model). The arrow indicates the model peak frequency. Model peak frequency=(a) 125.414 Hz, (b) 125.414+0.05 Hz, (c) 125.414−0.05 Hz.

tween the two FFT data points. Figure 2(b) and (c) shows the result of changing the model peak frequency by ±0.05 Hz, respectively, from the value used for Fig. 2(a). The peak fit in Fig. 2(a) with model frequency 125.414 Hz is clearly superior to the other two both visually and as measured by the RSS. The 0.014-Hz deviation from the true frequency is discussed below.

B. Accuracy estimates

The accuracy of peak frequency estimates was investigated experimentally by synthesizing pure single sinusoids, then adding Gaussian noise and, finally, synthesizing pairs of closely spaced sinusoids. Tracker errors were then observed for these known signals.

1. Single sinusoids

a. Synthesis errors. To measure synthesis errors arising from amplitude quantization, a single constant amplitude 125.500-Hz sinusoid was synthesized by direct and also by wave table calculation and written to files in several formats with different precisions. The sinusoid amplitude was set at ±2047 arbitrary units to permit direct comparison with the 12 bit timpani sound recordings. The files contained 5000 samples or approximately 245 sinusoid half-cycles. The synthesis process was carefully checked by measuring zero crossing interval durations in the file, using linear interpolation for locating the zero crossing positions between sample points. ''Instantaneous frequencies'' were computed from the reciprocals of the zero crossing interval durations. The statistics of their differences from the correct value are presented in Table I. The rows of the table show the synthesis error when the samples are computed first in double precision floating point form, then after having been truncated to 12-bit integers and, finally, after a further truncation to 10 bits. The table demonstrates that truncation errors can cause individual instantaneous frequencies to be in error by about 0.1 Hz but that averaging over 5000 data points reduces the sample mean to a negligibly small value.

TABLE I. ''Instantaneous frequency'' values for a 5000-sample, 125.500-Hz synthesized sinusoid computed from reciprocals of the zero crossing interval duration. Expressed as deviations in Hz from the correct value for individual zero crossing intervals when the sample value amplitudes are computed in double precision floating point form, 12-bit integer form and 10-bit integer form.

	Synthesis error (Hz)			
	Max	Min	Mean	Sigma
Double precision	0.004	−0.005	0.000	0.000
12-bit integer	0.021	−0.023	0.000	0.012
10-bit integer	0.099	−0.082	0.000	0.056

Since the spectral data presented in this paper derive from 1024-sample frames, the quantization errors discussed above are presented in Table II for a set of eight data frames of that size. Individual frames show a maximum quantization error on the order of 0.001 Hz when the quantization is as coarse as 10 bits. This frequency error amounts to about 0.01 cents which is completely negligible for timpani partial tracking. At this level of detail, the waveform synchronism effect discussed below becomes apparent so the data presented in Table II is further averaged over 8 successive 1024 sample frames spaced 512 samples as a final estimate of the effects of signal amplitude quantization.

b. Tracking errors. The 0.014-Hz polynomial fit error observed in Fig. 2 can be largely attributed to lack of waveform synchronism for the FFT frames. To quantify this effect, thirty 5000-sample sinusoidal records were synthesized with frequencies from 124.625 to 128.250 Hz in steps of 0.125 Hz. The step size is 1/10 of an FFT bin width and the selected frequencies span more than two bins. A series of

TABLE II. Deviations in Hz from correct value when the individual ''instantaneous frequencies'' used to compute Table I are averaged over 8 frames of 1024 samples each. Successive frames are offset 0,512,... samples.

	Synthesis error (Hz)			
	Max	Min	Mean	Sigma
12-bit integer	−0.000 09	−0.000 11	−0.000 11	0.000 01
10-bit integer	0.000 89	−0.001 09	−0.000 13	0.000 86

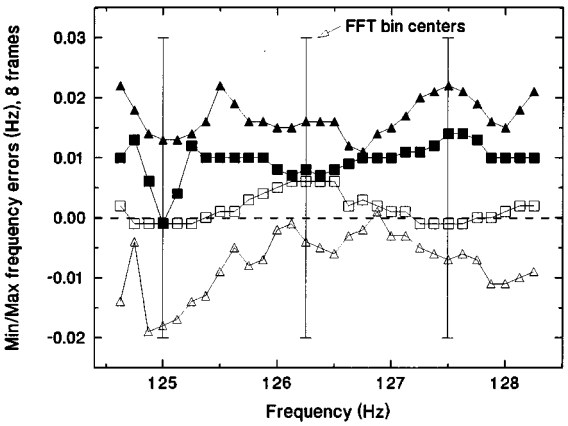


FIG. 3. Polynomial peak tracker frequency errors for synthesized sinusoids with frequencies between quantized FFT frequencies. Each vertical pair of points shows the minimum and maximum deviation from correct frequency for 8 FFT frames successively offset by 512 samples. Points □ and ■ are for pure signal only. Points △ and ▲ are for signal +5% Gaussian noise.

eight FFTs with frames successively offset by 512 samples (0.1 s) was created for each of the 30 records, and the polynomial peak tracking algorithm was applied to each FFT. Figure 3 shows the peak tracker minimum and maximum errors for a set of 8 successive FFTs at each of the 30 frequencies with and without the addition of Gaussian noise whose standard deviation is 5% of the signal peak value.

Figure 4 shows that the fit errors are smallest when the FFT frame starts and ends near zero crossings of the signal. This waveform/pitch synchronism problem has been studied extensively by the speech signal processing community (Rabiner and Schafer, 1978) and procedures have been developed to cope with it when necessary. For timpani partial studies the accuracy indicated in Fig. 3 is more than adequate. Therefore, it seems reasonable to state that for timpani spectral peak tracking the error in estimating the frequency of an isolated peak is less than ±0.05 Hz or about 1 cent. This level of accuracy compares favorably with the accuracies reported by Brown and Puckette (1993).

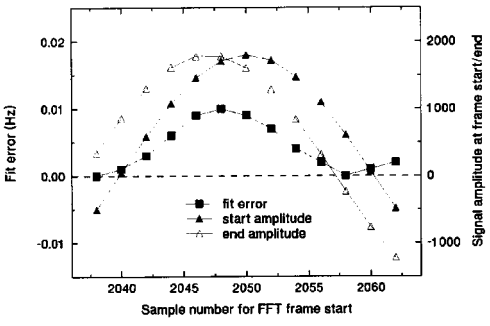


FIG. 4. Polynomial peak tracker errors on a 125.500-Hz synthesized signal related to the FFT frame start and end position with respect to the signal waveform. FFT frames are successively offset by two samples. Data is from the eighth FFT frame in Fig. 3, which has the largest error for 125.500 Hz. ▲=signal waveform amplitude at start of FFT frame; △=signal waveform amplitude at end of FFT frame; ■=fit error.

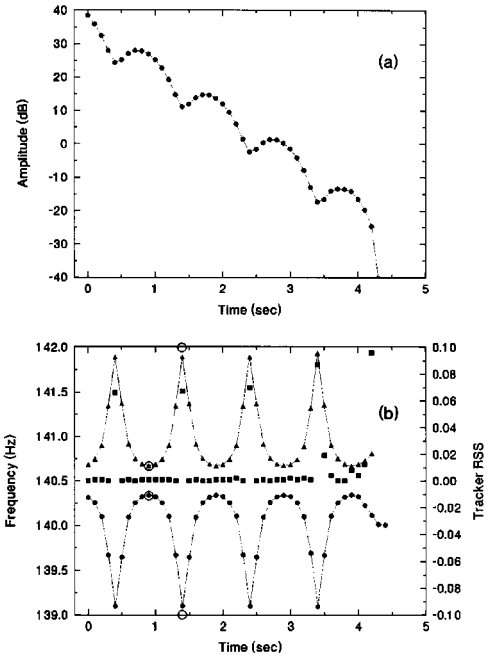


FIG. 5. Amplitude and frequency time histories reported by the tracker for a synthesized doublet consisting of decaying sinusoids with frequencies of 140 and 141 Hz and relative amplitudes of 2:1. Decay time constant=0.66 s (τ_{60}=4.6 s). (a) Amplitude. (b) Frequency: ●=140 Hz amplitude larger; ▲=141 Hz amplitude larger; ■=RSS; ○=frequency computed from Helmholtz formula.

2. Doublets

The musical modes of a vibrating kettledrum head are degenerate (Morse and Ingard, 1968) in that there are two eigenfunctions, one with sine and one with cosine azimuthal dependence, that satisfy the wave equation. In a situation where the head tension and mass distribution are homogeneous and isotropic, these two vibration modes have the same frequency. When either of those two conditions is not met, which is the usual case with timpani, the modes will have different frequencies giving rise to audible beats if the frequencies are close enough. The (31) mode in Fig. 1 has such a doublet with frequencies separated by about 1.3 Hz. The existence and cause of such doublets in bells was recognized early (Helmholtz, 1885) and, according to Rossing (1995), was known by Rayleigh to occur also in kettledrums.

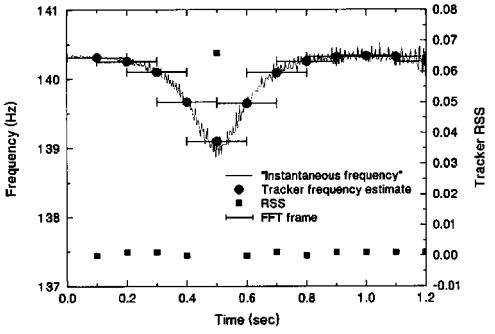


FIG. 6. Tracker frequency estimates for the synthesized doublet of Fig. 5 compared with ''instantaneous frequency'' estimates derived from signal zero crossing durations. ''Error bars'' show the 1024-sample frames used to compute FFTs for the tracker.

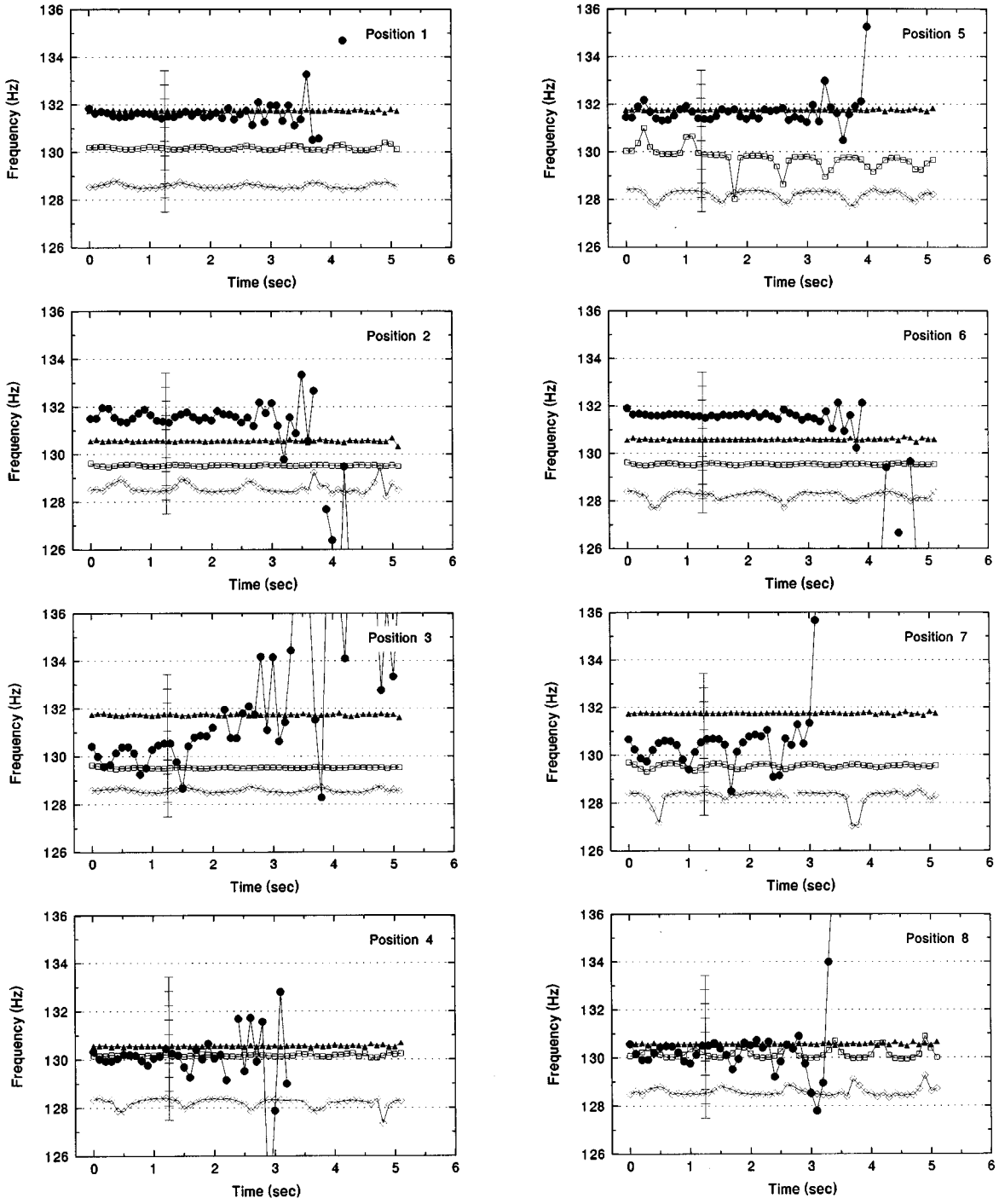


FIG. 7. Time histories of partial frequencies reported by the tracker from single strokes on a 26-in. Ludwig Professional Series drum at each of eight equally spaced azimuths. ●=principal; ▲=fifth; □=octave; ◇=tenth. Measured frequencies of the fifth, octave, and tenth are divided by 1.5, 2.0, and 2.5, respectively. The marks in the vertical string at 1.2 s are spaced at 10-cent intervals spanning a semitone.

With doublet modes the basic premise of the polynomial frequency tracker is violated. However, it turns out that the tracker performance degrades gracefully as the doublet frequency separation increases so that the tracking process remains quite useful though less accurate. This section illustrates tracker performance with doublets.

Figure 5(a) shows the amplitude time history reported by the tracker for a synthesized doublet consisting of the sum of two exponentially decaying sinusoids of different amplitudes, starting in phase and differing in frequency by 1 Hz. The amplitude ratio is 2:1 and the lower frequency component has the larger amplitude. The decay time constant of 0.66 s (τ_{60}=4.6 s) is representative of timpani partial decay rates. The lower curve in Fig. 5(b) shows the frequency time history reported by the tracker for this signal. An expression for the frequency variation in such beating signals was pro-

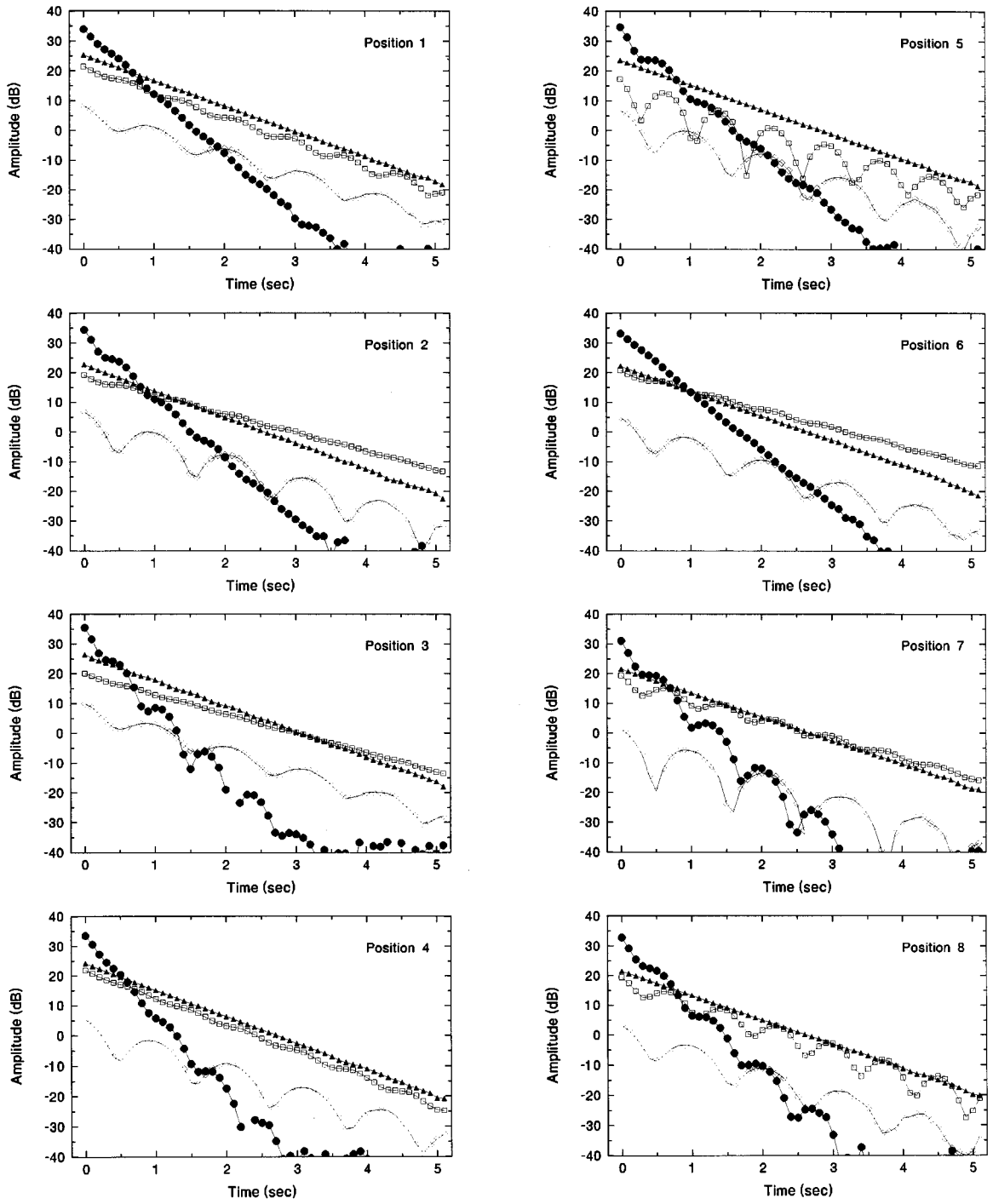


FIG. 8. Time histories of partial amplitudes reported by the tracker from single strokes on a 26-in. Ludwig Professional Series drum at each of eight equally spaced azimuths. ●=principal; ▲=fifth; □=octave; ◇=tenth.

vided by Helmholtz (1885, p. 414) as follows, with slightly revised notation. For a tone whose amplitude is of the form

$$v = A\ \sin(2\pi mt) + B\ \cos(2\pi nt + \phi),$$

where $m \approx n$ are the component frequencies in Hz and $A > B$, the frequencies of the tones at the amplitude maxima and minima are

$$f_{\max} = \frac{mA + nB}{A + B}, \quad f_{\min} = \frac{mA - nB}{A - B}. \tag{1}$$

When the component tone relative amplitudes are interchanged, as in the upper curve of Fig. 5(b), the same expression applies except that $f_{\max}$ and $f_{\min}$ are interchanged. The circles in Fig. 5(b) indicate the values of $f_{\max}$ and $f_{\min}$ predicted by Eqs. (1). When interference between the components is constructive, as at time 0.9 s, the tracker agrees with the equation within 0.004 Hz. When the interference is destructive the disagreement increases to 0.1 Hz for the reason illustrated in Fig. 6, which compares the peak tracker frequency estimates with ''instantaneous frequency'' measure-

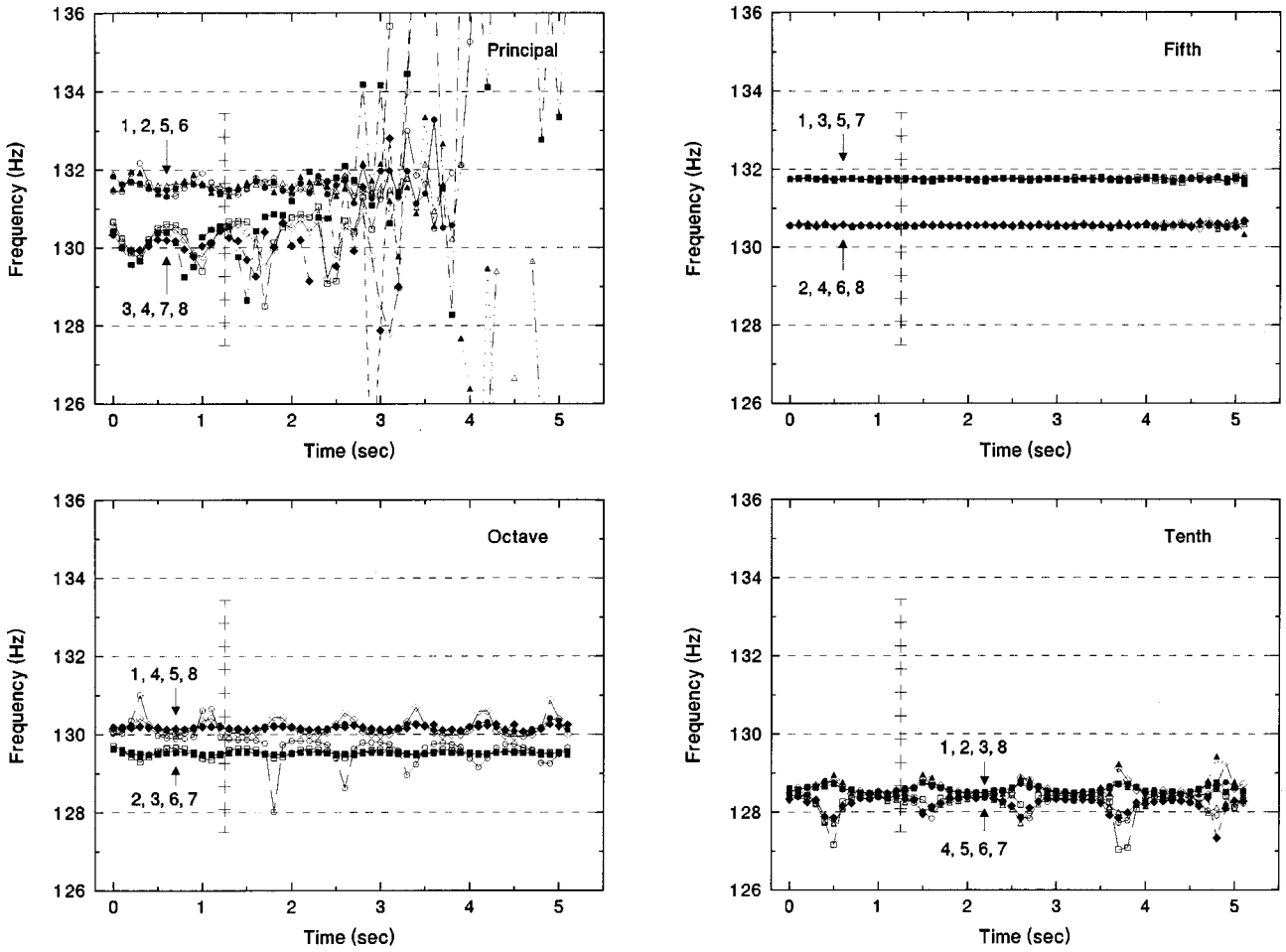


FIG. 9. Azimuthal dependence of partial frequencies reported by the tracker from single strokes on a 26-in. Ludwig Professional Series drum at each of eight equally spaced azimuths. The azimuthal positions corresponding to each partial time history are noted on the graphs. The wandering principal at position 3 observed in Fig. 7 is indicated in this figure by a ■.

ments on the same signal. The horizontal bars indicate the 1024-sample frames used to compute FFTs for the tracker and which limit the tracker's ability to follow rapidly changing frequencies.

As the frequency interval between the doublet components increases, the beat rate goes up and the tracker can no longer follow the rapid frequency changes at destructive interference. However, it can track the frequency changes at constructive interference reasonably well up to a beat rate of 2–3 Hz. A strictly empirical estimate of tracking accuracy at constructive interference is 0.05 Hz for a 1-Hz beat rate and 0.1 Hz for a 2-Hz beat rate. The RSS is a reasonable indicator for tracking accuracy as it increases rapidly when the polynomial fit gets poor for any reason. Values of RSS above 0.05 indicate that the fit is beginning to deteriorate.

A rejection threshold RSS of 3 has been used in the work reported here because it gets rid of really bad fits while retaining much visually interesting material. For example, when tracking a doublet, the peak shape can be either broader or narrower than the model function, giving rise to large RSSs. While the model fit is poor, the M–L process usually does an excellent job of locating the ''center'' of the peak so the tracker is tracking something real. Perception of the pitch of this ''something'' is beyond the scope of this paper. [See Rossing (1990) for a summary of pitch perception of complex sounds.]

III. OBSERVATIONS ON TIMPANI

The techniques described in this paper are presently being applied to various drums of different sizes and qualities in an attempt to confirm the hypothesis that the quality of timpani sound, particularly its clarity of pitch, is strongly dependent on the harmonicity of its partials. This is not a new idea. Kirby (1930) states: ''... The presence of the perfect fifth enormously increases the resonance and the beauty of tone of the drum note. ... A well-tuned drum should therefore always have the nominal and its fifth in perfect accord; if possible, the octave as well. On rare occasions I have succeeded in obtaining accurately the third, fifth, seventh, and even the double octave. ...The third, however, is almost always flat, but, as both it and the higher harmonics are not very prominent, its flatness does not affect the tone of the drum from a practical point of view.'' Kirby is clearly referring in the latter statement to the tenth since the third, as well as the seventh, is very inharmonic. Benade (1990) decribes

some partial frequency measurements on a well-tuned drum on which the fifth, octave, and tenth are within 5 cents of being harmonic and the double octave only 12 cents low.

Figure 7 shows time histories of partial frequencies from a 26-in. Ludwig Professional Series drum in the very good to excellent category. The microphone was placed 3 in. above the rim of the drum and pointing radially inward. The drum was rotated on its turntable to place the microphone at the azimuth of each of the eight tuning lugs and the drum was struck softly at the same azimuth and about 3 in. in from the rim. The azimuthal positions are numbered clockwise starting immediately left of the normal playing position. The graphs are placed so that diametrically opposed positions are horizontally adjacent.

The first four partials are plotted with the measured frequencies of the fifth, octave, and tenth divided by 1.5, 2.0, and 2.5, respectively. If all partials were harmonic they would coincide on the plot. The degree of their inharmonicity can be judged with the aid of the marks in the vertical string at 1.2 s, which are spaced at 10-cent intervals spanning a semitone.

One observes that the principal and fifth are coincident at positions 1 and 5, and nearly so at positions 4 and 8. In addition, at positions 4 and 8 the octave is also nearly coincident with the other two. The drum's owner, a professional timpanist, commented at the time these data were recorded that position 4 sounded particularly good. While this and similar comments favor the above hypothesis, a more rigorous correlation of sound quality with spectral detail would be most desirable and is being planned.

Beats at frequencies up to about 2 Hz are seen on some partials. A particularly interesting case is the octave at position 5 which shows both forms of destructive interference behavior depicted in Fig. 5(b). This is clear indication that the relative amplitudes of the two beating components are changing with time. The principal shows interesting behavior at position 3. In addition to having beats it wanders up in frequency. This phenomenon is common for the principal and sometimes observed for the fifth. Again, it indicates energy flowing from one of the two modes of a doublet to the other.

Figure 8 shows the partial decay rates corresponding to Fig. 7. Decay rates for the principal, fifth, octave, and tenth are 22, 8.2, 7.8, and 7.0 dB/s (τ_{60}=2.7, 7.3, 7.7, and 8.6 s) in this data set and are roughly independent of position. In all observed cases the principal decays much more rapidly than the other partials as predicted by Christian *et al.* (1984). This gives rise to an interesting phenomenon when the upper partials are inharmonic. If the fifth and octave, for example, are noticeably flat with respect to the principal the drum will sound flat once the principal has decayed to insignificance. This effect can be heard in many of the recorded sounds.

Figure 9 shows the data from Fig. 7 organized to highlight an azimuthal dependence of the measured partial frequencies. The frequencies of all partials fall into two groups separated by as much as 2 Hz with different patterns of high and low for each partial. The frequency difference between the two groups correlates quite well in most cases with the period of the beats with the exception of the fifth. If the data for the fifth in Fig. 9 are examined closely, the 1.8-Hz group separation frequency (1.2 Hz as plotted) is found only in the (very small) beats at positions 3 and 7.

It is important to keep in mind that the data in Figs. 7–9 were collected from eight separate strokes with the microphone and strike positions at the same azimuths. Two sets of experiments where the strike position was fixed at one azimuth and the microphone moved showed that the directional pattern relates to the microphone placement rather than the strike position. Single stroke, multiple microphone measurements are being planned. The measurements described here are near field, of course, and it is not obvious how the resulting sounds are perceived by the timpanist and his/her immediate neighbors let alone the conductor and audience.

IV. DISCUSSION

Opinions as to what constitutes good timpani sound are as varied as the timpanists but two of the most important factors appear to be the ability to produce a clear, focused pitch and the ability to achieve a good resonant sound, i.e., where partials such as the fifth and octave decay slowly. This paper has concentrated on methods to quantify the constituents of clear, focused pitch. Goals for future work are to correlate specifics of partial behavior with quality of sound and, if possible, to relate these to construction details and fine tuning procedures.

ACKNOWLEDGMENTS

The author is most grateful to timpanist Everett Beale for providing several drums, helping with the data collection from same, and participating in preliminary sound quality judgments, all with patience and good humor. Thanks are also due to timpanist Ed Sterling for helpful discussions and the use of his drums. Timpanist Rebecca Kite supplied some very interesting material on the process of clearing, or fine tuning, timpani heads, and timpanist Vic Firth started it all by questioning an early assertion regarding pitch changes following a heavy stroke.

Benade, A. H. (**1990**). *Fundamentals of Musical Acoustics* (Oxford U.P., New York, 1976; second, revised edition reprinted by Dover, New York, 1990), pp. 143–145.

Brown, J. C. (**1996**). ''Frequency ratios of spectral components of musical sounds,'' J. Acoust. Soc. Am. **99**, 1210–1218.

Brown, J. C., and Puckette, M. S. (**1993**). ''A high resolution fundamental frequency determination based on phase changes of the Fourier transform,'' J. Acoust. Soc. Am. **94**, 662–667.

Christian, R. S., Davis, R. E., Tubis, A., Anderson, C. A., Mills, R. I., and Rossing, T. D. (**1984**). ''Effects of air loading on timpani membrane vibrations,'' J. Acoust. Soc. Am. **76**, 1336–1345.

Fletcher, N. H., and Rossing, T. D. (**1990**). *The Physics of Musical Instruments* (Springer-Verlag, New York).

Helmholtz, H. L. F. (**1885**). *On the Sensations of Tone* (Longmans, London; second English edition translated by A. J. Ellis; Reprinted in 1954 by Dover, New York).

Kirby, P. R. (**1930**). *The Kettle-Drums, A Book for Composers, Conductors and Kettle-Drummers* (Oxford U.P., London), pp. 40–43.

Morse, P. M., and Ingard, K. U. (**1986**). *Theoretical Acoustics* (Princeton U.P., Princeton, NJ), p. 212.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (**1992**). *Numerical Recipes in C* (Cambridge U.P., New York), 2nd ed., p. 683.

Rabiner, L. R., and Schafer, R. W. (**1978**). *Digital Processing of Speech Signals* (Prentice-Hall, Englewood Cliffs, NJ).
Rossing, T. D. (**1982**). ''The physics of kettledrums,'' Sci. Am. **247** (5), 172–178.
Rossing, T. D. (**1990**). *The Science of Sound* (Addison-Wesley, Reading, MA), Chaps. 7–8.
Rossing, T. D. (**1995**). ''Lord Rayleigh's contributions to musical acoustics,'' J. Acoust. Soc. Am. **98**, 2406–2412.

Numerical simulations of xylophones. I. Time-domain modeling of the vibrating bars

Antoine Chaigne[a] and Vincent Doutaut
Ecole Nationale Supérieure des Télécommunications, Département SIGNAL, CNRS URA 820, 46 Rue Barrault, 75634 Paris Cedex 13, France

(Received 26 March 1996; revised 21 August 1996; accepted 5 September 1996)

A time-domain modeling of xylophone bars excited by the blow of a mallet is presented. The flexural vibrations of the bar, with nonuniform cross section, are modeled by a one-dimensional Euler–Bernoulli equation, modified by the addition of two damping terms for the modeling of losses and a restoring force for the modeling of the stiffness of the suspending cord. The action of the mallet against the bar is described by Hertz's law of contact for linear elastic bodies. This action appears as a force density term on the right-hand side of the bending wave equation. The model is completed by the equation of motion for the mallet, and by free–free boundary conditions for the bar. The bending wave equation of the bar is put into a numerical form by means of an implicit finite-difference scheme, which ensures a sufficient spatial resolution for an accurate tuning of the bar. The geometrical, elastic, and damping parameters of the model are derived from experiments carried out on actual xylophones and mallets. The validity of the numerical model is confirmed by three different procedures: First, a comparison is made between numerical results and analytical solutions. The second series of tests consist of examining the effects of the number of spatial steps on the convergence of the solution. Finally, various comparisons are made between measured and simulated impact forces and bar accelerations. The present model reproduces adequately the main features of a real instrument. The most significant physical parameters of bars, mallets and players' actions can be controlled independently for producing a remarkable variety of tones. [S0001-4966(97)02201-7]

PACS numbers: 43.75.Kk [WJS]

INTRODUCTION

The xylophone is a percussion instrument made of vibrating bars with free ends struck by a mallet. Wood is the most common material used for the bars, but other synthetic materials, such as fiberglass, may also be used.[1] The cross sections of xylophone bars are not uniform, for tuning purpose. In some musical circumstances, a tubular resonator may be placed below each bar in order to modify the sound-pressure field. In this case, the fundamental frequency of the resonator is generally tuned close to the lowest eigenfrequency of the bar. As a consequence, the resonator modifies markedly the temporal evolution of the emitted sound, with an ''aftersound'' which contains the fundamental frequency only. For a more complete description of the xylophone, and a more exhaustive discussion on the use of the instrument, one can refer, for example, to Moore.[2] A general presentation of the acoustics of mallet percussion instruments can be found in Refs. 3 and 4.

An extensive study of the tuning of the bars, and of their coupling with the resonators, has been made by Bork.[5,6] The same author also addressed the problem of measuring the properties of mallets, in connection with their musical use in percussion instruments.[7] More recently, a numerical method for determining the profile of the undercut, for a given tuning of the bar, has been proposed by Orduña-Bustamante.[8] However, the number of published papers devoted to xylophones, and to percussive instruments in general, is relatively small in comparison with the literature on wind and string instruments.

The equations governing the motion of waves in elastic bars can be found in many textbooks.[9] For many decades, the problem of the impact of a mass striking a beam has aroused a strong interest in the mechanical engineering community.[10,11] The most famous result related to this problem is Hertz's law of contact, which was established more than a century ago.[12] One example of application of Hertz's law in the field of musical acoustics can be found in a paper by Gridnev dealing with the impulsive excitation of a guitar top plate by means of the impact of a metal ball.[13] Surprisingly, the combination of these well–known results hasn't been applied in the past for the modeling of mallet percussive instruments.

The prime motivation of this paper is to test, by means of simulations, the basic principles that govern the physics of xylophone bars. Then, the numerical model is used for investigating the relevance of bar and mallet's properties, and of the main parameters of the player's action, on the tonal qualities of the produced sound. This preliminary model may be improved in the future, if evident deficiencies in the results clearly show the necessity of a higher degree of complexity. A similar approach has been used in the past for other families of instruments.[14–16] The relevance of the present model is assessed by a systematic comparison with an actual instrument, which serves here as a reference. However, the purpose of the present work is not only limited to

[a]Electronic mail: chaigne@sig.enst.fr

the mimicry of existing xylophones, but also to the exploration of a wide range of geometric and elastic properties for bar and mallets, which could be probably never achieved by makers.

The simplified model presented below is based on the one-dimensional Euler–Bernoulli equation for the flexural vibrations of a vibrating bar, with nonuniform cross section, excited by the blow of a mallet. In this model, it is assumed that the nonlinear bar–mallet interaction is governed by Hertz's law of contact. The bar model is refined by taking the damping of the flexural free vibrations into account, using a viscoelastic formulation for the losses. The determination of the damping constants is the result of experimental measurements on real wooden bars. The set of equations, which composes the model, is then put into a numerical form using standard finite difference methods.[17,18] The results of the numerical simulations can be heard after appropriate digital-to-analog conversion, and compared with actual waveforms measured on existing instruments. A time-domain approach appears to be particularly attractive here, due to the perceptual relevance of initial transients and decay times for percussive sounds.[19]

The paper is organized as follows: the time-domain one-dimensional model for the vibrating bar interacting with a mallet is presented in Sec. I. The corresponding numerical formulation of the problem is presented in Sec. II, where emphasis is put on the required spatial resolution, for an accurate tuning of the bar. The experiments carried out on actual xylophone and mallets for extracting the relevant physical parameters of the model are described in Sec. III. Finally, in order to confirm the validity of the numerical method, various tests are carried out, and presented in Sec. IV. The first test consists of a comparison between numerical and analytical results in two simplified situations: an infinite bar excited by a Gaussian pulse, and a finite bar excited by a Dirac delta function. The numerical tests are then extended by tests of convergence, where the goal is to examine the effects of increasing the number of spatial steps on the accuracy of the solution, with special emphasis on the tuning of the synthetic bar. In the same section, a systematic exploration of some mallet–bar parameters is presented, which is aimed at illustrating the wide variety of tones produced by the simulation program, with respect to the main parameters of control, in the normal use of the xylophone. This section ends with a comparison between real and simulated waveforms and spectra.

The present paper is limited to the presentation of the vibrational phenomena only. The method used for the modeling of the sound-pressure field, due to the radiation of the vibrating bar coupled with a tubular resonator, will be examined in a forthcoming study.

I. BASIC MODEL

A. Transverse motion of a thin bar

In the present paper, it is assumed that the vertical component $w(x,t)$ of the displacement of a xylophone bar (see Fig. 1) is governed by the following equations:

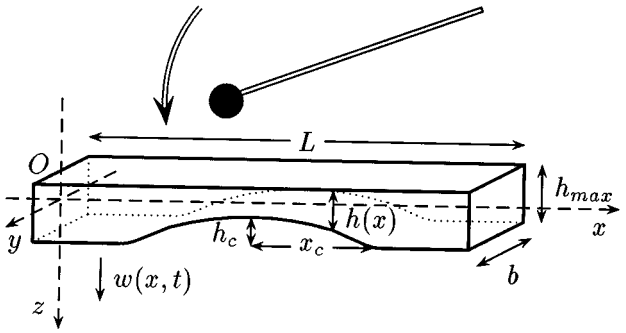


FIG. 1. Geometry of the xylophone bar.

$$M(x,t) = -EI(x)\left(1+\eta\frac{\partial}{\partial t}\right)\frac{\partial^2 w}{\partial x^2}(x,t),$$

$$\frac{\partial^2 w}{\partial t^2}(x,t) = \frac{1}{\rho S(x)}\frac{\partial^2 M}{\partial x^2}(x,t) - \gamma_B\frac{\partial w}{\partial t}(x,t) - \frac{\chi}{M_B}w(x,t) + f(x,x_0,t), \tag{1}$$

where $M(x,t)$ is the bending moment, and where $S(x)$ and $I(x)$ denote the cross section and its moment about the x axis, respectively. For a rectangular cross section we have

$$S(x) = bh(x) \quad \text{and} \quad I(x) = bh^3(x)/12, \tag{2}$$

where b is the width of the bar (see Fig. 1).

With the assumption of free–free boundary conditions, one obtains

$$\frac{\partial^2 w}{\partial x^2}(x,t) = \frac{\partial^3 w}{\partial x^3}(x,t) = 0\Big|_{x=0,L}. \tag{3}$$

This amounts to assuming, first, that the sound of a xylophone is due to the flexural motion of the bar in the vertical xz plane only, which amounts to neglecting the contribution of torsional and longitudinal waves as well as the flexural waves in the horizontal xy plane. Equation (1) secondly implies that this motion is conveniently described by the one-dimensional Euler–Bernoulli equation. A number of measurements carried out in the past on the instrument have shown that this latter assumption is valid, to a first degree of approximation, in the low-frequency range.[20] For current xylophones, it is generally observed that using Euler–Bernoulli model leads to a tuning error less than 2% in the lowest two octaves of the instrument, i.e., for fundamental frequencies below 1.2 kHz.[19]

In Eq. (1), E is the Young's modulus and ρ is the density. The material is assumed to be homogeneous and isotropic, which is clearly an approximation in the case of wood. However the bars are generally cut along the direction of fibers, in order to make them more resistant in case of strong impacts, and it turns out that Eq. (1) is an acceptable formulation for bending waves in wooden xylophone bars, assuming that E in this latter equation is taken equal to the longitudinal modulus of elasticity.[21] The thickness $h(x)$ is a function of the spatial coordinate x along the length of the bar, which accounts for the cutting of an arch under the bar in order to tune it.[4]

The damping of the flexural waves are represented in Eq. (1) by two terms. The first term, whose magnitude is proportional to the coefficient η, represents the viscoelastic losses, expressed by the following relationship between stress σ and strain ϵ in the material:

$$\sigma(x,t)=E\left(\epsilon(x,t)+\eta\frac{\partial\epsilon}{\partial t}(x,t)\right). \tag{4}$$

The coefficient η can also be defined as the time constant of the creep function of the material subjected to a suddenly applied step of stress.[22] The second term, whose magnitude is represented by the coefficient γ_B, is equivalent to a fluid damping. It is shown in Sec. III that the introduction of viscoelastic and fluid damping terms in Eq. (1) yields a good representation of losses in wooden bars. Both coefficients η and γ_B are derived from measurements of the partials' decay times on actual instruments.

Xylophones bars are usually supported elastically by means of a flexible thin cord passing through holes drilled at positions which correspond roughly to the nodal points for the lowest mode of the free–free bar.[6] In Eq. (1) the elasticity of the cord is modeled by spring-like forces equally distributed over the length L of the bar. The order of magnitude for the stiffness coefficient χ is derived from measurements of the natural frequency f_B of the spring–mass system made of the combination of the bar, of mass M_B, and the cord. This frequency is usually very low, typically 20 Hz, and thus it is reasonable to assume that the natural frequencies of the bars are not perturbed by the presence of the cord. For the lowest note (F_4=352 Hz) of a $3\frac{1}{2}$-oct xylophone keyboard, for example, the fundamental frequency of the bar is nearly 18 times the resonance frequency of the cord. The main function of the cord is to act as a high-pass filter, which eliminates, among other things, the zero-frequency modes corresponding to translation and rotation of the bar. It is also assumed that there is no coupling with the adjacent bars.

Finally, the force density term $f(x,x_0,t)$ in Eq. (1) represents the striking action of the mallet, which sets the bar into vibration. The contact duration is the result of a complex interaction between bar and mallet.

B. Interaction between bar and mallet

The model of the interaction between the bar and the mallet is similar in form to the one previously used by the first author for the modeling of hammer–string interaction in the piano.[16] The force density term $f(x,x_0,t)$ is related to the impact force $F(t)$ by the relation:

$$f(x,x_0,t)=f(t)g(x,x_0)$$

$$\text{with}\quad f(t)=\frac{F(t)}{\rho S(x)\int_{x_0-\delta x}^{x_0+\delta x}g(x,x_0)dx}. \tag{5}$$

The spatial window $g(x,x_0)$ represents the distribution of the force over the width $2\,\delta x$ of the impact, which can be estimated with the help of Hertz's model. Ordinary values for the impact width lie within the range 1 to 5 mm, depending on the initial velocity and the material of the mallet's head (see Appendix A). The advantage of the spatial window is to remove the discontinuities in the spatial distribution of the impact, and thus to limit the risk of parasitic oscillations in the solution. Because of the difficulty of defining the exact shape of $g(x,x_0)$ on the basis of physical criteria, it has been decided to use simple mathematical functions, such as a triangular or cosine windows, for spatially smoothing the impact. It has been observed in the simulations that small variations in the shape of the window have no significant influence on the solution. Finally, it has been assumed that the surface of contact between bar and mallet remains constant during the impact, which is not true in reality (see Appendix A). However, as it will be shown in Sec. IV, this simplifying assumption doesn't seem to have detectable consequences in simulated waveforms and spectra, except for a limited number of cases, for example when the contact point is situated near an antinode.

The impact force is given by Hertz's law of contact[23]

$$F(t)=K\left|\zeta(t)-w(x_0,t)\right|^{3/2}, \tag{6}$$

where $\zeta(t)$ is the displacement of the mallet's head. An important difference from the hammer–string case is that the nonlinear exponent 3/2 in the power law between force and deformation in Eq. (6) derives here from the general theory of elasticity (see Appendix A), whereas, in the case of pianos, the exponent of this nonlinear force-deformation relationship is derived from curve fitting on experimental data.[24,25] The experimental determination of the stiffness coefficient K, and the validity domain of Hertz's model for xylophones, are discussed in Sec. III.

The motion of the mallet is governed by Newton's law:

$$m_e\frac{d^2\zeta(t)}{dt^2}=-F(t)\quad\text{with}\quad\frac{d\zeta(0)}{dt}=V_0, \tag{7}$$

where m_e represents the equivalent mass of the mallet. This equivalent mass is defined and measured in Sec. III. In practice m_e is slightly larger than the mass of the head, due to the presence of the stick held by the player's hand. V_0 is the initial velocity of the mallet's head, at the time where it comes in contact with the bar. The flexibility of the stick is neglected.

Other complicating factors may exist in reality, which are not taken into account in the model. First, for a bar of finite length, it can be expected that the elastic waves reflected at both ends interfere with the blow. Second, due to the strength of the blow, the material of the bar may present a plastic behavior, which in turn leads to a permanent deformation of the bar. Finally for strong impacts with a soft mallet's head, for example, the strain may become so large that the assumptions of linear elasticity are no more valid.[22] Consequently, Hertz's law of contact doesn't apply in these cases. Measurements on mallets presented in Sec. III illustrate these different points.

In summary, the basic model presented above is able to account for the following three major physical aspects of the instrument:

(1) The introduction of a kinetic energy, localized in time and space, into the vibrating system.

(2) The influence of the initial velocity on both the contact duration and the impact force, due to the nonlinear force-deformation law, which determines the spectrum of the tone.

(3) The influence of the stiffness of the two materials in contact, which strongly determines the tone quality of the initial blow.

II. NUMERICAL FORMULATION

For a time-domain numerical formulation of a continuous problem, a necessary step is to define the spatial grid where the continuous equations are to be approximated, and the time interval (or time step) between two consecutive discrete values of the variables under examination. The numerical formulation of the xylophone model, presented in the previous section, is obtained here by means of finite difference methods. It is well known that these techniques are convenient if the geometry of the structure can be simply expressed in a standard system of coordinates, such as the Cartesian coordinates used here for the one-dimensional bar. In this case, the only parameter of the spatial grid to be determined is the elementary step $\Delta x = L/N$, the bar of length L being divided into N equally spaced segments. The time step Δt is simply the inverse of the sampling frequency f_s. The selection of these two parameters is generally imposed by stability and frequency warping criteria. In the case of sound synthesis, the general results of human auditory perception, especially those related to the sensation of pitch, can provide very useful additional guidelines, particularly for estimating the accuracy required for the estimation of frequencies.

For the xylophone model, the first idea was to use an explicit finite difference scheme of second order in time and space (referred to as a 2-2 scheme), similar to the one previously used for piano strings.[16] The main advantage of such a scheme is to transform the basic partial differential equation into a simple recurrence equation, which leads to very fast and efficient algorithms, with relatively low numerical dispersion. However, due to the presence here of fourth-order partial derivatives with respect to spatial coordinates and of viscoelastic losses in the bending wave equation, the stability condition of an explicit scheme becomes too drastic, and the method can only be used in restricted cases such as a bar with a constant section.

In the case of nonuniform xylophone bars, an implicit scheme has to be used, due to the spatial accuracy required for the modeling of the variable thickness. This method leads to a more precise, but also more time-consuming, matrix formulation of the numerical problem. In order to illustrate the stability and accuracy requirements of the numerical method, the reference case of a bar with a constant section is presented first, and followed by the presentation of the implicit numerical scheme used in our simulations.

A. Uniform bar of constant section

For a bar of constant section, one can write $h(x)=h$, and thus Eq. (1) becomes

$$\frac{\partial^2 w(x,t)}{\partial t^2} = -a^2\left[\frac{\partial^4 w(x,t)}{\partial x^4} + \eta\,\frac{\partial^5 w(x,t)}{\partial t\,\partial x^4}\right] - \gamma_B\,\frac{\partial w(x,t)}{\partial t} - \frac{\chi}{M_B}\,w(x,t) + f(x,x_0,t), \tag{8}$$

with

$$a^2 = \frac{EI}{\rho S}. \tag{9}$$

Approximating Eq. (8) by means of explicit finite differences of second order in time and space yields the following recurrence equation (see Appendix B):

$$\begin{aligned} w_i^{n+1} = {} & c_1 w_i^n + c_2 w_i^{n-1} + c_3[w_{i+2}^n - 4w_{i+1}^n - 4w_{i-1}^n \\ & + w_{i-2}^n] + c_4[w_{i+2}^{n-1} - 4w_{i+1}^{n-1} - 4w_{i-1}^{n-1} + w_{i-2}^{n-1}] \\ & + c_5 F(n) g(i,i_0), \end{aligned} \tag{10}$$

with

$$\begin{aligned} & c_1 = \frac{2 - 6r^2(1+\beta) - (\Delta t\,\omega_B)^2}{1+\gamma}, \quad c_2 = \frac{-1+\gamma+6\beta r^2}{1+\gamma}, \\ & c_3 = \frac{-r^2(1+\beta)}{1+\gamma}, \quad c_4 = \frac{\beta r^2}{1+\gamma}, \quad c_5 = \frac{N}{M_B f_s^2 (1+\gamma)}, \\ & \beta = \eta f_s, \quad \omega_B^2 = \frac{\chi}{M_B}, \quad \gamma = \frac{\gamma_B}{2 f_s}, \quad r = a\,\frac{N^2}{f_s L^2}, \end{aligned} \tag{11}$$

where i and n are the spatial and time indices, respectively, and where the notation w_i^n represents the value of the vertical displacement of the bar at position $x_i = i\Delta x$ and at time $t_n = n\Delta t$.

Using standard procedures in numerical analysis,[26] it can be shown that the explicit scheme expressed in Eq. (10) remains stable under the following condition (see Appendix B):

$$N \leqslant N_{\max} = \frac{3}{4}\,\sqrt{\frac{\pi f_s}{f_1(1+\eta f_s)}}, \tag{12}$$

where f_1 is the frequency of the lowest partial. This method yields satisfactory results as long as $\eta f_s \ll 1$. For wooden bars, the order of magnitude is $\eta f_s = 10^{-2}$.

Intuitively, one can expect that the stability condition expressed in (12) is also valid for a xylophone bar with an undercut, since the cutting helps to reduce f_1, and thus to increase $N_{\max}$, and this property is effectively observed in the simulations. However, the exact demonstration is tedious. It has been given by Cohen, for example, for the wave equation using energetic methods,[27] but, as far as we know, such demonstration remains to be done in the case of bars.

Two main conclusions can be drawn from Eq. (12). First, it can be seen that, due to the presence of terms involving a fourth-order partial derivative versus space in Eq. (8), the maximum number of spatial points $N_{\max}$ is roughly proportional here to the square root of the sampling frequency f_s, for low values of f_s, since $\eta \ll 1$. This means practically that it is necessary to multiply f_s by a factor of 4 in order to improve the spatial resolution by only a factor of 2. In comparison, recall that using the same finite difference scheme in the case of strings with low stiffness yields $N_{\max}$ roughly proportional to f_s, which is a more desirable condition.[16]

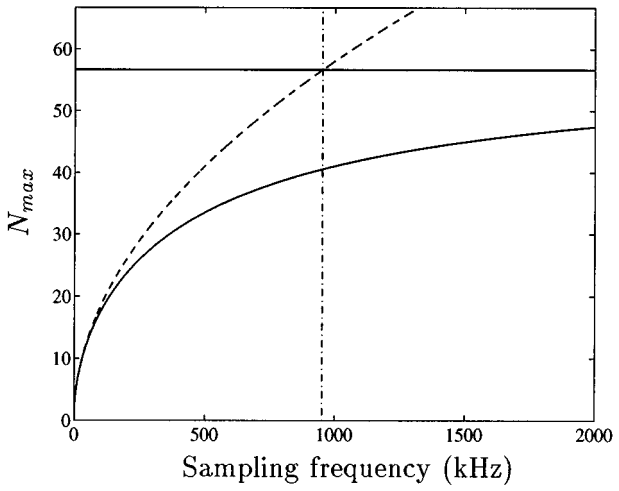


FIG. 2. The stability condition of the explicit finite difference scheme is illustrated by the maximum number of spatial steps $N_{\max}$ as a function of the sampling frequency f_s, for $\eta=0$ (dashed line) and $\eta=10^{-6}$ s (solid line). The horizontal line is the asymptotic limit for $N_{\max}$, as f_s tends to infinity. This example is given for a bar with fundamental $f_1=524$ Hz.

Second, Eq. (12) shows that $N_{\max}$ tends to the asymptotic limit $\frac{3}{4}\sqrt{\pi/f_1\eta}$ as f_s increases (see Fig. 2). Although such limit is physically unrealistic, it shows in practice that it becomes useless to increase f_s beyond a certain limit, since it will not improve substantially the spatial resolution.

These stability properties show the limits of the explicit second-order finite-difference scheme for the viscoelastic Euler–Bernoulli bending wave equation. A sampling frequency of 192 kHz, for example, yields a minimum spatial step Δx approximately equal to 1 cm. For this spatial resolution it is observed that the first natural frequencies of the bar are badly estimated, and consequently, that the pitch of the simulated bar can differ substantially from the one of the corresponding real bar. In addition, this spatial step is 2 to 10 times greater than the impact width, which means that the localization of the impact is not modeled correctly. Therefore, it can be concluded that the use of explicit schemes leads to prohibitive values of the spatial step, and that more accurate numerical schemes have to be used for xylophone bars.

B. Bar with variable section

An explicit system gives the unknowns w_i^{n+1} directly in terms of the known quantities w_i^n. If not, one must solve a set of simultaneous linear equations to obtain the w_i^{n+1}, and the system is called *implicit*. A general class of implicit finite difference schemes, usually called θ schemes, consists of approximating the fourth-order spatial derivative in Eq. (1) by the following three-level time average (see Appendix B):

$$\frac{\partial^2}{\partial x^2}\left[I(x)\,\frac{\partial^2 w}{\partial x^2}\,(x,t)\right] \approx \theta[\,\delta_x^2[I_i(\delta_x^2 w)_i]^{n+1}+\delta_x^2[I_i(\delta_x^2 w)_i]^{n-1}] + (1-2\theta)\,\delta_x^2[I_i(\delta_x^2 w)_i]^n, \tag{13}$$

with

$$\delta_x^2[I_i(\delta_x^2 w)_i]^n=\frac{I_{i+1}w_{i+2}^n-2(I_{i+1}+I_i)w_{i+1}^n+(I_{i+1}+4I_i+I_{i-1})w_i^n-2(I_i+I_{i-1})w_{i-1}^n+I_{i-1}w_{i-2}^n}{\Delta x^4} \tag{14}$$

and

$$0\leqslant\theta\leqslant 1.$$

This technique is equivalent to a ''smoothing'' of the spatial derivative with time, through application of a triangular window with parameter θ, centered at discrete time $t_n=n\Delta t$. The limiting case $\theta=0$ corresponds to the explicit scheme.

Using the approximation expressed in Eq. (13), then the numerical formulation of the general equation (1) becomes

$$(1+\gamma)w_i^{n+1}+(\omega_B^2\Delta t^2-2)w_i^n+(1-\gamma)w_i^{n-1} - \frac{\Delta t^2}{\Delta x}\,\frac{F(n)}{\rho S_i}\,g(i,i_0) = -\frac{E\Delta t^2}{\rho S_i}\,[\,\theta((1+\beta)\,\delta_x^2[I_i(\delta_x^2 w)_i]^{n+1}+(1-\beta) \times\delta_x^2[I_i(\delta_x^2 w)_i]^{n-1})+(1-2\theta)\,\delta_x^2[I_i(\delta_x^2 w)_i]^n]. \tag{15}$$

It can be shown that the θ scheme remains unconditionally stable if $\theta\geqslant 1/4$.[28] This latter result has a main advantage: it implies that the time step Δt and the spatial step Δx can be selected independently, without any risk of instability. The best estimation of natural frequencies is obtained for $\theta=1/4$.[19] A similar result has been previously obtained in the case of strings.[29] With a sampling frequency of 192 kHz, for example, it now becomes possible to spatially sample the bar with a better resolution than with the explicit scheme. In practice, a spatial resolution Δx of 1 to 3 mm has been selected in our simulations. This range is sufficient for accurately tuning the bar and for the modeling of the impact.

Due to the use of the θ approximation presented above, the numerical formulation of Eq. (1) becomes implicit, since approximations at time $n+1$ now appear on both sides of Eq. (15). Therefore, it is no longer possible to calculate the value of the displacement for one single discrete point w_i^{n+1} by means of a recurrence equation similar to Eq. (10), and a matrix formulation of the discrete problem has to be used.

The simulation program is written in C programming language and runs on a Sun Sparc10 workstation. It has been observed that the computing time is proportional to the product $N_s N^2$, where N_s is the duration in samples. For $N=100$

and $f_s=192$ kHz, the mean computing time is approximately equal to 1000 s per second of simulated signal. This corresponds equivalently to the calculation of 200 samples/s.

It must be stressed that the selection of the sampling frequency $f_s=192$ kHz is fully imposed by numerical considerations on the accuracy of the solution of the physical model presented in Eqs. (1)–(7). For audio applications, the data files are low-pass filtered at 48 kHz, and undersampled by a factor of 4, in order to be compatible with the standard digital equipment.

III. EXPERIMENTS ON REAL INSTRUMENTS

The basic equations (1)–(7) presented in Sec. I yield the general structure of the model. In order to simulate realistic tones, it is now essential to feed reasonable values of the physical parameters into the numerical formulation. These parameters have been measured on actual bars and mallets, so that the program can be tested through comparisons between simulated and measured waveforms. Another major interest of measurements on actual instruments is to show the limits of validity of the physical bases used in the model. This point is illustrated in this section with a discussion on the validity of Hertz's law for the modeling of the blow. This section starts with the presentation of measurements on the mallets and is followed by the presentation of elastic properties and damping factors of the bar.

A. Measurements on mallets

The action of the mallet is characterized by its effective mass m_e and by its initial velocity V_0 at the time of impact. In addition, the stiffness coefficient K [see Eq. (6)] depends on Young's moduli and Poisson's ratios of both solids in contact (see Appendix A).

The experimental determination of m_e is made as follows: the mallet's head first strikes a rigidly fixed impedance head which delivers two signals, proportional to the force $F(t)$ and to the acceleration $\Gamma(t)$, respectively. These two signals are very similar in form, and the effective mass is obtained by calculating the transfer function between acceleration and force on a dual-channel FFT analyzer. Measured values of m_e are given in Table I. It has been observed that these values are slightly higher (about 20%) than the mass of the head itself, and that the equivalent mass tends to increase with the magnitude of the force. This latter effect may be due to the inertia of the player's arm and handle. In order to make measurements of the striking force, a miniature accelerometer (B&K 4374) of mass $m_a=0.65$ g is glued on the stick near the head. The force is then obtained by multiplying the acceleration signal by the equivalent mass previously determined on the impedance head. A correction is applied, due to the presence of the accelerometer on the stick, which is responsible for an increase of 2% to 4% of the equivalent mass.

Integrating the acceleration signal provides us with the velocity $v_m(t)$ of the mallet from which the initial velocity V_0 is derived. The results obtained are in accordance with the values predicted by Hertz's law, using measurements of the interaction time τ and maximum force $F_{\max}$ (see Appendix A). In practice, experimental determination of the interaction time τ is carried out by measuring the interval between the origin of time and the first zero of $\Gamma(t)$. The remaining oscillations, visible in $\Gamma(t)$ for $t>\tau$, are due to the vibration of the stick after the impact.

TABLE I. Typical values of the parameters used in the simulations.

Bar No. 5(G_4: 397 Hz)—xylophone CONCORDE X4001—Padouk
Length $L=29.3$ cm
Width $b=3.7$ cm
Maximum thickness (edge) $h=1.9$ cm
Minimum thickness (center) $h=0.8$ cm
Total mass $M_B=142.04$ g
Young's modulus $E=9.54\times10^9$ N m^{-2}
Density $\rho=796$ kg m^{-3}
Viscoelastic constant $\eta=9.16\times10^{-8}$ s
Fluid damping coefficient $\gamma_B=12.44$ s^{-1}
Elastic parameters of a Rosewood bar
Young's modulus $E=2.13\times10^{10}$ N m^{-2}
Density $\rho=1015$ kg m^{-3}
Viscoelastic constant $\eta=2.35\times10^{-7}$ s
Fluid damping coefficient $\gamma_B=51.58$ s^{-1}
Mallets
Stiffness coefficient—Korogi No. 4 (rubber) $K=3.7\times10^7$ N m$^{-3/2}$
Stiffness coefficient—Vibrawell $\times7$ (boxwood) $K=1.31\times10^9$ N m$^{-3/2}$
Equivalent mass (rubber) $m_e=23.6$ g
Equivalent mass (boxwood) $m_e=26.2$ g
Player's action
Position of impact $x_0=15$ to 20 cm
Position of observation (acceleration) $x=9.3$ cm
Initial impact velocity $V_0\in[0.05;\quad 1.5]$ m s^{-1}
Sampling
Time sampling frequency $f_s=192$ kHz
Spatial step $\Delta x=1.0$ to 3.0 mm
Impact width $\delta x=2.0$ mm

The stiffness coefficient K is estimated by using the general results of contact theory. This problem has been completely solved by Landau in the case of the contact between two elastic spheres of different materials, assuming that the vibrations of the spheres are negligible.[23] One of the two spheres is replaced here by a sphere of infinite radius, which represents the bar. Within the framework of this theory, it is shown in Appendix A that K can be expressed as a function of τ and $F_{\max}$, as follows:

$$K=35.4\,\frac{1}{\tau^3}\,\sqrt{\frac{\mu^3}{F_{\max}}}, \tag{16}$$

where

$$\mu=\frac{m_e M_B}{m_e+M_B} \tag{17}$$

is the reduced mass between mallet and bar. In practice, a number of measurements of $F_{\max}$ are conducted, for various conditions of impact, and K is derived from a linear regression between $\log F_{\max}$ and $\log \tau$ (see Fig. 3). The obtained values are in agreement with those predicted by the model of elastic spheres given by Landau,[23] through application of Hertz's theory.

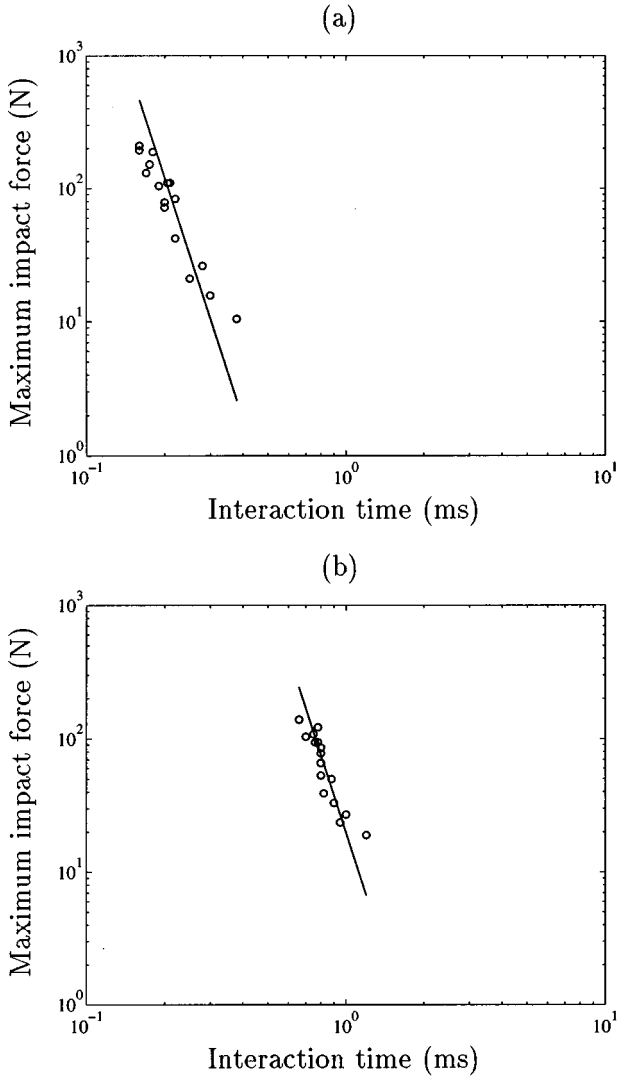


FIG. 3. Maximum of the impact force F_{max} as a function of the interaction time τ between bar and mallet, for a (a) hard mallet in boxwood and a (b) soft mallet in rubber. The experimental data are represented by circles, and the solid line is the theoretical prediction of Hertz's law.

The experimental data presented in Fig. 3 show interesting features with regard to the validity of Hertz's theory for xylophones. For low values of the interaction time τ, i.e., for strong impacts with hard mallets, the predicted maximum force is higher than the corresponding measured value. This might be a consequence of the plasticity of the bar, a fraction of the energy communicated by the blow being dissipated into a permanent deformation of the bar, clearly visible on its surface. In the case of weak impacts with soft mallets, especially those with rubber heads, the impact duration becomes comparable to the propagation time of the flexural waves, and thus measurements of τ may be affected by a wave returning from the closest end of the bar. Between these two limiting cases, experiments show good agreement with the contact theory of Hertz.

B. Measurements on wooden bars

The xylophone bar is characterized by its geometry and by its elastic and damping properties. The geometrical parameters of interest here are the length L, the width b, and the section profile $h(x)$ of the bar, respectively. With regard to this latter variable, two different strategies can be used in the simulations: first, if the purpose is to accurately reproduce the vibrations of existing instruments, then the successive values $h(x_i = i\Delta x)$ can be measured on a real bar, with a spatial resolution Δx corresponding to the numerical formulation of the model, and this set of measurements serves as input data for the simulations. This technique is used, for example, to compare measured and simulated waveforms and spectra. Another method consists of using a mathematical formula for $h(x)$—such as a parabolic profile, for example—the two parameters of the parabola then being related to the tuning of the first three partials of the bar, as recently proposed by Orduña-Bustamante.[8]

For a bar with a constant cross section, the Young's modulus E can be derived from measurements of the natural frequencies of the flexural waves (see Ref. 20). This is not the case for a bar with variable cross section, since, in this case, there is no simple analytical relationship between natural frequencies and Young's modulus.[9] For that reason, the Young's modulus has been derived here from measurements of the natural frequencies of the longitudinal waves, after experimental confirmation that the undercut doesn't affect substantially this set of frequencies,[30 21] and assuming further that the effects of lateral inertia can be neglected.[9] This determination of E has been validated by comparison of spectra between measured and simulated bar accelerations (see Sec. IV).

For the measurements of both the elastic and damping properties of the bars, an experimental setup with noncontact excitation and detection is used, so as to reduce as much as possible the causes of errors in the determination of natural frequencies and damping factors. The bar is freely suspended by means of rubber bands, and is set into vibration by means of an acoustic wave delivered by a loudspeaker, located at a distance of nearly 2 m from the bar, in an anechoic room. The acoustic excitation is maintained for several seconds, then suddenly removed. During the free regime of vibration, the transverse velocity of one particular point of the bar is measured by means of a laser vibrometer (Polytec OFV 2600-352). The electric signal delivered by the vibrometer is recorded on a digital audio tape (Sony TCD10-Pro) for further signal processing. In a first step, the acoustic excitation is made of a broadband noise, and a standard FFT analysis is carried out on the response of the vibrometer, in order to obtain a first estimate of the natural frequencies. In a second step, the bar is excited by a sine wave with a frequency corresponding to one peak of the previous FFT magnitude spectrum. The frequency of the sine wave is finely adjusted until the magnitude of the velocity, detected by the laser, reaches its maximum. After suppression of the sine wave excitation, the damping factor at this natural frequency is derived from measurements of the decay time t_d. The decay times are estimated by means of the matrix pencil method.[31] For wooden bars, the damping factors increase with frequency, as can be seen in Fig. 4. Using standard curve-fitting procedures on these experimental data shows that the fre-

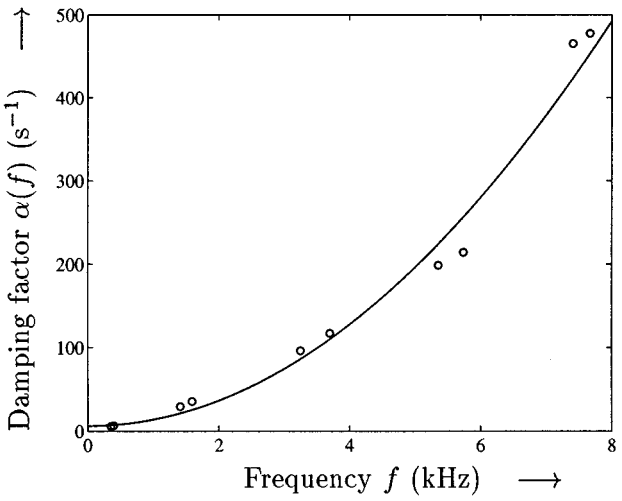


FIG. 4. Frequency dependence of the damping factors for a wooden bar (Padouk). Experimental points (○) and parabolic regression (solid line).

quency dependence of the damping factors can be fairly well described by an empirical law of the form:

$$\alpha(f\ ) = 1/t_d = a_0 + a_2 f^2, \tag{18}$$

where the constants a_0 and a_2 depend on the wood species. Solving Eq. (1) for a damped sinusoid of complex frequency $\omega + j\alpha$ shows that the viscoelastic parameter η and the fluid damping parameter γ_B in the theoretical model are related to the experimental constants by the following formulas:

$$\eta = a_2/2\pi^2 \quad \text{and} \quad \gamma_B = 2a_0 \tag{19}$$

which justifies *a posteriori* the general form of the basic equation (1), and shows that the decay process is mainly determined by internal losses, for wooden bars without resonators. The friction at the cord support is significant for frequencies below the fundamental of the lowest bar only, and can be neglected. The agreement between viscoelastic model and experimental data suggests further that radiation losses are small compared to internal losses. Equation (18) does not account for the influence of a resonator tuned to the fundamental on the time history of the radiated sound.

IV. RESULTS OF SIMULATIONS

One can identify three main causes of errors in the numerical simulations of mechanical structures in general, and of musical instruments in particular. First, a number of differences between measurements and simulations can be attributed to the limitations of the physical model itself. However, using a simple model deliberately may give significant insight into the main physical features of the instrument, and a high degree of complexity may not always be justified. The quality of a synthetic musical tone can be improved through successive steps of refinement, but it is clear that each new step will complicate the mathematical description, and thus increase the computing time. Therefore, a compromise has to be found between sound quality and computational burden. The prime objective here is to allow a clear recognition of the xylophone, and of its main timbral capabilities. Another interest of the model lies in the relatively limited number of parameters for the control of a wide variety of tones, each control parameter being closely related to the normal use and design of a real instrument. Our program is controlled by twelve different parameters, which are related to the geometry and material properties of the bar, to the inertial and stiffness properties of the mallet, and to the player's action, respectively.

A second class of errors is due to the numerical approximations of the equations. It has been shown in Sec. II that frequency warping is the prime drawback of an insufficiently small spatial sampling interval. The estimation of frequencies is highly relevant in the context of musical signals, since the ear is very sensitive to the frequency spectrum which conveys, among other things, the sense of pitch, and thus the degree of accuracy required for the numerical algorithm depends for the most part of its ability to control adequately the tuning of the simulated instrument. In practice, the spatial resolution of the bar has been selected so that the degree of accuracy for the three main partials is less than 0.5%, which roughly corresponds to the sensitivity of the human ear with regard to frequency.

Finally, significant differences between measured and simulated quantities may be due to insufficient precision in the experiments carried out for extracting the parameters, but may also indicate that the model is not accurate enough. For the modeling of short bars, for example, neglecting shear and rotary inertia may be a too drastic assumption, which may explain some discrepancies between real and simulated tones.

In order to discriminate between these various sources of errors, four specific points are addressed in the present section. In the first paragraph, the numerical results are evaluated through comparisons between simulated and analytical solutions. These first comparisons can be achieved only for a very simplified version of the model, and must be extended by tests of convergence for the complete model. In these procedures, presented in paragraph B, the spatial step decreases progressively, so as to examine whether the results tend asymptotically to a unique solution. The efficiency of the program, with regard to the variety of physical parameters used for controlling the timbre of the simulated tones, is discussed in paragraph C. Finally, comparisons between real and simulated forces, and between real and simulated bar accelerations, are presented in paragraph D.

A. Comparison between numerical and analytical solutions

Due to the complexity of the model described in Sec. I for the mallet–bar system, no analytical solution can be found in the general case. However, the basic model can be significantly simplified by removing from Eq. (1) the dissipative terms, the coupling with the cord, and the force density term, respectively. In this case, it has been shown by Graff that, given a Gaussian distribution of displacement at the origin of time (see Fig. 5) of the form

$$w(x,0) = w_0 \exp\left(-\frac{x^2}{4d^2}\right), \quad \frac{\partial w(x,0)}{\partial t} = 0, \tag{20}$$

then the following analytical solution for the problem of the infinite bar with constant cross section can be found:[9]

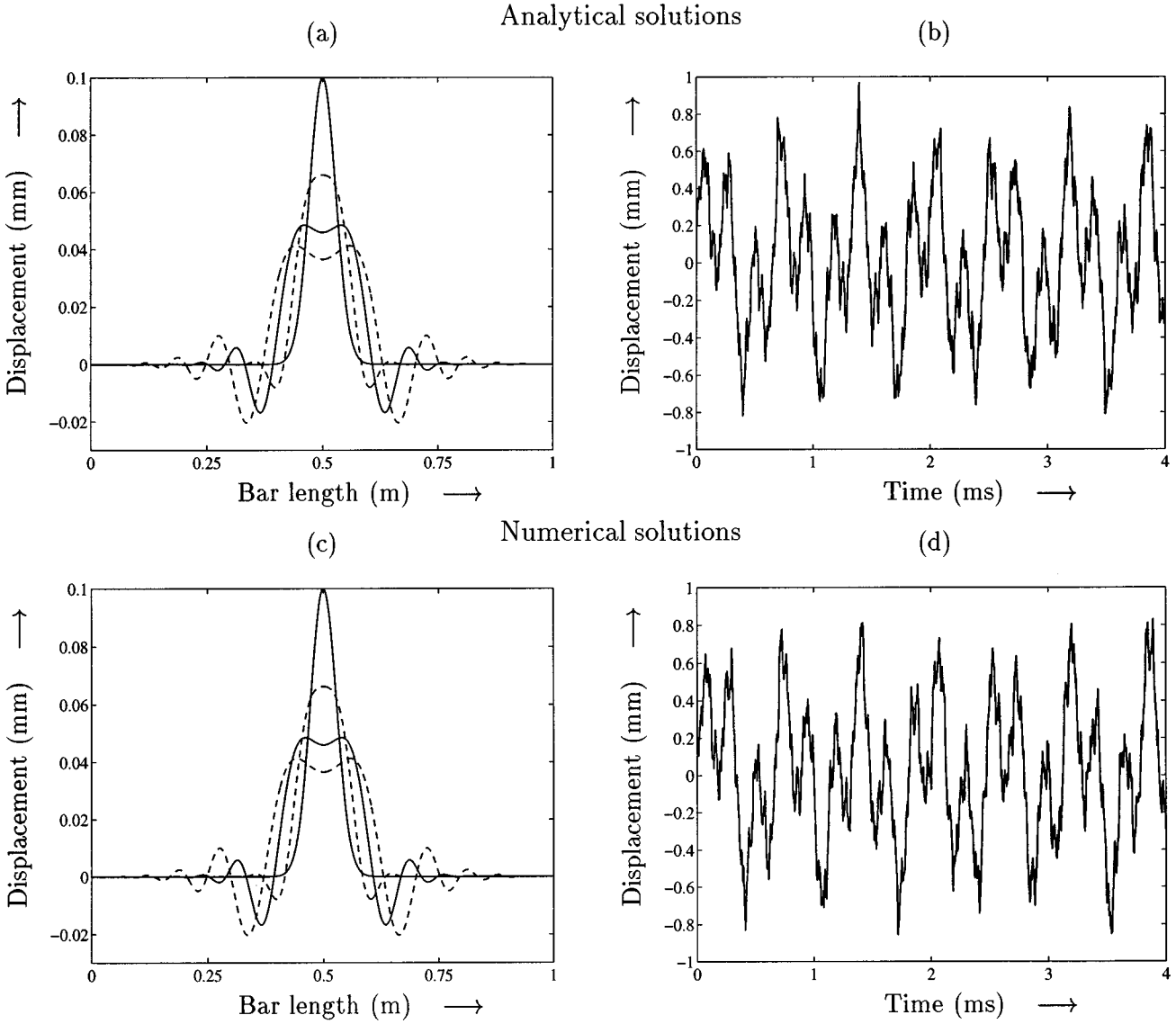


FIG. 5. Comparison between analytical (top) and numerical results (bottom). Displacement of a bar of constant section, struck at its mid-point, at four successive instants ($t=0$, $50\Delta t$, $100\Delta t$, and $150\Delta t$) before reflection at the ends (left). Time history of the displacement at $x=L/3$ for a bar of finite length excited by a Dirac delta function (right).

$$w(x,t)=\frac{w_0}{(1+a^2t^2/d^4)^{1/4}}\exp\left\{-\frac{x^2d^2}{4(d^4+a^2t^2)}\right\}\times\cos\left\{\frac{atx^2}{4(d^4+a^2t^2)}-\frac{1}{2}\arctan\left(\frac{at}{d^2}\right)\right\},\qquad(21)$$

where a is defined in Eq. (9). This analytical solution serves here as reference in order to validate the numerical formulation of the problem. The degree of discrepancy between analytical and numerical solution is evaluated by computing the following error criterion:

$$\epsilon_n=\frac{\Sigma_i(s_1-\hat{s}_i)^2}{\Sigma_i s_i^2},\qquad(22)$$

where s_i is the value of the analytical signal at position $x_i=i\Delta x$, and $\hat{s}_i$ is the corresponding value of the numerical signal, at a given time $t_n=n\Delta t$.

With $\Delta x=2$ mm and $f_s=768$ kHz, the error criterion ϵ_n defined in (22) is equal to 0.2% at $t_n=40\Delta t$, and equal to 0.3% at $t_n=80\Delta t$.

A similar criterion can be defined for evaluating the degree of error for the simulated displacement at a given position for a finite bar (see Fig. 5). In this case, the analytical expression of the displacement is obtained through a normal mode expansion over the first 50 modes of the free–free bar.

TABLE II. Influence of the spatial resolution on the tuning of a simulated G_4 bar.

Δx (mm)	N	f_1 (Hz)	f_2 (Hz)	f_3 (Hz)	f_4 (Hz)	f_5 (Hz)	f_6 (Hz)
20.9	14	393	1554	3386	5151	7150	9564
10.1	29	395	1588	3542	5541	7985	11067
5.05	58	397	1600	3584	5648	8196	11445
2.02	145	397	1604	3597	5680	8259	11560
1.01	290	397	1604	3599	5684	8283	11573
0.53	552	397	1604	3598	5683	8280	11570

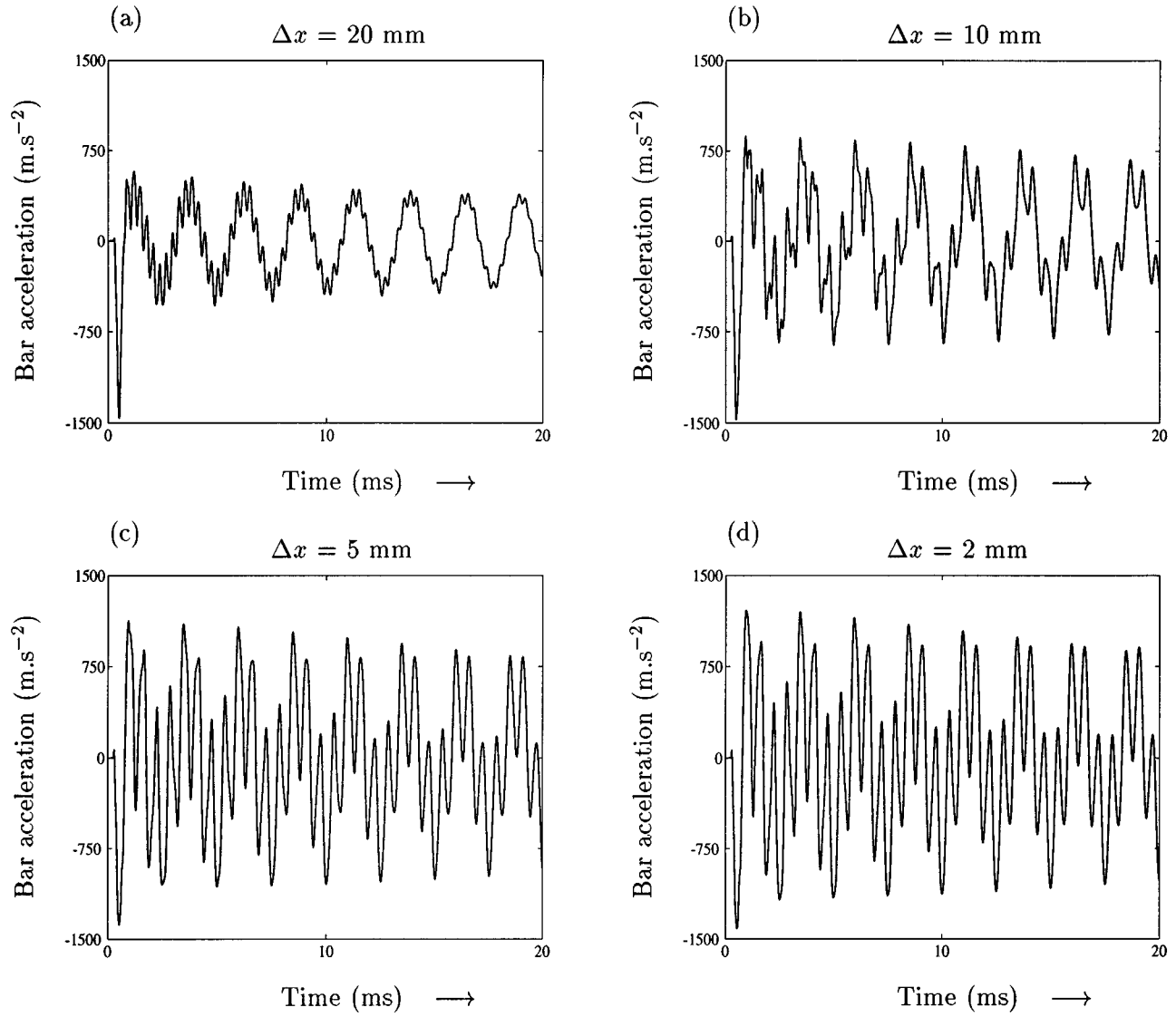


FIG. 6. Tests of convergence (I). Influence of the spatial resolution on the simulated bar acceleration (G_4: 397 Hz), at $x=L/3$, for large steps: (a) $\Delta x=20$ mm, (b) $\Delta x=10$ mm, (c) $\Delta x=5$ mm, (d) $\Delta x=2$ mm.

In this case, the impact force is a Dirac delta function, and thus the displacement can be viewed as the impulse response of the finite bar. At $x_0=L/2$, for example, the error criterion, with summation over the 80 first discrete instants to time, is equal to 0.4%.

B. Tests of convergence

The previous comparisons were restricted to a simplified version of the model. For the complete model described in Sec. I, another meaningful test consists of evaluating the convergence of the numerical scheme. In practice, the convergence of the scheme is assessed by increasing the number N of spatial points progressively, and examining the subsequent variations of the results in both the time and frequency domain.

Table II shows the frequencies of the first six partials of a simulated G_4 bar, for six different values of Δx, decreasing from 20.9 to 0.53 mm. This corresponds, for this bar, to a

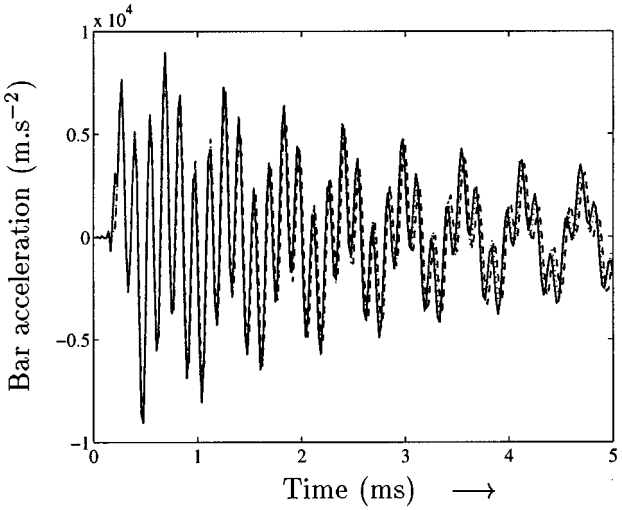


FIG. 7. Tests of convergence (II). Superimposed simulations of the bar acceleration (A_6: 1760 Hz), at $x=L/3$, for small steps: $\Delta x=5$ mm (dashed), $\Delta x=3$ mm (dotted), $\Delta x=5$ mm.

TABLE III. Natural frequencies and frequency ratios obtained for a simulated bar with various parabolic undercuts.

X_c	T_c	f_1 (Hz)	f_2 (Hz)	f_3 (Hz)	R_{21}	R_{31}
0.2687	0.8394	672	2012	4018	2.99	5.98
0.1564	0.5225	483	1927	3850	3.99	7.97
0.2378	0.4728	397	1584	3562	3.99	8.97
0.1241	0.3808	382	1905	3806	4.99	9.96
0.1564	0.3689	347	1731	3802	4.99	10.95
0.1869	0.3457	308	1540	3701	5.00	12.01

number N increasing from 14 to 552. It can be seen that this series of frequencies is convergent, and that the discrepancy is less than 0.3%, for all partials, for $\Delta x \leqslant 2$ mm.

The influence of the spatial resolution can also be seen on the waveforms shown in Fig. 6 and Fig. 7. For large values of Δx, the simulated acceleration is very distorted, when compared to the measured signal. Ripples and parasitic oscillations are clearly visible in Fig. 6, and are the consequences of the crude spatial approximation of the bar. With Δx smaller than or equal to 5 mm, the superimposed waveforms in Fig. 7 illustrate the slight deviations of the signals, which are mainly attributable to the dispersive properties of the numerical scheme.[27]

C. Systematic exploration of some mallet–bar parameters

The efficiency of the simulation program is further assessed by making systematic explorations of physical parameters, in connection with the usual manufacturing and normal use of the xylophone. It is observed, first, that variations of density and Young's modulus give rise to the expected changes in natural frequencies of the vibrations. Next, the simulations confirm that the damping model yields the characteristic timbre of wooden bars during the decay part of the sound and the expected decay times. In addition, the nonlinearity of the excitation is clearly audible, and visible in the

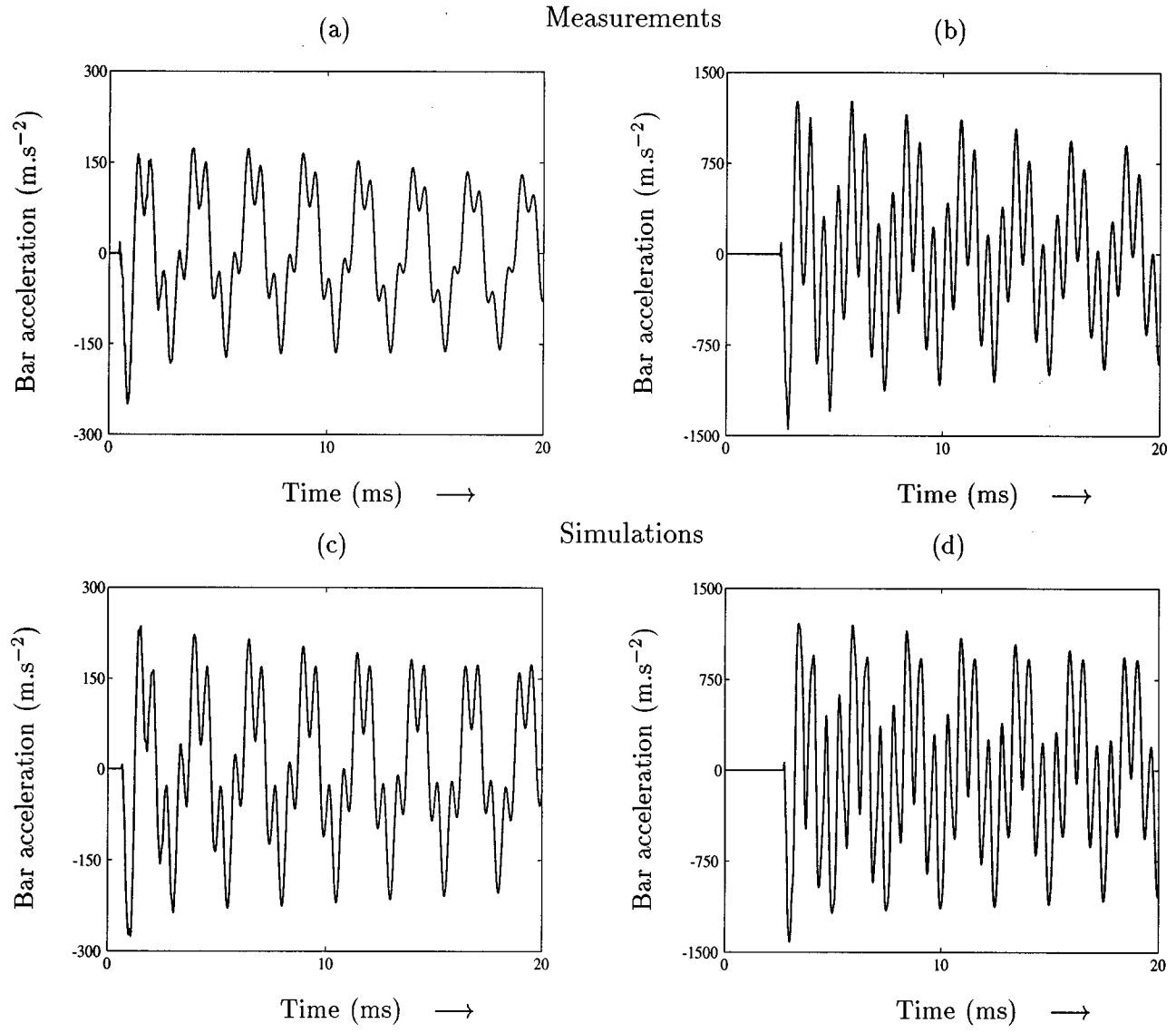


FIG. 8. Comparison between measured (top) and simulated (bottom) bar accelerations for a soft mallet with a rubber head. Weak impact *piano* (left), and strong impact *mezzo-forte* (right). Position of excitation $x_0 = 15.9$ cm, accelerometer position $x = 9.3$ cm.

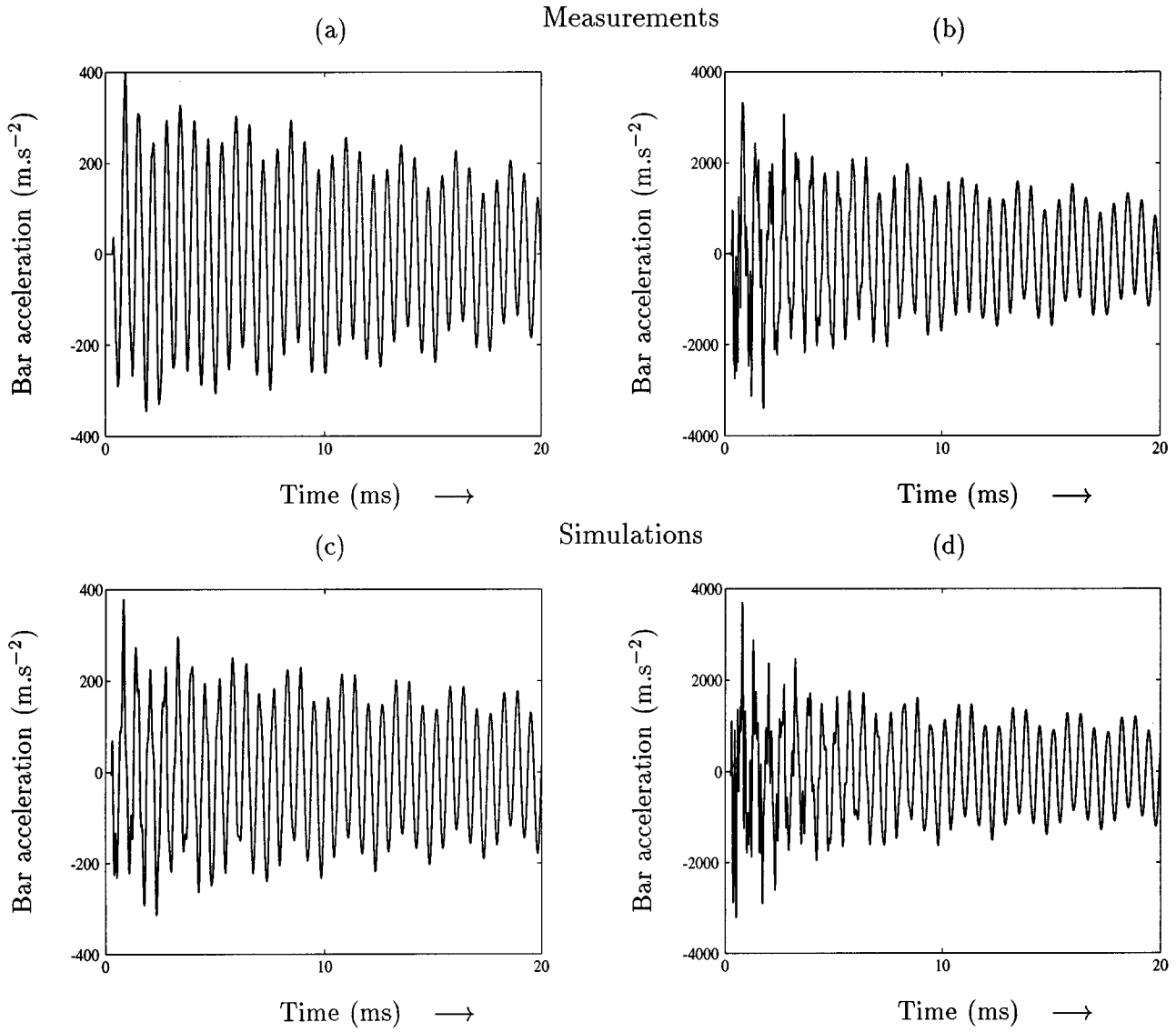


FIG. 9. Comparison between measured (top) and simulated (bottom) bar accelerations for a hard mallet with a boxwood head. Weak impact *piano* (left), and strong impact *mezzo-forte* (right). Position of excitation $x_0=18.2$ cm, accelerometer position $x=9.3$ cm.

spectra. The width τ of the impact force pulse decreases as the initial velocity of the mallet increases (see Appendix A). Variations of the impact position also produce the expected periodicities in the spectral envelope. These numerical experiments show that the model is able to reproduce and control the main characteristics of both the bar and player's action.

Another significant feature of the program is to allow the control of the mallet's stiffness by means of the coefficient K. It has been verified that the ear is very sensitive to variations of this particular parameter, which, together with V_0, determines the width of the initial pulse in the waveform and the number of excited modes. Auditory demonstrations made with series of simulated tones show that the timbral quality of the initial transient is notably influenced by modifications of the stiffness constant, and that some trained listeners are able to recognize the material of the mallet's head.

Finally, the numerical model has the interesting capability of simulating any section profile. One can take advantage of this property not only for modeling existing instruments, but also for testing new shapes for the cross section. To illustrate this point, Table III shows the relationship between various parabolic profiles used for the simulated undercut and the values obtained for the natural frequencies of the simulated acceleration. In this table, the dimensionless parameter $X_c=x_c/L$ is related to the width x_c of the undercut, and the dimensionless parameter $T_c=1-h_c/h_{\max}$ is the ratio between the thickness $h_{\max}-h_c$ of the bar at the center of the undercut, and the thickness $h_{\max}$ at both ends, using the same definitions as Orduña-Bustamante in Ref. 8. Similarly, R_{21} and R_{31} are the frequency ratios between the second partial and the fundamental, and between the third partial and the fundamental, respectively. It can be seen that the various tuning obtained for the simulated profiles are in excellent agreement with those predicted by this author. The simulations here were computed with $f_s=192$ kHz, and $\Delta x=2$ mm, using the elastic parameters of the G_4 bar (see Table I).

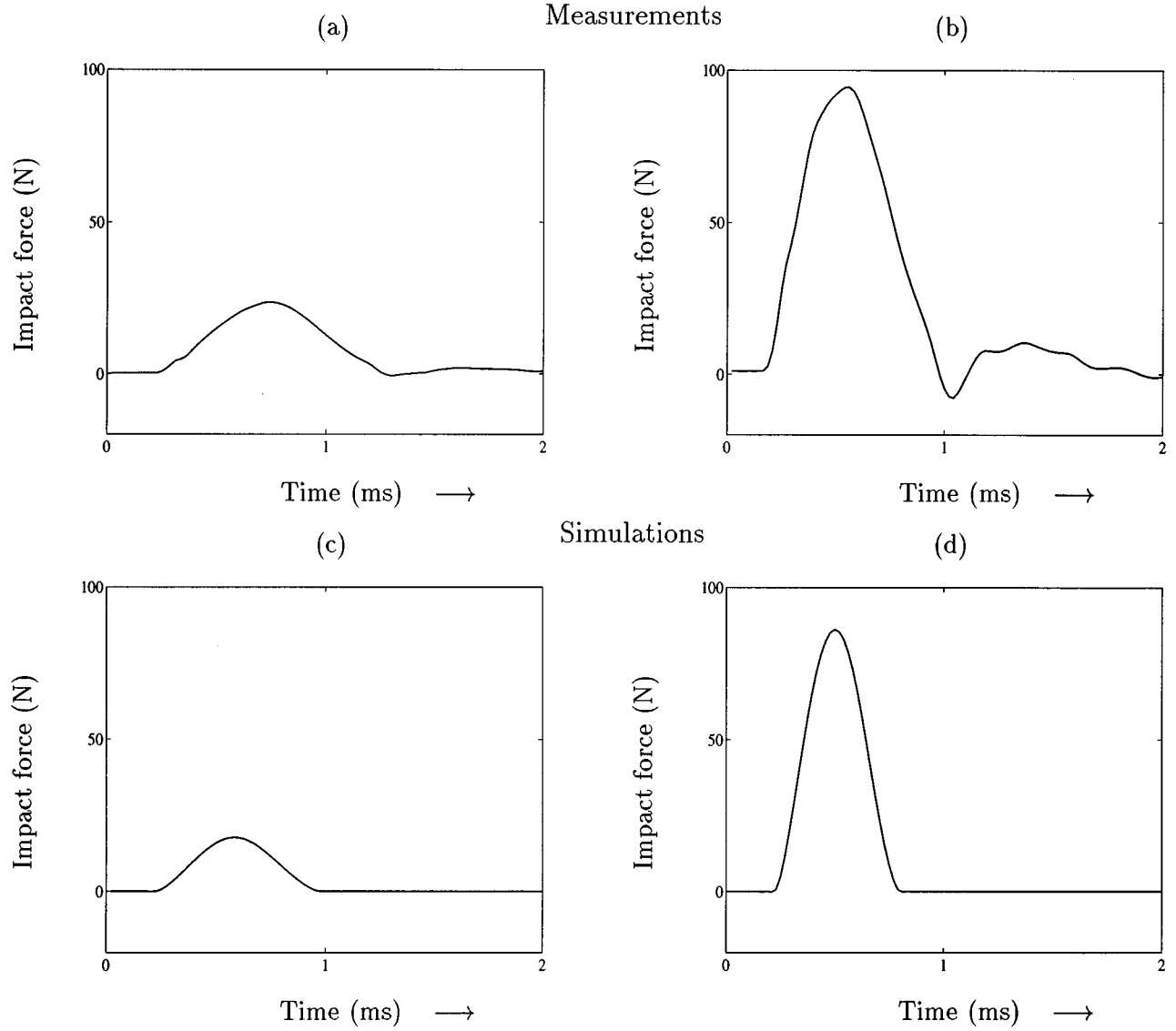


FIG. 10. Comparison between measured (top) and simulated (bottom) impact forces for a soft mallet with a rubber head. Weak impact *piano* (left), and strong impact *mezzo-forte* (right).

D. Comparison between measured and simulated force and acceleration waveforms

The last part of the experiment consists of comparing the results of simulations with force and acceleration waveforms recorded on a real instrument. As it has been described above in Sec. III, the force signal is derived from measurements of the mallet's head acceleration after multiplication by the equivalent mass defined in Sec. III. The acceleration of one particular selected point of the bar is either measured with the help of an accelerometer or derived from the velocity signal delivered by a laser vibrometer. This second method is preferred for small bars, for which the vibrations are affected by the presence of the accelerometer. The advantage for displaying acceleration waveforms and spectra, compared to displacement or velocity, is that the upper partials are more clearly visible.

Figure 8 shows a comparison between measured and simulated accelerations of a given point on a bar struck by a soft mallet with a rubber head, for two different impact velocities. For a weak impact (V_0=0.26 m s^{-1}), it can be observed that the amplitude and the general shape of the waveforms are similar, which is confirmed by spectral analysis (see Fig. 12). However, the upper partials seem to be damped more rapidly in the measured acceleration, which may be due to the fact that the contact is dissipative, a physical effect which has not been taken into account in the model. For a stronger impact (V_0=1.0 m s^{-1}), the measured and simulated waveforms are very similar in magnitude and shape, which indicates that the model is valid in this range. This result confirms the assumptions made in Sec. III on the validity domain of Hertz's law.

Similarly, Fig. 9 shows acceleration waveforms for the same bar struck by a hard mallet with a boxwood head. The comparisons between measured and simulated signals are made for a very weak impact (V_0=0.07 m s^{-1}) and for a relatively stronger impact (V_0=0.474 m s^{-1}), respectively. In both cases, one can see a high degree of similarity in magnitude and shape between measured and simulated wave-

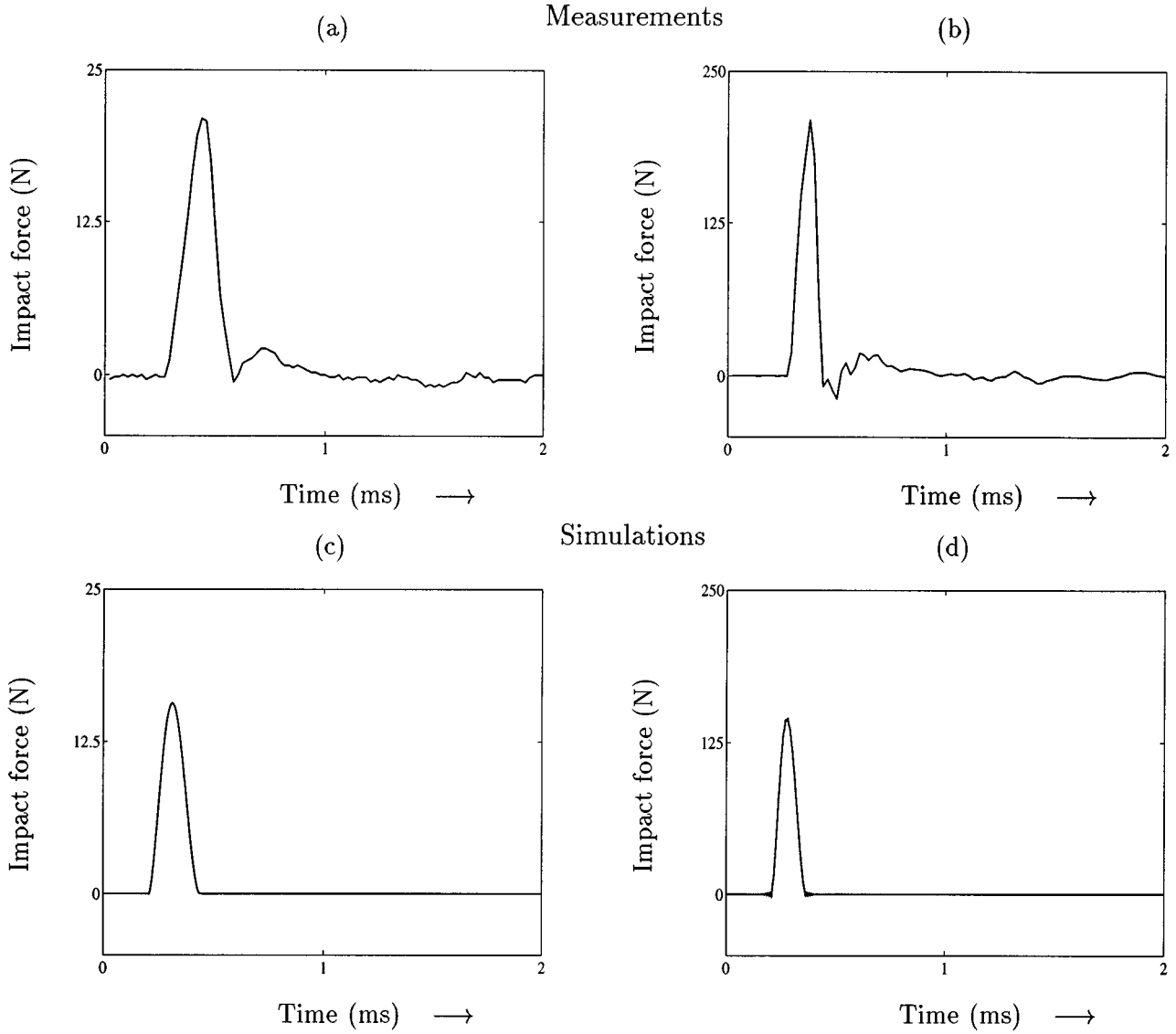


FIG. 11. Comparison between measured (top) and simulated (bottom) impact forces for a hard mallet with a boxwood head. Weak impact *piano* (left), and strong impact *mezzo-forte* (right).

forms, and this visual appearance is confirmed by spectral analysis (see Fig. 13) and listening tests. For the strong impact, the maximum acceleration is 3400 m s^{-2} in the measured signal, and 3217 m s^{-2} in the simulation, whereas for the weak impact, the maximum accelerations are 400 and 378 m s^{-2}, respectively.

The time history of the impact forces corresponding to the previous accelerations can be seen in Fig. 10 for the rubber maller (soft), and in Fig. 11 for the boxwood mallet (hard). Figure 10 shows that the order of magnitude for both amplitude and shapes are fairly reproduced, but also that the simulated impact durations are systematically shorter (about -20%) than the measured ones. This feature illustrates further to what extent the behavior of rubber during the contact with the bar is governed by a more complicated law than the Hertz's model. For the strong impact, the shape of the measured force is slightly distorted, which might be a consequence of the relatively large deformation of the head. The Hertz's theory fails in accounting for this effect, since it is based on the assumption of infinitesimal strain.

In the case of an impact with the hard (boxwood) mallet, the impact durations are identical for both measured and simulated forces. The duration is equal to 250 μs for the weak impact, and 160 μs for the strong impact (see Fig. 11). Here, the discrepancies primarily affect the maximum of the force, which is nearly 30% higher in the measurements than in the simulations. These discrepancies can be due to the lack of accuracy in the experimental determination of K, since the simulations presented in this figure were conducted with a mean value of $K=1.31\times10^9$ N $m^{-3/2}$, the standard deviation being roughly equal to 34% of this value. By readjusting this parameter, the magnitudes of both measured and simulated force become comparable.

Finally, Figs. 12 and 13 show comparisons in the frequency domain between measured and simulated spectra, for identical excitation and observation points. These spectra were obtained by means of an FFT analysis on the first 20 ms of the acceleration signals presented in Figs. 8 and 9,

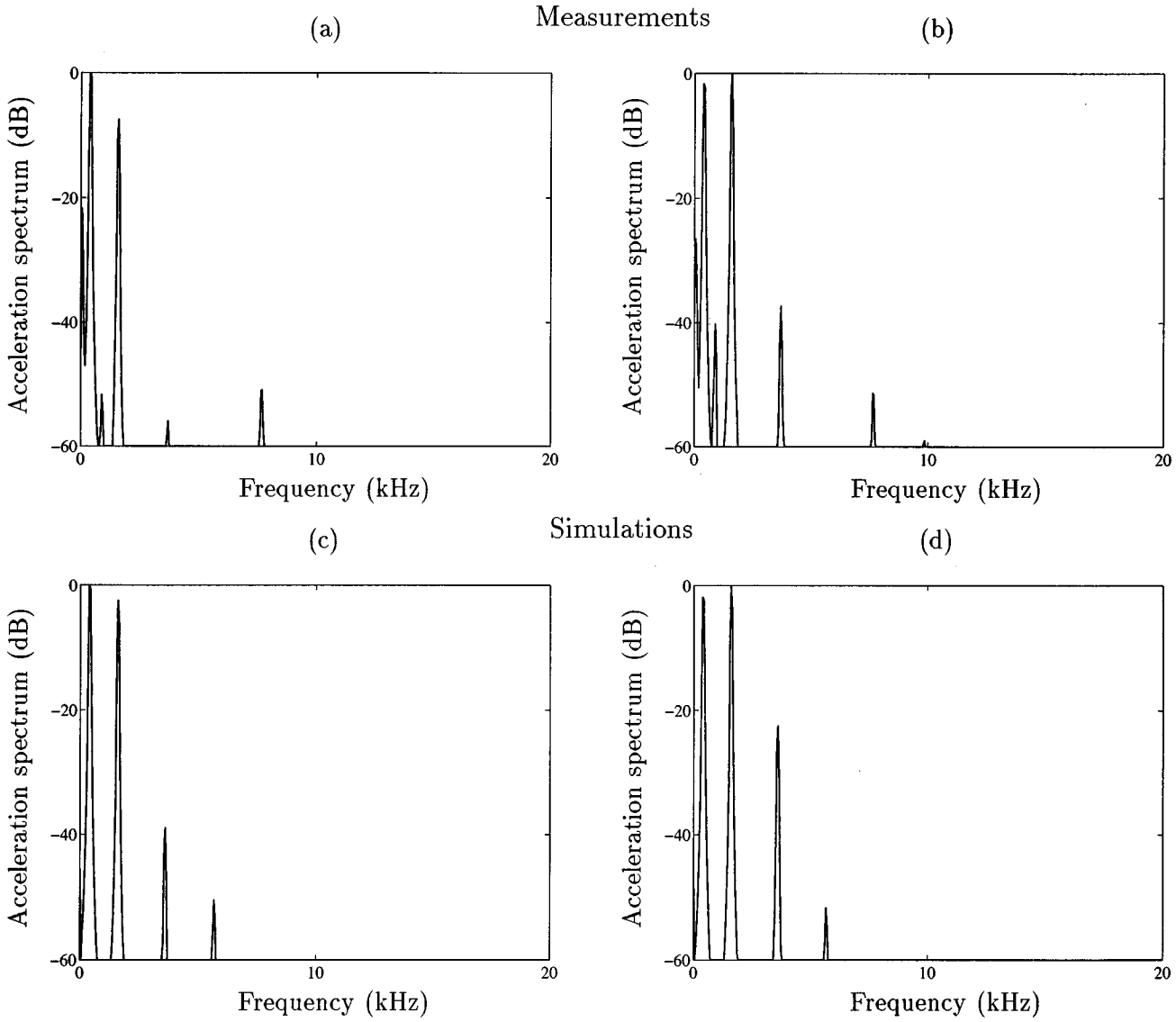


FIG. 12. Comparison between measured (top) and simulated (bottom) bar acceleration spectra for a soft mallet with a rubber head (spectral analysis of waveforms presented in Fig. 8). Weak impact *piano* (left), and strong impact *mezzo-forte* (right).

respectively. In Fig. 12 (rubber head), the simulated magnitude of the third partial is about 15 dB below the magnitude of the corresponding partial in the measured signal. This may be a consequence of the above-mentioned differences in impact force durations, and/or of the fact that the damping of the mallet's head itself is not taken into account in the model. In Fig. 13 (boxwood head), it can be seen that the magnitudes of the first three partials are reproduced almost perfectly, with discrepancies less than or equal to 2 dB. The values of the frequencies, and thus the perception of pitch, are nearly identical. The discrepancies become visible for the frequencies above 6 kHz. In this domain, the assumptions of the Euler–Bernoulli model are no more valid, and it is necessary to include shear and rotary inertia in the model in order to account for the variation of flexural wave velocity with frequency[9] more closely. In addition, the measured spectra frequently exhibit some peaks which don't correspond to the flexural wave series. Careful examination show that these additional peaks can be mostly attributed to torsional waves, which are excited by striking the bar off its symmetry axis Ox. Here again, it is plain that the model is not yet able to account for these waves.

V. CONCLUSION

The present model of vibration for a thin xylophone bar excited by a mallet yields simulated tones of high quality, and reproduces adequately the main physical attributes of a real instrument. The use of Hertz's model for the excitation of the bar yields good agreement between real and simulated sounds, except for very strong impacts. In this case, the permanent deformation of the bar, and/or nonlinear stress–strain relation of the mallet's head, are probably the main causes of deviation from Hertz's law.

The necessity of defining the geometry of the bar with sufficient accuracy for the tuning and for the modeling of the impact surface, imposes the use of implicit, rather than explicit, finite difference schemes. However, the numerical for-

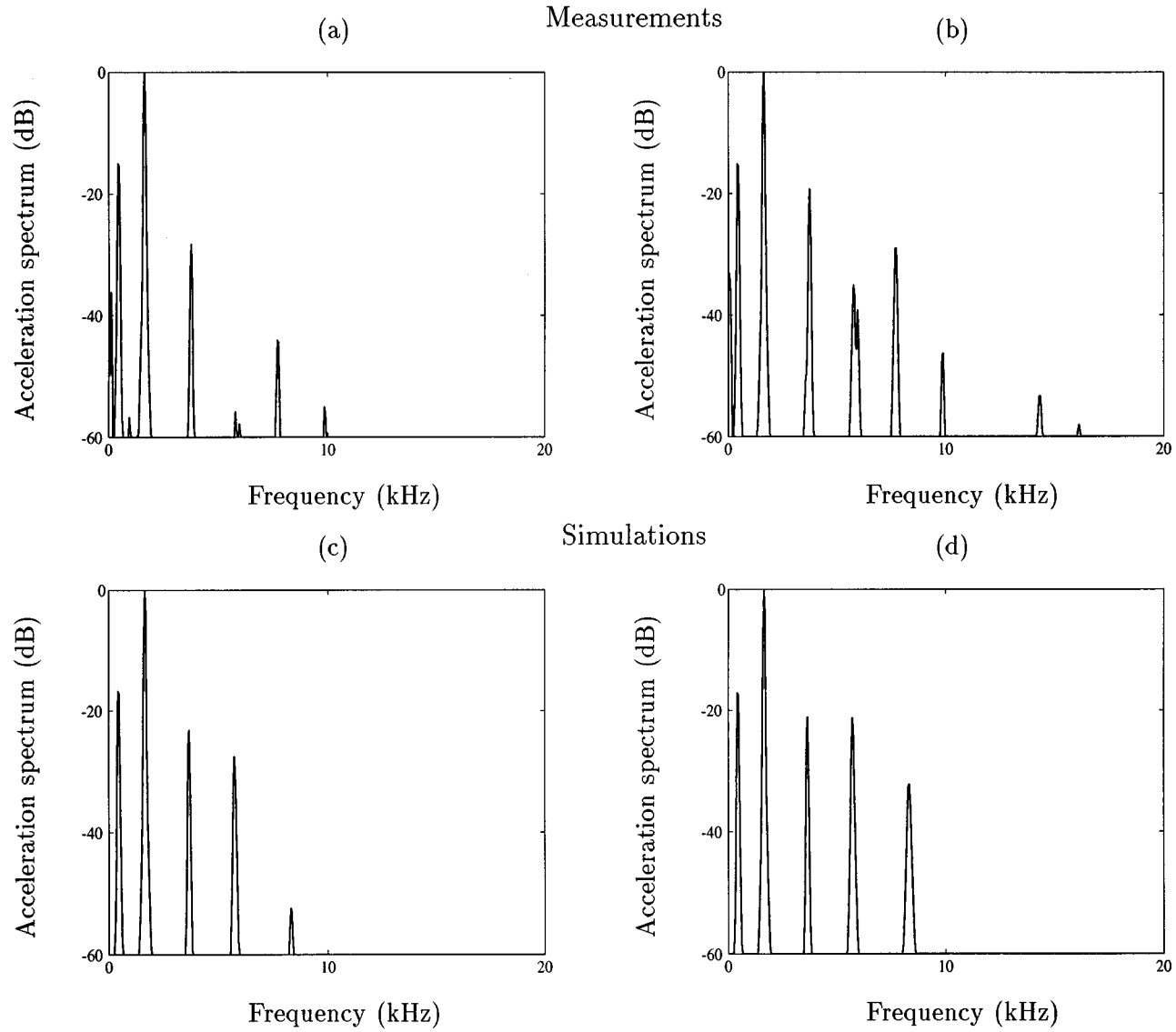


FIG. 13. Comparison between measured (top) and simulated (bottom) bar acceleration spectra for a hard mallet with a boxwood head (spectral analysis of waveforms presented in Fig. 9). Weak impact *piano* (left), and strong impact *mezzo-forte* (right).

mulation of the model remains relatively simple in form. The average computing time on a standard workstation is 10^3 for 1 s of sound, at 192 kHz.

A practical application of the model, in the context of instrument making, is the simulation of various shapes of the undercut for obtaining given frequency ratios of overtones. One illustration of this feature has been given in Table III for a family of parabolic undercuts. Other families of profiles are currently under examination, which should give more flexibility for the tuning of the notes in the high register. Another work in progress consists of testing systematically the human sensitivity with regard to the physical parameters of the vibrating source. For this purpose, series of simulated tones are equalized in pitch and loudness, and only one parameter of timbre is modified at a time. Results of these psychoacoustical tests will be reported in the next future. The question of preferred tuning, similar to the one investigated in the past by Bork and Meyer,[32] has not been addressed in our study.

Visual and auditory comparisons between real and simulated bar accelerations show a very high degree of similarity. The model is able to control independently the most significant parameters of bar, mallet, and player's action, respectively, for producing a remarkable variety of tones. As in previous simulations devoted to piano strings, the use of a time-domain modeling allows the possibility of simulating repeated notes on the same bar and musical pieces.[16] The next step of the work will now consist of the modeling of the sound-pressure radiated by the xylophone bar. This model will also include the coupling of the bar with a one-dimensional tubular resonator.

Although the presentation has been restricted to the case of xylophones, the model can be adapted to other mallet percussion instruments, such as marimba, glockenspiel, and vibraphone. This extension implies a generalization of the viscoelastic model presented in Sec. I in order to reproduce the complicated frequency dependence of damping for materials such as aluminum, steel, glass, and carbon fibers. Such a presentation is beyond the scope of the present paper.

The major limitation of the physical description for the flexural waves of the bar follows from the fact that both shear and rotary inertia are neglected. These assumptions become inadequate in the upper range of the instruments, where the width and the thickness of the bar become comparable to the spatial wavelength. As a consequence, the frequencies of the partials are overestimated with the Euler–Bernoulli model, when compared to actual xylophone bars. For this reason, a Timoshenko model, which takes shear and rotary inertia into account, is now in progress for the time-domain modeling of a complete instrument.

ACKNOWLEDGMENTS

Part of this work has been supported by a grant from French Ministry of Culture and Francophony. The authors wish to thank Dr. T. Rossing for helpful comments on the manuscript.

APPENDIX A: HERTZ'S LAW APPLIED TO BAR–MALLET INTERACTION

Most of the theoretical results presented below can be found in Ref. 23. This Appendix introduces the main formulas used for the modeling of the bar–mallet interaction in xylophones.

During the time of contact, the compression δ_M of the mallet's head of radius R_M, Young's modulus E_M, and Poisson's ratio ν_M striking a bar with radius of curvature R_B, Young's modulus E_B, and Poisson's ratio ν_B, is given by

$$\delta_M = \left[F^2 D^2 \left(\frac{1}{R_B} + \frac{1}{R_M} \right) \right]^{1/3} \tag{A1}$$

with

$$D = \frac{3}{4} \left(\frac{1-\nu_B^2}{E_B} + \frac{1-\nu_M^2}{E_M} \right), \tag{A2}$$

where F is the interaction force between the two solids in contact. In our case, it is assumed that $R_B \to \infty$, so that (A1) becomes

$$F = K \delta_M^{3/2} \quad \text{with} \quad K = \sqrt{R_M}/D. \tag{A3}$$

With $R_M = 1$ cm, $E_B = E_M = 2.0 \times 10^{10}$ N m^{-2}, $\nu_B = \nu_M = 0.1$, Eq. (A3) yields $K = 1.35 \times 10^9$ N m$^{-3/2}$, which is comparable to the experimentally obtained value for the boxwood head ($K = 1.31 \times 10^9$ N m$^{-3/2}$).

The conservation of mechanical energy, written in the referential where the center of gravity of both solids is at rest, imposes

$$\mu \left(\frac{d\delta_M}{dt} \right)^2 + \frac{4}{5} K \delta_M^{5/2} = \mu V_0^2, \tag{A4}$$

where μ is the reduced mass of bar and mallet, and V_0 the velocity of the mallet's head at the beginning of the impact. From (A4), one can derive the maximum value $\delta_{M\,\max}$ of the compression:

$$\delta M_{\max} = \left(\frac{5}{4} \frac{\mu}{K} \right)^{2/5} V_0^{4/5} \tag{A5}$$

and the interaction time:

$$\tau = 2 \int_0^{\delta_{M\max}} \frac{d\delta_M}{\sqrt{V_0^2 - \frac{4}{5}(K\delta_M^{5/2}/\mu)}} = 2 \left(\frac{25\mu^2}{16K^2 V_0} \right)^{1/5} = \int_0^1 \frac{d\xi}{\sqrt{1-\xi^{5/2}}} \tag{A6}$$

using properties of the Gamma function:

$$\int_0^1 \frac{d\xi}{\sqrt{1-\xi^{5/2}}} = \frac{2}{5} \frac{\Gamma(2/5)\Gamma(1/2)}{\Gamma(2/5+1/2)} = \frac{2}{5} \sqrt{\pi} \frac{\Gamma(2/5)}{\Gamma(9/10)} = 1.4716,$$

which yields

$$\tau = 3.2181 \left(\frac{\mu^2}{K^2 V_0} \right)^{1/5}. \tag{A7}$$

Equation (A7) shows that the duration of impact is weakly decreasing with the impact velocity. This result is in agreement with the experimental values obtained by Bork for a large variety of mallets.[7]

Inverting (A7) allows an estimation of the stiffness constant K from τ and V_0:

$$K = 18.5775 \frac{\mu}{\tau^{5/2}\sqrt{V_0}}. \tag{A8}$$

Similarly, one can derive from (A3) and (A5) the expression of K as a function of the maximum force $F_{\max}$ and V_0:

$$K = F_{\max}^{5/2} \left(\frac{4}{5 V_0^2 \mu} \right)^{3/2} = 0.7155 \frac{F_{\max}^{5/2}}{V_0^3 \mu^{3/2}}. \tag{A9}$$

Since it is not straightforward to measure V_0 directly, another interesting alternative consists of deriving K from measurements of $F_{\max}$ and τ, using (A8) and (A9), which yields

$$K = 35.4 \frac{1}{\tau^3} \sqrt{\frac{\mu^3}{F_{\max}}}. \tag{A10}$$

The maximum value of the compression is calculated by means of (A5), using measured values of K, V_0, and μ for the boxwood and the rubber mallet's head, respectively. Ordinary values of $\delta_{M\,\max}$ usually lie within the range $[2.0 \times 10^{-5}; 3.0 \times 10^{-4}]$ m. Finally, the radius of contact area is given by[23]

$$a_{M\,\max} = \left(F_{\max} D \frac{R_B R_M}{R_B + R_M} \right)^{1/3}, \tag{A11}$$

which gives, in our application,

$$a_{M\,\max} = \sqrt{\delta_{M\,\max} R_M}. \tag{A12}$$

With the previous numerical values, in the cases of boxwood and rubber mallet heads, Eq. (A12) gives values of the contact radius between 0.5 and 2.5 mm. The impact width δx in the model is taken equal to $2a_{M\,\max}$.

Replacing $\delta_{M\,\max}$ by $\delta(t)$ in Eq. (A12) shows that the contact area varies with time.

APPENDIX B: FINITE DIFFERENCE APPROXIMATIONS

1. Explicit scheme

Using linear combinations of Taylor series allows us to approximate the partial derivatives of first and second order with respect to time in (1) and (8) by the following discrete operators:

$$(\delta_t w)_i^n = \frac{w_i^{n+1} - w_i^{n-1}}{2\Delta t} + O(\Delta t^2)$$

and (B1)

$$(\delta_t^2 w)_i^n = \frac{w_i^{n+1} - 2w_i^n + w_i^{n-1}}{\Delta t^2} + O(\Delta t^2).$$

Similarly, one can approximate the partial derivative of fourth order with respect to spatial coordinates in (1) by the operator

$$(\delta_x^4 w)_i^n = \frac{w_{i+2}^n - 4w_{i+1}^n + 6w_i^n - 4w_{i-1}^n + w_{i-2}^n}{\Delta x^4} + O(\Delta x^2). \quad \text{(B2)}$$

In order to keep the explicit formulation of the recurrence equation (10), one has to approximate the derivative $\partial^5 w/\partial x^4 \partial t$ by the operator

$$\frac{(\delta_x^4 w)_i^n - (\delta_x^4 w)_i^{n-1}}{\Delta t} + O(\Delta t). \quad \text{(B3)}$$

The approximation (B3) is only of the first order with respect to Δt. However, this derivative is multiplied by a parameter η of relatively small magnitude compared to the other terms of (8), and the accuracy of the numerical scheme (10) is not affected by this low-order approximation. The recurrence equation for a bar with constant section (10) is obtained by means of the approximations (B1) to (B3).

For this explicit scheme, the stability condition written in Eq. (12) is obtained by the Fourier method,[26] on the simplified equation

$$\frac{\partial^2 w(x,t)}{\partial t^2} = -a^2 \left[\frac{\partial^4 w(x,t)}{\partial x^4} + \eta \frac{\partial^5 w(x,t)}{\partial t \partial x^4} \right] \quad \text{(B4)}$$

considering that the other terms do not affect the stability. This assumption is exact for the fluid damping term and for the force density, provided that this latter term is bounded in time and space, as it is the case here. The effect of the restoring force of the cord on stability is of second order with respect to the sampling rate, and will be neglected. The Fourier method consists of examining the propagation of a solution of the form

$$w_i^n = W_n \exp(jik\Delta x) \quad \text{with} \quad j = \sqrt{-1}. \quad \text{(B5)}$$

The difference equation corresponding to Eq. (B4) then becomes

$$W_{n+1} - 2W_n \left[1 - 8r^2(1 + \eta f_s) \sin^4\left(\frac{k\Delta x}{2} \right) \right] + W_{n-1} \left[1 - 16r^2 \eta f_s \sin^4\left(\frac{k\Delta x}{2} \right) \right] = 0. \quad \text{(B6)}$$

The stability criterion is obtained by writing the condition under which W_n does not increase with time. This is obtained if the discriminant of the characteristic equation corresponding to Eq. (B6) is negative or equal to zero, which yields

$$r^2 \leqslant \frac{1}{4(1 + \eta f_s)\sin^4(k\Delta x/2)}. \quad \text{(B7)}$$

Finally the term $\sin^4(k\Delta x/2)$ is replaced by unity in Eq. (B7), considering that this condition must be fulfilled for all k. The condition expressed in Eq. (12) is then obtained after straightforward algebraic transformations.

2. Implicit scheme

For a bar with a variable cross section, the partial derivative of second order with respect to spatial coordinates $\partial^2 w/\partial x^2$ in (1) is approximated by the discrete operator

$$(\delta_x^2 w)_i^n = \frac{w_{i+1}^n - 2w_i^n + w_{i-1}^n}{\Delta x^2} + O(\Delta x^2), \quad \text{(B8)}$$

while the derivative $\partial/\partial x^2 (I(x)(\partial^2 w/\partial x^2))$ is approximated by the θ scheme presented in (13). Here, it is not necessary to keep the explicit formulation of the numerical scheme, and thus the derivative $\partial^3 w/\partial x^2 \partial t$ in (1) is approximated by

$$\frac{(\delta_x^2 w)_i^{n+1} - (\delta_x^2 w)_i^{n-1}}{2\Delta t} + O(\Delta t^2). \quad \text{(B9)}$$

Equation (B9), together with (13) and (14), yields the implicit numerical formulation of the bar equation presented in (15).

In order to illustrate the efficiency of this scheme, its dispersion properties and the stability condition are calculated on the simplified bar equation

$$\frac{\partial^2 w(x,t)}{\partial t^2} = -a^2 \frac{\partial^4 w(x,t)}{\partial x^4}. \quad \text{(B10)}$$

With the use of the θ scheme presented in Sec. II, Eq. (B10) becomes

$$(\delta_t^2 w)_i^n = r^2 [\theta(\delta_x^4 w)_i^{n+1} + (1 - 2\theta)(\delta_x^4 w)_i^n + \theta(\delta_x^4 w)_i^{n-1}]. \quad \text{(B11)}$$

Using again the Fourier method yields the difference equation

$$W_{n+1} \left[1 + 16r^2 \theta \sin^4\left(\frac{k\Delta x}{2} \right) \right] - 2W_n [1 - 8r^2(1 - 2\theta) \times \sin^4\left(\frac{k\Delta x}{2} \right)] + W_{n-1} \left[1 + 16r^2 \theta \sin^4\left(\frac{k\Delta x}{2} \right) \right] = 0. \quad \text{(B12)}$$

The modulus of the roots in Eq. (B12) are bounded under the condition

$$4 \sin^4\left(\frac{k\Delta x}{2} \right) \left[(1 - 4\theta) \sin^4\left(\frac{k\Delta x}{2} \right) - 1 \right] \leqslant 0, \quad \text{(B13)}$$

which shows that the scheme is unconditionally stable for $\theta \geqslant 1/4$.

The estimation of frequencies with this scheme can be calculated by examining the propagation of a solution of the form

$$w_i^n = \exp\, j(ki\Delta x - \omega n \Delta t). \tag{B14}$$

Replacing w_i^n by its expression in Eq. (B11) yields the so-called ''dispersion'' relation between numerical frequency and wave number

$$\sin^2\left(\omega \frac{\Delta t}{2}\right) = \frac{r^2 \sin^4(k\Delta x/2)}{1+16\theta r^2 \sin^4(k\Delta x/2)}. \tag{B15}$$

Comparing Eq. (B15) with the continuous relation between ω and k:

$$\omega^2 = a^2 k^4, \tag{B16}$$

shows that the numerical scheme gives the best approximation if θ is minimal. The minimum value compatible with unconditional stability is therefore $\theta = 1/4$.

It is interesting to make here a parallel with some well-known results in the signal processing community, noticing that the case $\theta = 1/4$ corresponds to a discrete-time low-pass filter (1/4, 1/2, 1/4) which has a frequency response given by $H(\omega) = \cos^2(\omega T/2)$, i.e., which varies from 1 at frequency 0 down to 0 at frequency $f_s/2$.

[1] D. Holz, ''Investigations on acoustically important qualities of xylophone-bar materials: Can we substitute any tropical woods by European species?,'' in *Proceedings of the International Symposium on Musical Acoustics* (Jouve, Paris, 1995), pp. 351–357.

[2] J. L. Moore, ''Acoustics of bar percussion instruments'' Ph.D. thesis, Ohio State University, Columbus, Ohio, 1970.

[3] T. D. Rossing, ''Acoustics of percussion instruments: Part I,'' Phys. Teach. **14**, 546–556 (1976).

[4] N. H. Fletcher and T. D. Rossing, *Physics of Musical Instruments* (Springer–Verlag, New York, 1991).

[5] I. Bork, ''Zur Abstimmung und Kopplung von schwingenden Stäben und Hohlraumresonatoren,'' Ph.D. thesis, Dissertation, Technischen Universität Carolo-Wilhelmina, Braunschweig, 1983.

[6] I. Bork, ''Practical tuning of xylophone bars and resonators,'' Appl. Acoust. **46**, 103–127 (1995).

[7] I. Bork, ''Measuring the acoustical properties of mallets,'' Appl. Acoust. **30**, 207–218 (1990).

[8] F. Orduña-Bustamente, ''Nonuniform beams with harmonically related overtones for use in percussion instruments,'' J. Acoust. Soc. Am. **90**, 2935–2941 (1991); Erratum: J. Acoust. Soc. Am. **91**, 3582–3583 (1992).

[9] K. F. Graff, *Wave Motion in Elastic Solids* (Dover, New York, 1991).

[10] E. H. Lee, ''The impact of a mass striking a beam,'' Trans. ASME J. Appl. Mech. **62**, A–129 (1940).

[11] A. C. Eringen, ''Transverse impact on beams and plates,'' Trans. ASME J. Appl. Mech., 461–468 (1953).

[12] H. Hertz, ''Über die Berührung fester elasticher Körper,'' J. Math. **92**, 156 (1881).

[13] M. V. Gridnev, ''Impulsive excitation of the resonance boxes of plucked stringed instruments,'' Sov. Phys. Acoust. **20**, 431–434 (1975).

[14] G. Weinreich, ''Violin sound synthesis from first principles,'' J. Acoust. Soc. Am. Suppl. 1 **74**, S52 (1983).

[15] G. Weinreich, ''Synthesis of piano tones from first principles,'' J. Acoust. Soc. Am. Suppl. 1 **69**, S88 (1981).

[16] A. Chaigne and A. Askenfelt, ''Numerical simulations of piano strings. Part I. A physical model for struck string using finite difference methods,'' J. Acoust. Soc. Am. **2**, 1112–1118 (1994).

[17] S. H. Crandall, ''Numerical treatment of a fourth order parabolic partial differential equation,'' J. Assoc. Comput. Mach. **1**, 111–118 (1954).

[18] L. Collatz, ''Zur Stabilität des Differenzenverfahrens bei der Stabschwingungsgleichnung,'' Z. Angew. Math. Mech. **31**, 392–393 (1951).

[19] V. Doutaut, ''Etude expérimentale et simulation numérique d'instruments de percussion à clavier,'' Ph.D. thesis, Ecole Nationale Supérieure des Télécommunications, Paris, 1996.

[20] T. D. Rossing and D. Russell, ''Laboratory observation of elastic waves in solids,'' J. Acoust. Soc. Am. **58**, 1153–1162 (1990).

[21] V. Bucur, *Acoustics of Wood* (CRC, Boca Raton, FL, 1995).

[22] R. M. Christensen, *Theory of Viscoelasticity* (Academic, New York, 1982), 2nd ed.

[23] L. Landau and E. Lifchitz, *Théorie de l'élasticité* (Physique théorique, Mir, Paris, 1967).

[24] D. A. Russel and T. D. Rossing, ''Testing the nonlinear character of piano hammers using shock spectra,'' J. Acoust. Soc. Am. (submitted).

[25] T. Yanagisawa and K. Nakamura, ''Dynamic compression characteristics of piano felt,'' J. Acoust. Soc. Jpn. **40**, 725–729 (1984) (in Japanese).

[26] W. F. Ames, *Numerical Methods for Partial Differential Equations* (Academic, London, 1992), 3rd ed.

[27] G. Cohen (Editor), ''Méthodes numériques d'ordre élevé pour les ondes en régime transitoire,'' Collection Didactique, Ecole INRIA des ondes, INRIA Ed., Le Chesnay, 1994.

[28] R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-Value Problems* (Interscience, New York, 1967), 1st ed.

[29] A. Chaigne, ''On the use of finite differences for musical synthesis. Application to plucked stringed instruments,'' J. Acoust. **5**, 181–211 (1992).

[30] F. Orduña-Bustamante and A. Quintanar Isaias, ''A preliminary determination of the mechanical properties of four species of tropical wood from Mexico,'' J. Sound Vib. **154**, 365–368 (1992).

[31] J. Laroche, ''The use of the matrix pencil method for the spectrum analysis of musical signals,'' J. Acoust. Soc. Am. **4**, 1958–1965 (1994).

[32] I. Bork and J. Meyer, ''Zur klanglichen Bewertung von Xylophonen,'' Das Musikinstrument **31**, 1076–1081 (1982).

Ultrasonic wave propagation in bovine cancellous bone

A. Hosokawa and T. Otani
Department of Electrical Engineering, Faculty of Engineering, Doshisha University, Tanabe-cho, 610-03 Kyoto-fu, Japan

(Received 1 September 1995; accepted for publication 29 August 1996)

The acoustic properties of bovine cancellous (spongy) bone have been experimentally studied *in vitro* by the pulse transmission technique. Fast and slow longitudinal waves have been clearly identified when the acoustic wave propagates parallel to the direction of the trabeculae. Propagation speeds and attenuation of the fast and slow waves were observed in the frequency range of 0.5–5 MHz. Theoretical discussion is given to Biot's theory and the propagation of sound waves in fluid-saturated porous media. © *1997 Acoustical Society of America.* [S0001-4966(97)03201-3]

PACS numbers: 43.80.Cs [FD]

INTRODUCTION

Bone tissues are elastic, and can be classified into two types depending on macroscopic and microscopic features. Bone with a low volume fraction of solid (less than 70%) is called cancellous (spongy) bone and above 70% cortical (compact) bone.[1] Because of the complex anisotropy and inhomogeneity of cancellous and cortical bone, it is difficult to characterize bone tissues *in vivo* or *in vitro*. In particular, cancellous bone, being a two-component material comprised of a hard spongy core with cavities filled by soft tissue, has both strong anisotropy and inhomogeneity and is therefore especially difficult to analyze.

Biot's theory is a comprehensive theory for the ultrasonic propagation in fluid-saturated porous media,[2–4] such as porous rocks or sediments, and predicts the presence of both fast and slow longitudinal waves.

In previous attempts to study cancellous bone by ultrasound, results have been limited by the short bandpass used in the experimental studies, or Biot's theory has been applied in studies of only the fast wave.[5,6] Using wide band transducers (0.1–10 MHz), we report on the propagation of ultrasonic pulse waves in bovine cancellous bone. To our knowledge, for the first time, both fast and slow longitudinal waves in cancellous bone containing less than 70% solid fraction can be clearly identified in *in vitro* experimental observation. The acoustic propagation of both waves agree with Biot's theory.

I. THEORY

Biot initially proposed a general theory of elastic wave propagation in a system composed of a porous elastic solid saturated by a viscous fluid.[2–4] In Biot's theory, it is predicted that there should exist two longitudinal waves denoted by Biot as ''waves of the first and second kind.'' The former is a wave of fast speed corresponding to the wave in the solid and fluid moving in phase, while the latter is a slow speed corresponding to motion out of phase. The dispersion relation of both waves for isotropic materials is

$$\begin{vmatrix} Hl^2-\rho\omega^2 & \rho_f\omega^2-Cl^2 \\ Cl^2-\rho_f\omega^2 & m\omega^2-Ml^2-j\omega F(\kappa)\eta/k \end{vmatrix}=0, \tag{1}$$

with $l=l_r+jl_i$. The phase velocities ω/l_i and attenuation l_r can be calculated by solving Eq. (1). Here, H, C, and M are generalized elastic coefficients deduced by Biot and expressed in terms of the bulk moduli of the solid material (skeletal frame) K_s and pore fluid K_f, the bulk K_b, and shear modulus μ of the skeletal frame, and porosity β.[2,3,7] According to Johnson, the bulk K_b and shear moduli μ of the skeletal frame of cancellous bone change as a function of bone volume fraction $V_f(=1-\beta)$[1]:

$$K_b=\frac{E_s}{3(1-2\nu_b)}V_f^n, \tag{2}$$

$$\mu=\frac{E_s}{2(1+\nu_b)}V_f^n, \tag{3}$$

where n is a variable depending on the geometrical structure of the cancellous bone, and υ_b is the Poisson's ratio of the skeletal frame. In terms of the porosity β and the density of the solid ρ_s and fluid ρ_f, the total density ρ of the saturated medium is given by

$$\rho=(1-\beta)\rho_s+\beta\rho_f. \tag{4}$$

The density parameter m is used to account for the fact that not all of the pore fluid moves in the direction of the pressure gradient because of tortuosity or sinuosity, and is calculated by the equation

$$m=\alpha\rho_f/\beta. \tag{5}$$

Here, α is the structure factor and it is determined by the Berryman:[8]

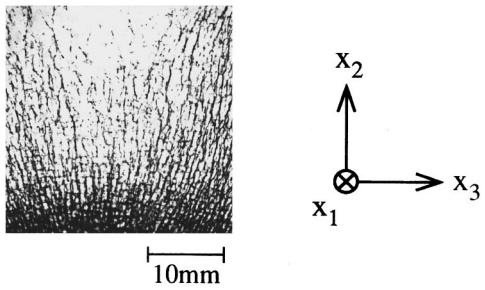
$$\alpha=1-r(1-1/\beta), \tag{6}$$

where r is a variable calculated from a microscopic model of a frame moving in the fluid. $F(\kappa)\eta/k$ takes the viscous resistance to fluid flow into account. η is the fluid viscosity, k is the permeability, and $F(\kappa)$ is defined by

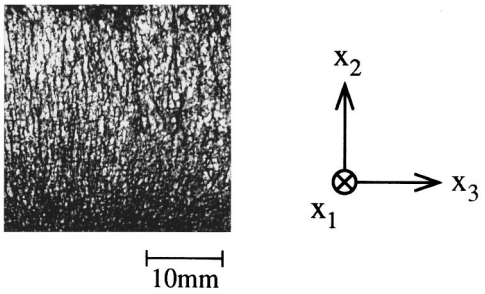
$$F(\kappa)=\frac{1}{4}\frac{\kappa T(\kappa)}{[1-2T(\kappa)/j\kappa]}, \tag{7}$$

where

$$T(\kappa)=\frac{\mathrm{ber}'(\kappa)+j\ \mathrm{bei}'(\kappa)}{\mathrm{ber}(\kappa)+j\ \mathrm{bei}(\kappa)}, \tag{8}$$



(a)



(b)

FIG. 1. X-ray photographs of cancellous bone: (a) low density; (b) high density.

$$\kappa = a(\omega\rho_f/\eta)^{1/2}. \tag{9}$$

Here the function ber(κ) and bei(κ) are the real and imaginary parts of the Kelvin function, and a is a parameter with dimension of length and depends on both the size and shape of the pores.

If the frequency is high enough to satisfy the condition,

$$\omega \gg 2\eta/\rho_f a_0^2, \tag{10}$$

where η is the fluid viscosity and a_0 the pore size, the attenuation mechanism has little effect on the propagation speeds and the speeds of the fast and slow waves are nondispersive.[9,10]

II. EXPERIMENTS

A. Specimens

Cancellous bone specimens, 20–30 mm in size and 9 or 7 mm thickness, were cut from the distal epiphysis of bovine femora, with soft tissue *in situ*. To observe the effect of changes in bone density, two specimens were taken with densities of 1120 and 1200 kg/m^3 [Fig. 1(a) and (b)]. As we predict that air penetrates the narrow space between the trabeculae and soft tissue of the specimens during the cutting and forming processes, prior to all experiments, each specimen was saturated with water followed by decompression to remove air bubbles. The total densities of these specimens were measured using Archimedes' principle, and the trabecular arrangement was estimated by x-ray photography. X-ray photographs of a typical trabecular structure in cancellous bone used in our experiments and its local coordinate system are shown in Fig. 1. The x_1, x_2, and x_3 are equivalent to the longitudinal, transverse, and sagittal axes in the general anatomic coordinates.

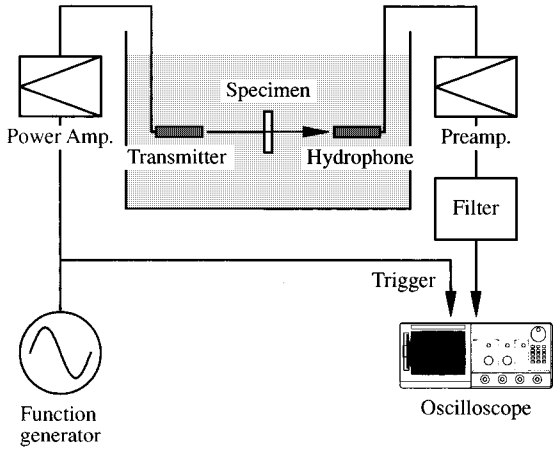


FIG. 2. Experimental arrangement for transmission of ultrasonic pulse.

B. Methods

The experimental arrangement for transmission of ultrasonic pulse is shown in Fig. 2. A wide band (0.1–10 MHz) PVDF transmitter and hydrophone were submerged in distilled water at 23±0.5 °C, and the specimens mounted between the transmitter and hydrophone at normal incidence. The transmitter was driven by a single sinusoidal impulse voltage of 50 V peak to peak in the frequency range of 0.5–5 MHz. Pulse waves propagating through the water/specimen/water system are detected by the hydrophone and the output signal amplified by a 40-dB preamplifier. The signal is then displayed on a digital storage oscilloscope. The phase and amplitude spectra of the received waveforms were calculated by an FFT algorithm, and the propagation speed ν and attenuation l_r were derived using the equation

$$\nu = \frac{\Delta d\,\nu_0}{[\Delta d - (\Delta\phi/\omega)\nu_0]}, \tag{11}$$

$$l_r = \ln \Delta V/\Delta d, \tag{12}$$

where Δd is the difference of the specimen thickness d_1 (=9 mm) and d_2 (=7 mm), ν_0 is the propagation speed in water, $\Delta\phi$ is the differences of phase spectra at an angular frequency ω of the signals traveling through the specimen of d_1 and d_2 thickness, and ΔV is the ratio of amplitude spectra for the two different thickness.

C. Results

Figure 3 shows a pulsed waveform at 1 MHz traveling in water. Figure 4 shows typical waveforms traveling through the cancellous bone in the direction of trabecular alignment (x_2 direction). Figure 4(a) is the waveform for the low density (ρ=1120 kg/m^3) specimen and (b) is the high density (ρ=1200 kg/m^3). In both Fig. 4(a) and (b), the fast and slow longitudinal waves can be clearly observed in the time domain. As the density increases (i.e., as the volume fraction of the solid bone increases), the amplitude of the fast

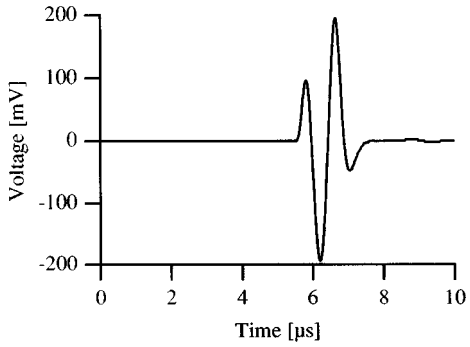
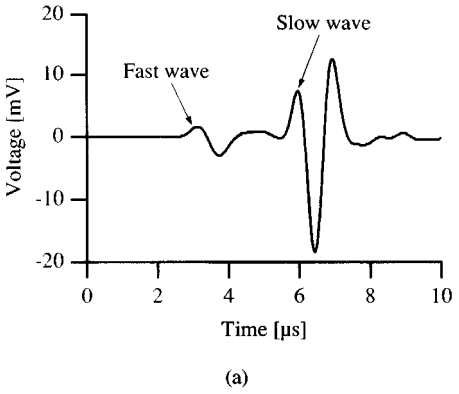


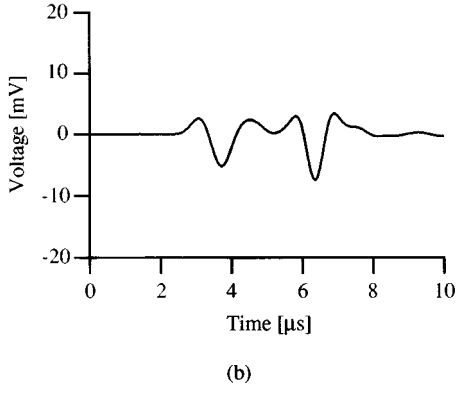
FIG. 3. Pulsed waveform at 1 MHz traveling in water.

wave becomes greater. At the same, the amplitude of the slow wave decreases. Accordingly, it can be deduced that the fast wave is associated with the solid core in cancellous bone, and the slow wave with the propagation in soft tissue.

Figure 5 shows the propagation speeds of the fast and slow waves in cancellous bone at 1 MHz as a function of bone volume fraction V_f. The speed of the fast wave varies from 2200 to 2700 m/s as the bone volume fraction increases. The slow wave remains constant at about 1400 m/s. The speed 1400 m/s of the slow wave is close to the propagation speed of 1450 m/s in bone marrow (see Table AI in the Appendix). Because of this, the ultrasonic properties of the soft tissue in cancellous bone can be expected to be similar to the bone marrow (the intact medullary portion of the



(a)



(b)

FIG. 4. Pulsed waveform at 1 MHz traveling through cancellous bone: (a) low density; (b) high density.

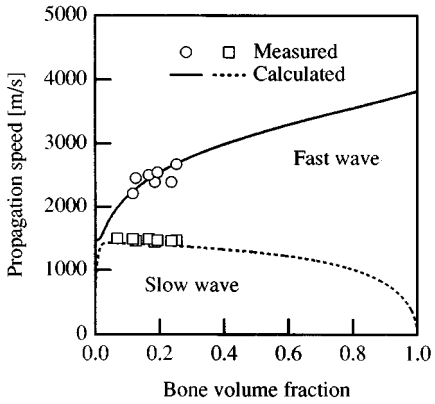


FIG. 5. Propagation speeds of fast and slow waves in cancellous bone at 1 MHz as a function of bone volume fraction V_f.

bone). It can therefore be assumed that the slow wave is mainly dominated by soft tissue. The speed 2200–2700 m/s of the fast wave is much slower than the propagation speed 3400–4200 m/s of cortical bone (Table AI). This can be explained by the fact that the cancellous bone is not solid but has a spongy structure.

Figure 6 shows the frequency dependence of the propagation speeds for one specimen of bone volume fraction V_f=0.19 (density ρ=1140 kg/m^3). No data for the fast wave at frequencies over 1.5 MHz was obtained because the weak amplitude of the signal made further measurement difficult. In Fig. 6, the propagation speeds of both the fast and slow waves are considered nondispersive in the range 0.5–5 MHz. Figure 7 shows the attenuation for three specimens of V_f=0.17, 0.19, and 0.25 (density ρ=1120, 1140, and 1200 kg/m^3). From Fig. 7(a), it can be seen that the data for the fast wave from bone volume fraction V_f=0.17, 0.19, and 0.25 are about the same. On the other hand, the attenuation of the slow wave [Fig. 7(b)] increases with frequency, and the variation of the slow wave with bone volume fraction V_f is large. The attenuation in cortical bone and bone marrow were also obtained (as shown in Table AII), and both the fast and slow waves in cancellous bone show much higher attenuation than the bulk wave in cortical bone (ρ=1960 kg/m^3) or bone marrow (ρ=930 kg/m^3). Accordingly, both

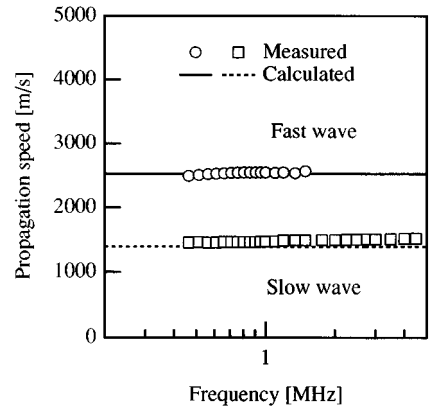
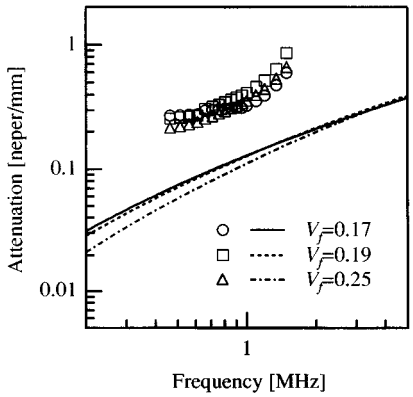
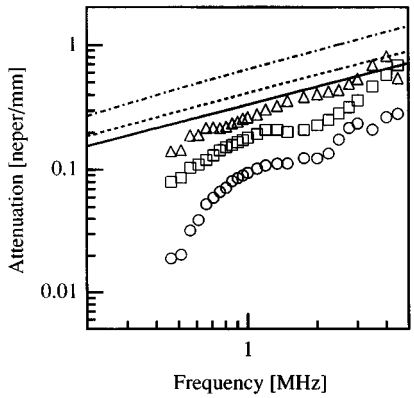


FIG. 6. Propagation speeds of fast and slow waves in cancellous bone at a bone volume fraction of V_f=0.19 as a function of frequency.



(a)



(b)

FIG. 7. Attenuation of fast and slow waves in cancellous bone at three bone volume fraction (V_f=0.17, 0.19, and 0.25) as a function of frequency: (a) fast wave; (b) slow wave.

the fast and slow waves are attenuated not only by the solid core or soft tissue component in the cancellous bone, but also by the structure of the cancellous bone.

III. DISCUSSION

Biot's theory[2–4] predicts two longitudinal waves propagating in fluid-saturated porous media. Because of the porous nature of cancellous bone, it is expected that the observed fast and slow waves in the previous section should correspond to the predicted two waves. Cancellous bone is composed of a solid core formed by the trabeculae and soft tissue which fills the pore spaces. The soft tissue is a sort of elastic body at normal temperature, but can be assumed to behave as a viscous fluid in Biot's theory.

The trabecular structure of cancellous bone is regarded as anisotropic, however Williams[3] showed that Biot's theory for isotropic media is applicable to cancellous bone, when the trabeculae is aligned along the propagation direction. We therefore have estimated the ultrasonic wave propagation in cancellous bone along the trabecular alignment by the same explanation.

Table I lists the parameters used in Biot's theory. The Young's modulus E_s=22 GPa and the Poisson's ratio ν_s=0.32 of the solid bone were taken from work by Williams and Johnson,[11] and are equivalent in value to cortical bone. The density of solid bone ρ_s, is given by Lang[12] as 1960 kg/m^3, and r=0.25 by Williams.[5] The bulk modulus K_f=2.0 GPa and density ρ_f=930 kg/m^3 of bone marrow are taken from our own experimental values as a substitute for the soft tissue of cancellous bone. The Poisson's ratio of the skeletal frame ν_{b23} (2 and 3 correspond to the x_2 and x_3 direction shown earlier in Fig. 1) is taken from studies by Williams and Johnson and equals 0.32.[11]

TABLE I. Biot's model parameters of cancellous bone.

Young's modulus of solid bone E_s	22 GPa
Poisson's ratio of solid bone ν_s	0.32
Density of solid bone ρ_s	1960 kg/m^3
Bulk modulus of bone marrow K_f	2.0 GPa
Density of bone marrow ρ_f	930 kg/m^3
Poisson's ratio of skeletal frame ν_{b23}	0.32
Variable r	0.25
Viscosity of bone marrow η	1.5 N s/m^2

The exponent n in Eqs. (2) and (3) depends on the direction of loading (namely, the direction of the ultrasonic propagation). Gibson derived analytically that n has a value of 1 when the material is loaded along the direction of trabecular alignment, and a value between 2 and 3 in the transverse direction.[1] However, the trabecular orientations of the tested specimens are partially random as shown in Fig. 1, and thus the value of n was estimated in order to fit the experimental data of the speed of the fast wave as a function of bone volume fraction V_f. This resulted in a value of n=1.46. Figure 5 shows the calculated curves for the speeds of the fast and slow waves at 1 MHz as a function of bone volume fraction V_f, with the measured data. Although the value of n=1.46 was obtained for the fast wave, good agreement could also be found with the slow wave. This estimating method was originally introduced by Williams, and a value of n=1.23 was estimated for water-saturated cancellous bone with aligned trabeculae obtained from bovine tibia.[5] The value of n=1.46 obtained in our experiments is higher than the value of 1.23 obtained by Williams. This could be explained if one assumes that the trabecular structure of our specimens are more random than that of the bovine tibial specimens, tested by Williams.

The propagation speeds of the fast and slow waves are theoretically nondispersive under Eq. (10). (For the viscosity of soft tissue of cancellous bone, a value of η=1.5 N s/m^2 was taken from the paper of McKelvie and Palmer[6] and is equivalent in value to bone marrow. The pore sizes a_0 of cancellous bone used in our experiments vary from about 0.5 to 1.5 mm. Then, a value of $2\eta/\rho_f a_0^2$ ranges between 10^3 and 10^4.) As shown in Fig. 6, the calculated curves of wave speed as a function of frequency are in good agreement with

TABLE II. Pore size and permeability.

Volume fraction V_f	Pore size a_0	Permeability k
0.17	1.35 mm	3×10^{-8}
0.19	1.20	2×10^{-8}
0.25	0.80	7×10^{-9}

measured values in the frequency range between 0.5 and 5 MHz. The attenuation of the fast and slow waves in the range of 0.5–5 MHz were also calculated from the data listed in Table I, $n=1.46$ and the pore size parameter a and permeability k data. The shape of the pores was ignored and the averaged pore size of each specimen, which was estimated by x-ray photography, was used as the pore size parameter. The permeability k is estimated in reference to the data of McKelvie and Palmer.[6] The pore sizes a_0 and permeabilities k for bone volume fraction $V_f=0.17$, 0.19 and 0.25 are listed in Table II. Calculated and measured attenuation results are shown in Fig. 7(a) and (b). Surprisingly throughout the frequency range, there is large difference between experimental and theoretical values of both the fast and slow waves. Additionally, the attenuation of the fast wave is much larger than that of the slow wave in spite of the opposite description in Biot's theory.[2,3] The primary reason for this could be due to water saturation in cancellous bone (strictly speaking, water saturation in the narrow spaces between the trabeculae and soft tissue in cancellous bone). Biot's theory in its basic form is developed considering the relative motion of the fluid in the pore spaces to the skeletal frame. The relative motion is, however, affected if water permeates the soft tissue. Consequently, the affect of the water saturation must be taken very seriously in any attempts to compare experiment to theory.

IV. CONCLUSIONS

Ultrasonic wave propagation through bovine cancellous bone has been experimentally examined. It was shown that both fast and slow longitudinal waves propagate through cancellous bone in the direction of the trabecular alignment, and that the fast wave is associated with the solid core of cancellous bone. The slow wave is associated with soft tissue in the pore spaces. The propagation of the fast and slow waves was estimated by Biot's theory for an isotropic medium by considering the elastic moduli in the propagation direction. The propagation speeds and attenuation of the fast and slow waves were measured and compared quantitatively. The measured speeds of the fast and slow waves are in good agreement with estimated results, and the waves were confirmed to correspond to the two waves predicted by Biot. A wide difference between measured and estimated attenuation values was observed because of the presence of water in the cancellous bone.

We conclude that Biot's theory is adequate to the ultrasonic wave propagation in cancellous bone although a more thorough examination is necessary to explain all observed phenomena.

APPENDIX

The acoustic properties of two components from which cancellous bone is made, solid bone and soft tissue component, was experimentally examined over a frequency range of 0.5–5 MHz. Cortical bone (90% solid core) cut from the diaphysis of bovine femur was used as the solid bone specimen. The intact medullary portion of the bone was used as the specimen instead of soft tissue. The propagation speeds and attenuation were measured in the same method as the cancellous bone, and are shown in Tables AI and AII. The propagation speeds both in cortical bone and in bone marrow could be regarded as nondispersive. The speeds in cortical bone are about 4200 m/s in the longitudinal direction, 3400 m/s in the radial, and 3500 m/s in the tangential. The speeds in bone marrow are about 1450 m/s both in the longitudinal and radial direction. The attenuation in cortical bone are 5.0×10^{-2} neper/(mm MHz) in the longitudinal direction, 2.5×10^{-2} neper/(mm MHz) in the radial, 2.7×10^{-2} neper/(mm MHz) in the tangential, and in bone marrow 1.3×10^{-2} neper/(mm MHz) in the longitudinal direction, and 1.2×10^{-2} neper/(mm MHz) in the radial.

TABLE AI. Propagation speeds in cortical bone and bone marrow.

Cortical bone	Longitudinal direction	4200 m/s
	Radial direction	3400
	Tangential direction	3500
Bone marrow	Longitudinal direction	1450 m/s
	Radial direction	1450

TABLE AII. Attenuation in cortical bone and bone marrow.

Cortical bone	Longitudinal direction	5.0×10^{-2} neper/(mm MHz)
	Radial direction	2.5×10^{-2}
	Tangential direction	2.7×10^{-2}
Bone marrow	Longitudinal direction	1.3×10^{-2} neper/(mm MHz)
	Radial direction	1.2×10^{-2}

[1] J. L. Gibson, ''The mechanical behavior of cancellous bone,'' J. Biomech. **18**, 317–328 (1985).
[2] M. A. Biot, ''Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low-frequency range,'' J. Acoust. Soc. Am. **28**, 168–178 (1956).
[3] M. A. Biot, ''Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range,'' J. Acoust. Soc. Am. **28**, 179–191 (1956).
[4] M. A. Biot, ''Generalized theory of acoustic propagation in porous dissipative media,'' J. Acoust. Soc. Am. **34**, 1254–1264 (1962).
[5] J. L. Williams, ''Ultrasonic wave propagation in cancellous and cortical bone: Prediction of some experimental results by Biot's theory,'' J. Acoust. Soc. Am. **91**, 1106–1112 (1992).
[6] M. L. McKelvie and S. B. Palmer, ''The interaction of ultrasound with cancellous bone,'' Phys. Med. Biol. **36**, 1331–1340 (1991).
[7] R. D. Stoll, ''Acoustic waves in saturated sediments,'' in *Physics of Sound in Marine Sediments*, edited by L. Hampton (Plenum, New York, 1974), pp. 19–39.
[8] J. G. Berryman, ''Confirmation of Biot's theory,'' Appl. Phys. Lett. **37**, 382–384 (1980).
[9] T. J. Plona and D. L. Johnson, ''Experimental study of the two bulk compressional modes in water-saturated porous structures,'' in *1980 Ultrasonic Symposium Proceedings*, edited by B. R. McAvoy (IEEE, New York, 1980), Vol. 1, pp. 868–872.
[10] D. L. Johnson and T. J. Plona, ''Acoustic slow waves and the consolidation transition,'' J. Acoust. Soc. Am. **72**, 556–565 (1982).
[11] J. L. Williams and W. J. H. Johnson, ''Elastic constants of composites formed from PMMA bone cement and anisotropic bovine tibial cancellous bone,'' J. Biomech. **22**, 673–682 (1989).
[12] S. B. Lang, ''Ultrasonic method for measuring elastic coefficients of bone and results on fresh and dried bovine bones,'' IEEE Trans. Biomed Eng. **BME-17**, 101–105 (1970).

Anisotropy of the apparent frequency dependence of backscatter in formalin fixed human myocardium

Christopher S. Hall, Edward D. Verdonk, Samuel A. Wickline, Julio E. Perez, and James G. Miller
Washington University, Department of Physics and School of Medicine, St. Louis, Missouri 63130

(Received 16 January 1996; accepted for publication 1 August 1996)

Measurements of the frequency dependence of ultrasonic backscatter are presented for specific angles of insonification for regions of infarcted and noninfarcted human myocardium. A 5-MHz transducer was used to insonify cylindrical cores taken from 7 noninfarcted regions and 12 infarcted regions of the left ventricular free wall of 6 formalin-fixed human hearts explanted because of ischemic cardiomyopathy. The dependence of apparent (uncompensated for diffraction effects and attenuation) backscatter on frequency was approximated by a power-law dependence, $|B(f)|^2 = af^n$. Under ideal conditions in a lossless medium, the effect of not compensating for the effects of diffraction and attenuation leads to the value of n to be 2.0 for Rayleigh scatterers while the frequency dependence of the fully compensated backscatter coefficient would be f^4. The value of n was determined over the frequency range, 3–7 MHz. Both noninfarcted and infarcted myocardium exhibited anisotropy of the frequency dependence of backscatter, with maxima occurring at angles that were perpendicular to the predominant myofiber direction and minima when parallel to the fibers. Perpendicular insonification yielded results for n of 1.8 ± 0.1 for noninfarcted myocardium and 1.2 ± 0.1 for infarcted myocardium while parallel insonification yielded results of 0.4 ± 0.1 for noninfarcted and 0.0 ± 0.1 for infarcted myocardium. The functional form of the angle-dependent backscatter is similar for both noninfarcted and infarcted myocardium, although the frequency dependence is clearly different for both tissue states for all angles of insonification. The results of this study indicate that the anisotropy of the frequency dependence of backscatter may play a significant role in ultrasonic imaging and is an important consideration for ultrasonic tissue characterization in myocardium.
[S0001-4966(97)02712-4]

PACS numbers: 43.80.Ev, 43.80.Jz [FD]

INTRODUCTION

A major goal of tissue characterization is to extend ultrasonic imaging beyond merely delineating the silhouette and gross structure of organs. Many clinical and laboratory animal studies have shown that ultrasound sensitively differentiates a variety of pathophysiologic states of tissue.[1–3] Ultrasound has been used successfully for imaging and tissue characterization of the heart,[1–15] liver,[16] kidney,[17,18] and other organs. Cardiovascular medical ultrasound has primarily been employed to detect structural and functional aspects of the myocardium, such as the evaluation of ventricular ejection fraction, the presence of abnormal ventricular wall motion, the presence of valvular disease, etc. Furthermore, we and others have shown that useful diagnostic parameters that relate to tissue structure and function can be obtained by analyzing the interaction of ultrasound with the muscle of the heart wall.[2,3,5,11]

Specifically, ultrasonic backscatter has been used to differentiate pathologies such as acute ischemia,[19,20] completed infarct,[21,22] myocarditis,[3] and the extent of myocardial reperfusion.[19] Anisotropy of the acoustic properties of myocardium plays a role in the observed reduction in signal intensity (i.e., echo ''dropout'') in specific views of the heart, as well as in the ability to utilize tissue characterization for diagnostic purposes.[4] In particular, recent studies have demonstrated that the time delay of myocardial cyclic variation with respect to the global timing of cardiac contraction may be dependent upon the relative angle of insonification with respect to the muscle fiber direction.[5] We and others have shown that myocardium exhibits an anisotropy of backscatter,[7,8,23,24] longitudinal velocity,[13,25] and attenuation coefficient.[26,27] It has been demonstrated that the extent of variation of these ultrasonic parameters correlates well with the microstructure of the myocardium.[9,28,29]

Previous studies have shown that the extent of anisotropy can be used to differentiate infarcted from noninfarcted myocardium both *in vitro* as well as *in vivo*.[3,9] The variation of the frequency dependence of myocardial backscatter with the heart cycle was reported by Wear *et al.*, demonstrating that the frequency dependence is greater at end-diastole and less at end-systole.[14] It would be desirable to obtain *in vivo* measurements of the myocardial backscatter coefficient, but this is difficult to accomplish because the exact frequency dependence of the tissue's attenuation and the transducer's diffraction are not known. Hence, a more applicable measurement for a clinical setting is the use of the ''apparent'' frequency dependence of backscatter which does not compensate for attenuation or diffraction effects. In this study, we report results for measurements of the anisotropy of the apparent frequency dependence of backscatter for both noninfarcted and infarcted myocardium.

The myocardium can be described in terms of an ar-

rangement of nonisotropic scatterers associated with the myofibers. These myofibers exist in a collagen matrix and consist of bundles of cells called myocytes. Although histology has provided information about the anisotropy of the myofibers throughout the heart wall, the exact source of the anisotropy of backscatter has remained a perplexing question. We have developed an ultrasonic scattering model which treats the myocyte as an ellipsoidal shell consisting of collagen-like material representing the interstitial collagen which is hypothesized to be the primary scatterer in the myocardium.[12] The acoustic impedance mismatch between the surrounding medium and the collagen sheath of the myocyte provides the acoustic contrast, and the ellipsoidal shape of the myocyte accounts for the anisotropy of backscatter. A previous study demonstrated good agreement between previously measured values of the anisotropy of frequency-averaged backscatter and model predictions.[12]

The goal of the present work was to measure the anisotropy of the frequency-dependent backscatter from human myocardium and to clarify possible morphological and ultrasonic mechanisms for the observed anisotropy of backscatter from the myocardium. Section II deals with the methods used to measure and to model the frequency dependence of backscatter. Section III reports our results. Section IV discusses possible mechanisms for the anisotropy of the frequency-dependent backscatter and identifies some of the potential limitations of this study.

I. METHODS: ULTRASONIC MEASUREMENTS

A. Pathologic specimens

Six explanted, formalin-fixed human hearts were obtained from patients who underwent transplantation for ischemic cardiomyopathy. Specimens were prepared from these hearts as described in an earlier report from our laboratory.[21] Transmural cylindrical samples were taken from the left ventricular free wall with a 14-mm cork borer. The predominant orientation of the myofibers is perpendicular to the long axis of the cylindrical plug so that it is possible to insonify at any one transmural level in such a way as to go from perpendicular to parallel incidence by rotating the plug around its axis of symmetry. The samples were attached to styrofoam plugs with cyanoacrylate and stored in a 10% formalin solution for several months to several years. Seven samples were taken from regions of noninfarcted myocardium and 12 samples were taken from regions of infarct. Histologic examination of the tissue was used to classify the samples into groups of noninfarcted and infarcted tissue.

B. Acquisition of ultrasonic data

The method of obtaining backscatter data has been previously described by us[15,23] and is summarized below. Heart tissue plugs were placed in a room-temperature, degassed water bath and allowed to attain thermal equilibrium. A fiducial mark was made on the styrofoam plugs that indicated a rough estimate of the fiber direction in the midmyocardium. The tissue plugs were mounted on an apparatus that allowed the sample to be rotated through a full 360° in 2° increments. The experimental setup is shown in a block diagram in Fig. 1. A 5-MHz, 2-in. focal length, 0.5-in.-diam transducer (Panametrics V309) was used to insonify the sample at each angle. A total of five independent site averages occurred at each angle by translating the transducer laterally in 1-mm steps while maintaining the same angle of insonification. The focal region of the transducer was placed at the front surface of the sample. The backscattered signal was received and amplified in a Metrotek receiver (model 106). A 14-μs (approximately 10.8 mm) segment of the rf signal, corresponding to the tissue throughout the entire sample excluding the front and back wall specular reflections, was gated (Metrotek 701 stepless gate). The gated signal was passed to a spectrum analyzer (Hewlett–Packard 8557A) where the power spectrum was captured and sent to a Macintosh II computer via IEEE-488 GPIB to be stored off-line. Each power spectrum consisted of 481 samples from 0 to 10 MHz. A total of 29 scans of noninfarcted myocardium and 33 scans of infarcted myocardium were performed.

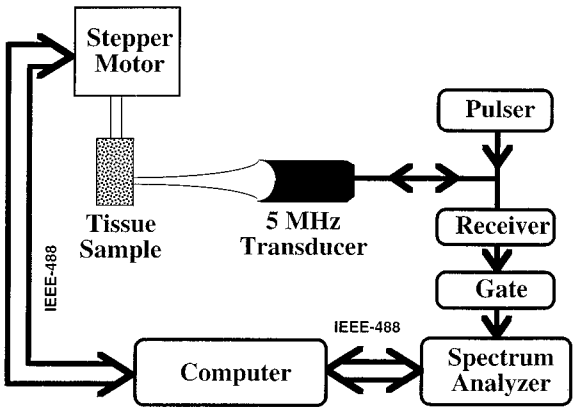


FIG. 1. Experimental setup used to measure the anisotropy of the frequency dependence of backscatter.

C. Analysis of ultrasonic data

In order to compensate for the electromechanical effects of the measurement system in the backscatter data from the tissue, a calibration spectrum was obtained from the reflection from a polished stainless steel plate. The steel plate was placed in the focal region of the transducer and aligned so that its surface was perpendicular to the ultrasonic beam. The reflection coefficient for polished steel in water is very close to 1.0 so the assumption is made that all of the transmitted signal is reflected back and received by the transducer. With some approximations, the backscattered signal can be modeled as a convolution of effects due to attenuation, diffraction, backscatter intrinsic to the material being insonified, and the electromechanical response of the transducer and electronics. This can be expressed as a function of the frequency, f, and the distance from the transducer to the insonified volume, r; in equation form $|E_{\text{tissue}}(f)|^2 = |P(f)|^2|D(f,r)|^2|A(f,r)|^2|S(f)|^2$, where $|P(f)|^2$ is the power spectrum due to the electronics, $|D(f,r)|^2$ is the effect of the diffraction of the transducer's beam, $|A(f,r)|^2$ is the effect of attenuation throughout the insonified volume of tissue, and $|S(f)|^2$ is the backscatter intrinsic to the material.

The expression above is valid in the case when the tissue of interest is placed in the focal zone or far field of the transducer where the largest component of the insonifying wave is planar. This approximation allows the diffraction pattern to be decoupled from the effects due to attenuation. By normalizing the spectra returned from the tissue by the spectrum of the steel plate, $|E_{\mathrm{ref}}(f)|^2=|P(f)|^2$, effects of the measurement system can be removed $|B(f)|^2=|E_{\mathrm{tissue}}(f)|^2/|E_{\mathrm{ref}}(f)|^2$. The resulting calculation, $|B(f)|^2$, contains the frequency-dependent effects of backscatter, attenuation, and diffraction. In our laboratory, we employ the terminology ''apparent'' backscatter transfer function to indicate that no compensation is made for the frequency dependence of attenuation and diffraction.

The apparent backscatter transfer function, $|B(f)|^2$, can be modeled with a power-law expression, $|B(f)|^2=af^n$, where f represents frequency, a is a constant of proportionality, and n is the power-law frequency dependence. From this equation, it is possible to derive the power-law frequency dependence by making a linear least-squares fit to a plot of the logarithm of the apparent backscatter transfer function versus the logarithm of frequency. The results presented in the next section are plots of this value, n, versus angle of insonification. Under ideal conditions in a lossless medium, the value of n would be 2.0 for Rayleigh scatterers whose fully compensated frequency dependence would be f^4. Under the assumptions in the previous paragraph, the expression for $|B(f)|^2$ contains components due to attenuation, diffraction, and intrinsic scattering. The backscatter coefficient is proportional to the intrinsic scattering, $|S(f)|^2$, and for Rayleigh scatterers this should have a frequency dependence of f^4. If the medium is lossless then the component due to attenuation does not contribute any frequency dependence of backscatter. The diffraction term in the above expression has a frequency dependence of f^{-2}. Hence, the frequency dependence for $|B(f)|^2$ must be 2.0 for the case of Rayleigh scatterers.

Although the fiducial mark is a useful method for aligning the samples in the water tank, it is too coarse for use in averaging multiple plots of anisotropy. In order to improve the averaging procedure, the separate anisotropy curves of apparent backscatter transfer functions averaged over frequencies 3–7 MHz were cross correlated.[15] Integrated backscatter (IB) was used because broadband averaging allows a low noise, robust method for identifying the angle of insonification with respect to the fibers. The maximum value of the correlation was used to determine the angular shift that produced the greatest overlap between anisotropy functions. With this information of angular orientation, the apparent backscatter transfer functions of similar angles of insonification were averaged to obtain one apparent backscatter transfer function for each angle of insonification. A linear least-squares fit was made to the plot of the logarithm of the average apparent backscatter transfer function versus the logarithm of frequency to obtain the power-law frequency dependence for every angle.

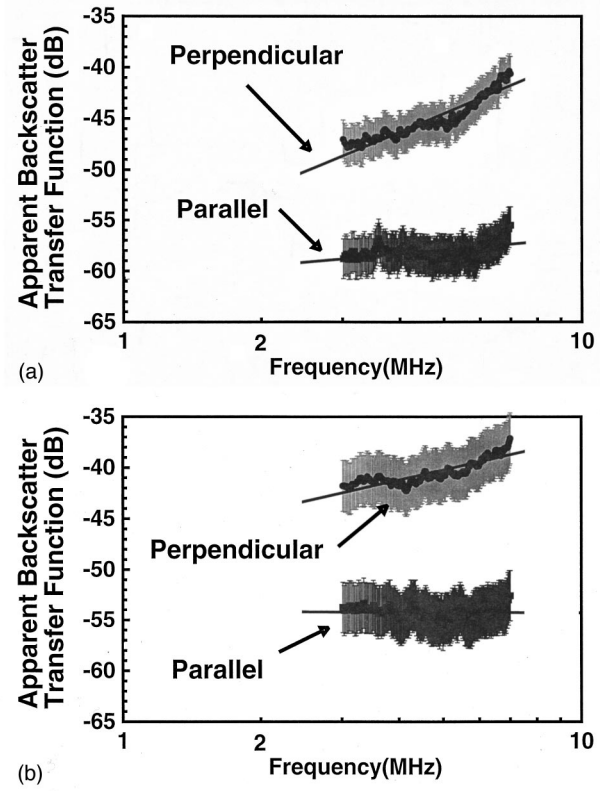


FIG. 2. (a) Average apparent backscatter transfer function for noninfarcted myocardium for both parallel and perpendicular insonification. (b) Average apparent backscatter transfer function for infarcted myocardium for both parallel and perpendicular insonification.

II. RESULTS

A. Frequency dependence of backscatter

The results of 29 scans of noninfarcted myocardium and 33 scans of infarcted myocardium are reported. The scans for both tissue types were averaged in the manner summarized in Sec. I. The resulting average apparent backscatter transfer functions are shown for both perpendicular and parallel incidence for both noninfarcted and infarcted tissue in Fig. 2. The power-law frequency dependence for noninfarcted myocardium is $n=1.8\pm0.1$ perpendicular to the myofibers and $n=0.4\pm0.1$ parallel. The results for infarcted myocardium yielded a similar differentiation between perpendicular ($n=1.2\pm0.1$) and parallel ($n=0.0\pm0.1$) orientations. The error bars reflect the slopes of the lines within the 95% confidence limits (StatView 4.0, Abacus Concepts).

Figure 3(a) shows the measurement of the power-law frequency dependence for noninfarcted myocardium in 2° increments. The functional form of the anisotropy of the frequency dependence of backscatter differs somewhat near perpendicular from near parallel. The shape of the curve suggests that the function may vary more rapidly near perpendicular incidence than near parallel. However, the amount of variation in the curves makes any quantitative measurements of this difference problematic. Figure 3(b) shows the results for infarcted myocardium which exhibit a similar functional form to the results of noninfarcted myocardium. Figure 3(c) compares anisotropy plots of the slopes of the apparent backscatter transfer function for both infarcted and noninfarcted

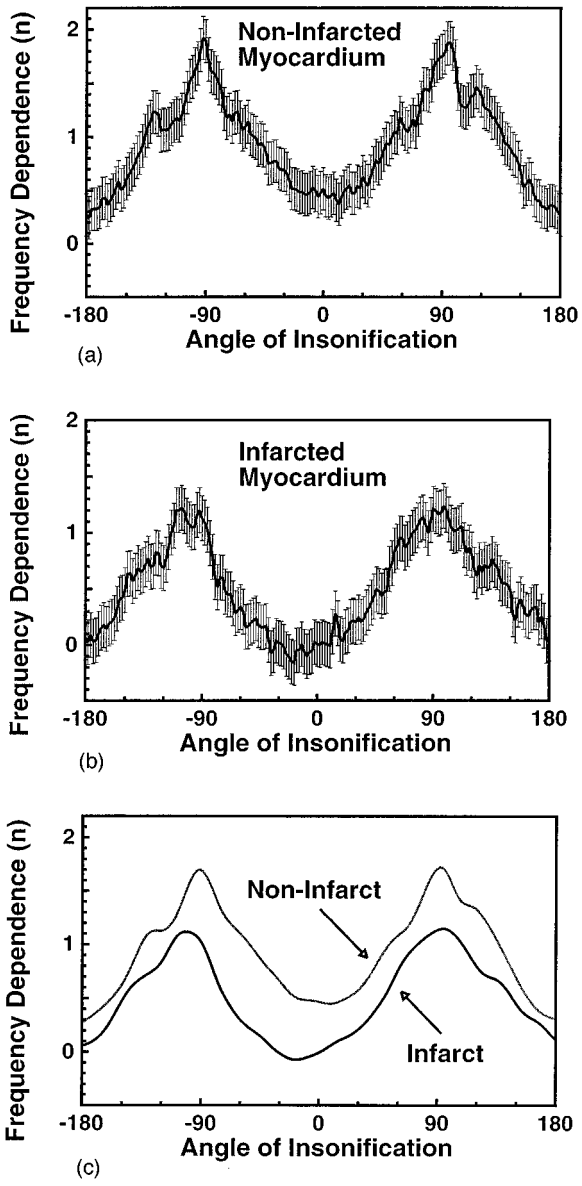


FIG. 3. (a) Anisotropy of the frequency dependence of backscatter for noninfarcted myocardium. (b) Anisotropy of the frequency dependence of backscatter for infarcted myocardium. (c) Anisotropy of the frequency dependence of backscatter for both infarcted and noninfarcted myocardium filtered with 20 passes of a three-point binomial filter.

myocardium. The data are the same as in Fig. 3(a) and (b), but have been smoothed with 20 passes of a three-point binomial filter in order to emphasize the general features.

III. DISCUSSION

The left ventricular wall of human myocardium exhibits a complex three-dimensional fiber architecture as the fibers change orientation transmurally through the heart wall and longitudinally around the left ventricle. This results in a dependence of many ultrasonic parameters on the angle of insonification. The anisotropy of attenuation[26] and longitudinal velocity[13,25] have been previously measured by our laboratory for human and canine myocardium.

In the present study, we have demonstrated that the dependence of the apparent backscatter transfer function with frequency changes with the angle of insonification with respect to the myofibers. The apparent backscatter transfer function, unlike the backscatter coefficient, is not compensated for frequency-dependent attenuation and the effects of diffraction. Hence, our reported values for n are less than those associated with the backscatter coefficient. For scattering in soft tissues and in the absence of resonances at specific frequencies, the frequency dependence of the backscatter coefficient ($n_{\text{b.c.}}$) can range between 0 and 4, where $n_{\text{b.c.}}=0$ represents scatterers much larger than the wavelength of sound and $n_{\text{b.c.}}=4$ represents scatterers much smaller than the wavelength. However, for a lossless medium containing Rayleigh scatterers, the theoretical value of the power-law dependence of the apparent backscatter transfer function is $n=2$ since it still has the effects of the frequency-dependent attenuation and diffraction convolved.

The apparent backscatter transfer function behaves differently when the insonification is near perpendicular incidence than when parallel to the myofibers. The dependence of n with the angle of insonification varies faster with angles near perpendicular incidence compared with angles near parallel incidence. This result could affect the interpretation of the measurements of frequency dependence of backscatter and to a lesser extent the use of integrated backscatter in an *in vivo* setting where the exact orientation of the transducer to myofiber direction is often unknown and can vary over the heart cycle. A variation of a few degrees off perpendicular incidence might yield misleading results that mistake normal tissue for infarcted tissue. Therefore, exact knowledge of the angle of incidence of the insonifying sound wave with respect to the predominant myofiber direction is important before the diagnostic use of the apparent frequency dependence of backscatter becomes feasible.

This angular dependence of the frequency dependence of the apparent backscatter transfer function may be used to infer the relative scatterer size at each angle of insonification and as a result, the shape of the fundamental scatterer in the myocardium. The shape of the anisotropy curves suggests a fundamental scatterer that has a geometric cross section which changes with angle, e.g., a cylindrical or ellipsoidal shape. Because of the polarization of the incoming ultrasonic wave (i.e., longitudinal), the cross section of interest is determined by the distance that the wave has to traverse through the scatterer. In the case of parallel incidence, the incoming wave sees a scatterer with a larger cross section because it must transverse the entire length of the myocyte. Hence, the frequency dependence for parallel incidence is smaller than for perpendicular incidence because the effective scatterer size is larger. Previous studies have modeled the fundamental scatterer in myocardium with similar shapes.[12,30]

The effective differentiation of infarcted tissue from noninfarcted tissue on the basis of the frequency dependence of backscatter has been demonstrated in a previous study.[21] In this study, we have chosen to employ the *apparent* frequency dependence, a measurement which is susceptible to differences in attenuation between infarcted and noninfarcted tissue. However, measurements of the frequency-dependent attenuation for the same tissue used in this study were unable

to differentiate infarcted from noninfarcted myocardium.[27] This finding differs from previous measurements made in this and other laboratories where a successful differentiation was made between infarcted and noninfarcted tissue on the basis of the frequency dependence of attenuation.[31] The apparent disagreement might be due to the effects of formalin fixation.[27] The similarity of the measurement of attenuation for infarcted and noninfarcted myocardium makes any compensation for the effects of attenuation and diffraction identical for both tissues. Therefore, the differences in the apparent frequency dependence of backscatter between infarcted and noninfarcted tissue are not attributable to attenuation. Figure 3(c) illustrates that infarcted tissue has a frequency dependence which is less than that of the noninfarcted tissue for each angle of insonification. The presence of collagen crosslinking as well as other morphological changes in the infarcted tissue may have induced the size of the fundamental scatterer to increase and would explain the smaller apparent frequency dependence of backscatter. Another interesting observation can be made about the structure of the infarct zone. The anisotropy of both integrated backscatter and the frequency dependence of backscatter remains in zones of infarct, which implies that there is still a great deal of organization. This suggests that the fibers' alignment is preserved even though the fundamental size of the scatterer might change.

In summary, we have demonstrated that human myocardium exhibits an anisotropy of the frequency dependence of the apparent backscatter transfer function. This anisotropy cannot be solely explained by the known anisotropy of the frequency-dependent attenuation because previous measurements of the attenuation for these tissues yielded results that were similar for infarcted and noninfarcted myocardium. Results of this study stress the importance of the knowledge of transducer orientation with respect to the myocardial fiber direction in order to employ measurements of the frequency dependence of myocardial backscatter to differentiate infarcted from noninfarcted tissue. Such knowledge will be difficult to obtain in a clinical setting. However, advances in both conventional two-dimensional echocardiography and three-dimensional ultrasonic imaging may make clinical diagnostic use of anisotropy of ultrasonic parameters feasible based on the knowledge of the spatial position of the transducer with respect to the beating heart.

ACKNOWLEDGMENTS

This work was supported in part by NIH Grants Nos. HL40302 and HL42950, and an AHA Established Investigator Award.

[1] D. T. Linker, B. A. Angelsen, and R. L. Popp, ''Autocorrelation length of normal and myopathic human myocardium measured by acoustic microscopy: implications for clinical ultrasonic tissue characterization,'' Ultrasound Med. Biol. **16**, 793–799 (1990).

[2] K. B. Sagar, L. R. Pelc, T. L. Rhyne, R. A. Komorowski, L. S. Wann, and D. C. Warltier, ''Role of ultrasonic tissue characterization to distinguish reversible from irreversible myocardial injury,'' J. Am. Soc. Echocardiogr. **3**, 471–477 (1990).

[3] 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).

[4] M. Aygen and R. L. Popp, ''Influence of the orientation of myocardial fibers on echocardiographic images,'' Am. J. Cardiol. **60**, 147–152 (1987).

[5] A. E. Finch-Johnston, H. M. Gussak, J. Mobley, M. R. Holland, O. Petrovic, J. E. Perez, and J. G. Miller, ''Effect of time delay on the apparent magnitude of cyclic variation of myocardial ultrasonic backscatter in standard echocardiographic views,'' Ultrason. Imag. **17**, 77 (1995).

[6] C. S. Hall, E. D. Verdonk, J. E. Perez, S. A. Wickline, and J. G. Miller, ''Effects of myocardial anisotropy on differentiating acutely ischemic myocardium from mature infarct,'' Ultrason. Imag. **17**, 76 (1995).

[7] E. I. Madaras, J. Perez, B. E. Sobel, J. G. Mottley, and J. G. Miller, ''Anisotropy of the ultrasonic backscatter of myocardial tissue: II. Measurements *in vivo*,''J. Acoust. Soc. Am. **83**, 762–769 (1988).

[8] J. G. Mottley and J. G. Miller, ''Anisotropy of the ultrasonic backscatter of myocardial tissue: I. Theory and measurements *in vitro*,''J. Acoust. Soc. Am. **83**, 755–761 (1988).

[9] M. O'Donnell, J. W. Mimbs, and J. G. Miller, ''The relationship between collagen and ultrasonic backscatter in myocardial tissue,''J. Acoust. Soc. Am. **69**, 580–588 (1981).

[10] D. Recchia, C. Hall, R. K. Shepard, and J. G. Miller, ''Mechanisms of the view-dependence of ultrasonic backscatter from normal myocardium,''IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42**, 91–98 (1995).

[11] T. L. Rhyne, K. B. Sagar, S. L. Wann, and G. Haasler, ''The myocardial signature: Absolute backscatter, cyclical variation, frequency variation, and statistics,'' Ultrason. Imag. **8**, 107–120 (1986).

[12] J. H. Rose, M. R. Kaufmann, S. A. Wickline, C. S. Hall, and J. G. Miller, ''A proposed microscopic elastic wave theory for ultrasonic backscatter from myocardial tissue,''J. Acoust. Soc. Am. **97**, 656–668 (1995).

[13] E. D. Verdonk, S. A. Wickline, and J. G. Miller, ''Anisotropy of ultrasonic velocity and elastic properties in normal human myocardium,''J. Acoust. Soc. Am. **92**, 3039–3050 (1992).

[14] K. A. Wear, M. R. Milunski, S. A. Wickline, J. E. Perez, B. E. Sobel, and J. G. Miller, ''Contraction-related variation in frequency dependence of acoustic properties of canine myocardium,''J. Acoust. Soc. Am. **86**, 2067–2072 (1989).

[15] S. A. Wickline, E. D. Verdonk, and J. G. Miller, ''Three-dimensional characterization of human ventricular myofiber architecture by ultrasonic backscatter,'' J. Clin. Invest. **88**, 438–446 (1991).

[16] E. Madsen, J. Zagzebski, M. Insana, T. Burke, and G. Frank, ''Ultrasonically tissue-mimicking liver including the frequency dependence of backscatter,''Med. Phys. **9**, 703–710 (1982).

[17] J. M. Rubin, P. L. Carson, and C. R. Meyer, ''Anisotropic ultrasonic backscatter from the renal cortex,'' Ultrasound Med. Biol. **14**, 507–511 (1988).

[18] M. F. Insana, R. F. Wagner, D. G. Brown, and T. J. Hall, ''Describing small-scale structure in random media using pulse ultrasound,''J. Acoust. Soc. Am. **87**, 179–192 (1990).

[19] M. R. Milunski, G. A. Mohr, J. E. Perez, Z. Vered, K. A. Wear, C. J. Gessler, B. E. Sobel, J. G. Miller, and S. A. Wickline, ''Ultrasonic tissue characterization with integrated backscatter: Acute myocardial ischemia, reperfusion, and stunned myocardium in patients,''Circulation **80**, 491–503 (1989).

[20] J. W. Mimbs, D. Bauwens, R. D. Cohen, M. O'Donnell, J. G. Miller, and B. E. Sobel, ''Effects of myocardial ischemia on quantitative ultrasonic backscatter and identification of responsible determinants,'' Circ. Res. **49**, 89–96 (1981).

[21] S. A. Wickline, E. D. Verdonk, B. E. Sobel, and J. G. Miller, ''Identification of human myocardial infarction *in vitro* based on the frequency dependence of ultrasound backscatter,''J. Acoust. Soc. Am. **91**, 3018–3025 (1992).

[22] 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).

[23] B. K. Hoffmeister, A. K. Wong, E. D. Verdonk, S. A. Wickline, and J. G. Miller, ''Comparison of the anisotropy of apparent integrated ultrasonic backscatter from fixed human tendon and fixed human myocardium,''J. Acoust. Soc. Am. **97**, 1307–1313 (1995).

[24] B. K. Hoffmeister, A. K. Wong, E. D. Verdonk, S. A. Wickline, and J. G. Miller, ''Anisotropy of ultrasonic backscatter from human tendon compared to that from normal human myocardium,'' Proc. IEEE Ultrason. Symp. 91, CH 3079-1, 1127–1131 (1991).

[25] B. K. Hoffmeister, E. D. Verdonk, S. A. Wickline, and J. G. Miller,

''Effect of collagen on the anisotropy of quasi-longitudinal mode ultrasonic velocity in fibrous soft tissues: A comparison of fixed tendon and fixed myocardium,''J. Acoust. Soc. Am. **96**, 1957–1964 (1994).
[26] J. G. Mottley and J. G. Miller, ''Anisotropy of the ultrasonic attenuation in soft tissues: Measurements *in vitro*,''J. Acoust. Soc. Am. **88**, 1203–1210 (1990).
[27] E. D. Verdonk, B. K. Hoffmeister, S. A. Wickline, and J. G. Miller, ''Anisotropy of the slope of ultrasonic attenuation in formalin fixed human myocardium,''J. Acoust. Soc. Am. **99**, 3837–3843 (1996).
[28] K. K. Shung and J. M. Reid, ''Ultrasonic scattering from tissue,'' Proc. IEEE Ultrason. Symp. 77, CH 1264-1SU, 230–233 (1977).
[29] D. Y. Fei and K. K. Shung, ''Ultrasonic backscatter from mammalian tissues,''J. Acoust. Soc. Am. **78**, 871–876 (1985).
[30] K. N. Kumar and J. G. Mottley, ''A computer algorithm for calculating scattering cross sections from cylindrical and spherical scatterers using a time-domain born approximation,'' Ultrason. Imag. **13**, 201 (1991).
[31] M. O'Donnell, J. W. Mimbs, and J. G. Miller, ''The relationship between collagen and ultrasonic attenuation in myocardial tissue,''J. Acoust. Soc. Am. **65**, 512–517 (1979).

In vitro and *in vivo* enhancement of sonodynamically active cavitation by second-harmonic superimposition

Shin-ichiro Umemura,[a] Ken-ichi Kawabata, and Kazuaki Sasaki
Advanced Research Laboratory, Hitachi Ltd., Hatoyama, Saitama 350-03, Japan

(Received 29 September 1995; accepted for publication 6 August 1996)

Acoustic cavitation, the primary mechanism of sonochemical effects, is known to be induced more easily by standing waves than by progressive waves. It has been found that acoustic cavitation can be an order of magnitude enhanced by superimposing the second harmonic on the fundamental. Significant synergistic effects between the fundamental and the second harmonic were observed in both *in vitro* and *in vivo* experiments employing a progressive wave field. Second-harmonic superimposition induces *in vitro* sonochemical reaction as well as fractional harmonic emission at a relatively low ultrasonic intensity even in a progressive wave field. The effect of second-harmonic superimposition was also investigated using exteriorized mouse livers suspended in degassed saline. The intensity threshold for the production of focal tissue damage, paired with fractional harmonic emission was significantly lowered by second-harmonic superimposition especially when a sonodynamically active agent had been administered to the mouse. Insonation with second-harmonic superimposition in combination with such administration may have potential use for selective tumor treatment. © *1997 Acoustical Society of America.* [S0001-4966(97)05612-9]

PACS numbers: 43.80.Gx, 43.80.Sh [FD]

INTRODUCTION

Ultrasound has a low enough tissue attenuation coefficient to penetrate intervening tissues and deliver its energy to nonsuperficial objects while maintaining the ability to focus the energy into small volumes. This is a unique advantage when compared to electromagnetic modalities such as laser beams or microwaves in the applications for noninvasive and selective treatment of nonsuperficial tumors. Use of ultrasonically induced thermal effects in such applications has been studied mainly in two categories: ultrasonic hyperthermia[1] and high-intensity focused ultrasound surgery.[2,3] In the former approach, a tumor is heated above 43 °C for more than a period in the order of 10 min. In the latter, the tumor tissue in a focal spot is heated above its coagulation temperature for a period in the order of a second. In comparison with the thermal effects, therapeutic use of nonthermal effects of ultrasound has been much less studied.[4–10]

Ultrasonically induced cavitation is known to be the primary cause of sonoluminescence and sonochemical reactions,[11] and besides lithotripsy, may have the highest potential for therapeutic applications among the nonthermal effects of ultrasound if it can be somehow controlled. Recently, *in vitro* and *in vivo* experiments have demonstrated that ultrasound can activate certain porphyrins such as hematoporphyrin (Hp) and thereby induce significant antitumor effects. Suspended tumor cells were lysed by ultrasound at a much higher rate in the presence of such porphyrins than in their absence at concentrations for which the porphyrins by themselves show no cytotoxicity.[9,12,13] Experimental tumors implanted into mice have been treated with ultrasound in combination with administration of such porphyrins and the tumor growth was significantly inhibited at an ultrasound intensity for which ultrasound alone showed no significant effect.[7,9,10] These experimental results allowed us to propose a new modality of tumor treatment, which we call ''sonodynamic therapy.'' [14]

The above described *in vitro* and *in vivo* experiments were performed under standing wave conditions because acoustic cavitation can more easily be induced with reproducibility by standing waves than by progressive waves. In a standing wave field, microbubbles smaller than the resonant size are driven by ultrasonic radiation force toward a pressure antinode, where they can efficiently grow and collapse.[15] However, insonation with standing waves does not seem to be widely applicable to various therapeutic situations. In early studies on the biological effects of high-intensity focused ultrasound, in which a progressive wave field was employed, the tissue damage due to so-called collapse cavitation was only observed at ultrasonic intensity an order of magnitude higher than the intensity threshold for the thermal tissue damage.[16] Some new method to efficiently generate acoustic cavitation even by progressive waves seems to be necessary to make sonodynamic treatment clinically useful.

Microbubbles of acoustic cavitation in liquids or biological tissues are oscillators with such high nonlinearity that they radiate not only the integral harmonics but also the fractional harmonics of the insonated ultrasound.[17–19] Therefore, it may be natural to expect that simultaneous insonation at harmonically related multiple frequencies might modify acoustic cavitation.[20–26] In this paper, *in vitro* and *in vivo* effects of simultaneous insonation of the fundamental and the second harmonic on acoustic cavitation in a progressive wave field was studied. The synergistic effects between the fundamental and the second harmonic on sonochemical reaction and on the threshold for inducing tissue damage were

[a]Electronic mail: umemura@harl.hitachi.co.jp

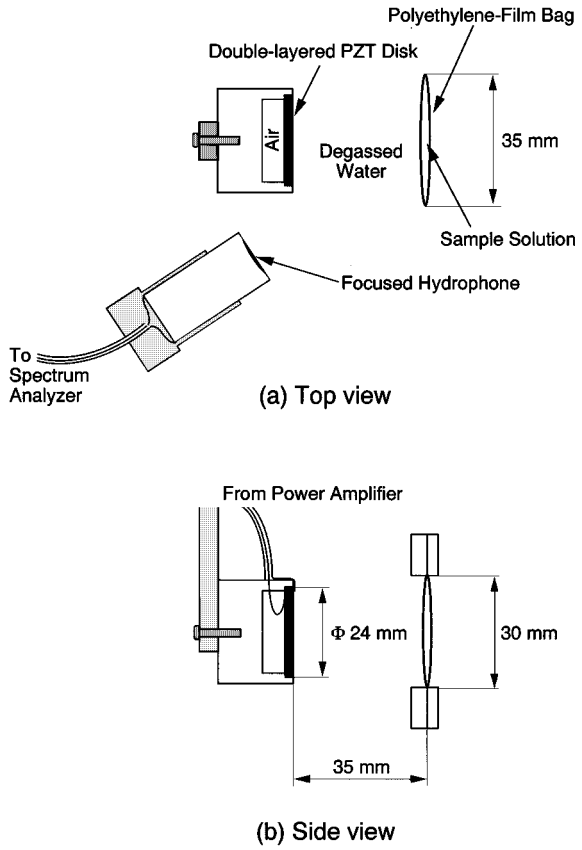


FIG. 1. *In vitro* insonation setup with plane-wave transducer. Cross-sectional (a) top view and (b) sideview. The double-layered transducer is piezoelectrically active at both fundamental and second-harmonic frequencies and can generate ultrasonic waves with second-harmonic superimposition. A small thin-film bag filled with sample solution is suspended in degassed water in such a way that its surface is parallel to that of the transducer and ultrasound propagates perpendicularly through it. A wideband focused hydrophone is directed so as to detect acoustic emission from the sample.

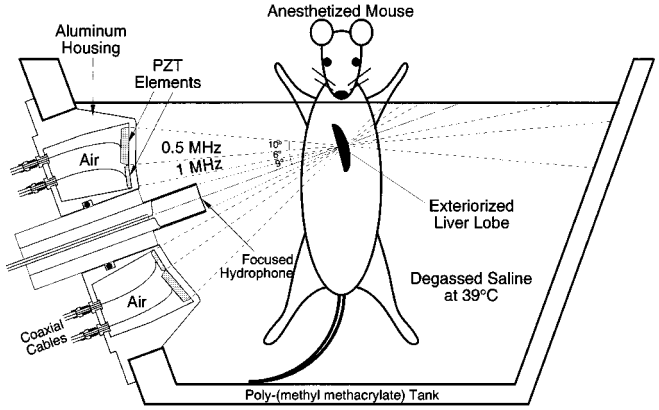


FIG. 2. *In vivo* insonation setup with focused array transducer. The fundamental and the second harmonic are separately generated by the co-focally aligned piezoelectric elements, respectively, on the outer and inner tracks of the array transducer, and then superimposed onto each other in the focal zone. An exteriorized mouse liver is placed at the focus in degassed saline in such a way that its surface is parallel to the focal plane and the focused ultrasound propagates perpendicularly through it. A wideband focused hydrophone is co-focally located in the central hole of the transducer so as to detect acoustic emission from the focal zone.

investigated in *in vitro* and *in vivo* insonation experiments, respectively. Acoustic emission was simultaneously detected during both insonation experiments and its spectral components were analyzed since its fractional harmonic components such as subharmonics or higher-order subharmonics are considered to be direct evidence of cavitation.[17–19]

I. MATERIALS AND METHODS

A. Ultrasonic transducer for second-harmonic superimposition

An acoustic field with second-harmonic superimposition can be described by

$$A(\mathbf{r},t)=a_1(\mathbf{r})\cos(\omega t+\Phi_1(\mathbf{r})) + a_2(\mathbf{r})\cos[2(\omega t+\Phi_1(\mathbf{r}))+\Phi_2(\mathbf{r})], \quad (1)$$

where $A(\mathbf{r},t)$ is the acoustic pressure at point $\mathbf{r}$ at time t, and $a_1(\mathbf{r})$ and $a_2(\mathbf{r})$ are the amplitude of the fundamental and the second harmonic at point $\mathbf{r}$, respectively; ω is the angular frequency of the fundamental; $\Phi_1(\mathbf{r})$ the phase of the fundamental at point $\mathbf{r}$; and $\Phi_2(\mathbf{r})$ the phase of the second harmonic relative to the fundamental. In order to synthesize a plane-wave field with second-harmonic superimposition, a disk-type ultrasonic transducer piezoelectrically active at both fundamental and second harmonic was devised. Since the field produced thereby has a uniform distribution of second-harmonic phase $\Phi_2(\mathbf{r})$ throughout a relatively large volume, it was used for *in vitro* sonochemical experiments. In order to synthesize a focal field with second-harmonic superimposition for *in vivo* experiments, a focused array transducer with a co-focal alignment of its elements at the fundamental and those at the second harmonic was devised.

The plane-wave transducer consisted of two layers of PZT (C-213, Fuji Ceramics, Japan): a piezoelectrically active layer with two thirds of the total thickness and a piezoelectrically inactive layer with a third of the total thickness. This design makes its electromechanical coupling for the second harmonic as good as that for the fundamental. The air-backed transducer was 24 mm in diameter and had its fundamental resonance at 1.03 MHz. Its cross sections are shown in Fig. 1. It was driven with an rf power amplifier (2100L, ENI, New York). The input signal to the amplifier was generated by an arbitrary waveform generator (AG4100, Yokogawa, Japan) operating at a clock frequency of 400 MHz.

A cross section of the focused array transducer is shown in Fig. 2. The array consisted of two tracks and 16 sectors. The fundamental and second-harmonic acoustic waves were generated from piezoelectric elements on the outer and inner tracks, respectively. Here, 0.5 MHz was chosen for the fundamental frequency, $\omega/2\pi$. The air-backed PZT elements (C-213, Fuji Ceramics) were tightly bonded with a conductive epoxy adhesive (Cho-bond 584-29, Chomerics, Massachusetts) onto the machined flats formed on the back surface of a spherical aluminum shell. The shell is 100 mm in diameter and has a spherical radius of curvature of 108 mm for geometric focusing. The thickness of the shell was adjusted to half a wavelength at 0.5 MHz, which equals a full wavelength at 1 MHz.

Each transducer element was driven by a high-voltage

amplifier consisting of a complementary pair of power MOSFET's. The phase of the input signal to each amplifier was computer controlled by digitizing based on a clock at 8 MHz. The phase of the second harmonic relative to the fundamental $\Phi_2(\mathbf{r})$ could thereby be controlled with the digitizing step of $\pi/4$. In order to control the acoustic amplitudes $a_1(\mathbf{r})$ and $a_2(\mathbf{r})$ independently of each other, the amplifier units were constructed so that their supply voltages to the power MOSFET amplifiers were separately adjustable for the outer and inner tracks. The detailed design of the circuitry is described elsewhere.[27]

The focused transducer has an axial hole 32 mm in diameter. A wideband focused hydrophone (Toray Techno, Japan) was co-focally located in the hole so as to detect acoustic emission from the focal zone. The hydrophone consisted of an 11-mm diameter, 90-μm-thick poly-(vinyledene difluoride-trifluoro ethylene) film having a spherical curvature 80 mm in radius with a 25-mm thick poly-(methyl methacrylate) (PMMA) backing. The closeness in acoustic impedance between the piezoelectric film and the backing material made the frequency response of the hydrophone wide enough to detect acoustic emissions in a 0.1- to 10-MHz range. An identical hydrophone was also used to detect acoustic emission in the *in vitro* experiments. Its configuration was shown in Fig. 1. These hydrophones were used without calibrating their sensitivity.

B. Acoustic field measurement

The distribution of the acoustic pressure at each frequency, $a_n(\mathbf{r})$ $(n=1,2)$, produced with each transducer on the plane where a specimen should be placed, was measured in degassed water with a 0.5-mm-diam needle-type poly-(vinyledene difluoride) (PVDF) hydrophone (Imotec, Germany), directed to the transducer. For the plane-wave transducer, the hydrophone was scanned on a plane 35 mm away from the transducer surface in the range of 40 mm$\times$40 mm square, with a 0.4-mm mesh separation. For the focused array transducer, it was scanned on the focal plane in the range of 20 mm$\times$20 mm square, with a 0.2-mm mesh separation. The phase of the second-harmonic relative to the fundamental, $\Phi_2(\mathbf{r})$, was calibrated by measuring the second-harmonic superimposed waveform $A(\mathbf{r},t)$ with the same hydrophone. The waveform was then fit with Eq. (1) by a least-squares method. In order to minimize the effects of nonlinear propagation and the possibility of damage to the hydrophone, these measurements were performed at a low acoustic pressure level. The generation of harmonics due to nonlinear propagation was checked with a spectrum analyzer (3588A, Hewlett–Packard, California) at an actual intensity level used in the *in vitro* and *in vivo* experiments.

The absolute acoustic power at each frequency P_n, defined as

$$P_n = \int a_n(\mathbf{r})^2/(2\rho c)\, d^2\mathbf{r} \quad (n=1,2) \tag{2}$$

was calibrated for each transducer by measuring the radiation force on a 2-mm-thick hollow aluminum plate in degassed water. Here, the integration is on a surface perpendicular to the direction of wave propagation, and ρc is the acoustic impedance of the water. The axis of the transducer was set horizontally and the plate was suspended in front of the transducer at a certain angle by four thin threads equal in length. The other ends of the threads were held with a bracket parallel to the plate so that the plate was kept at the constant angle while it moves along the transducer axis. Instead of measuring the deflection of the plate, the position of the bracket holding the other ends of the four threads was adjusted so that the plate was kept at the same position. The radiation force was then calculated from the coordinates of the adjusted bracket position. For the plane-wave transducer, a hollow plate with an area of 40 mm$\times$70 mm was used at an angle of 35° to the axis. For the focused array transducer, a hollow plate with an area of 20 mm$\times$28 mm was suspended in the focal zone at an angle of 45° to the axis. In both cases, the horizontal projection to the plane of specimen is the same as the area of the acoustic pressure distribution measurement.

The magnitude of the acoustic intensity distribution, $a_n(\mathbf{r})^2/(2\rho c)$ $(n=1,2)$, on the plane of specimen was calibrated at each frequency by comparing the measured absolute acoustic power and the surface integral of the square of the hydrophone output voltage. The spatial average intensity I_{S_n} $(n=1,2)$, defined as

$$I_{S_n} = \int_{S_n} a_n(\mathbf{r})^2/(2\rho c)\, d^2\mathbf{r}/S_n \quad (n=1,2) \tag{3}$$

was then calculated at each frequency, where S_1 and S_2 denote the areas of surface integration for the fundamental and second harmonic intensities, respectively. For the plane-wave transducer, a circular area approximately corresponding to the -3 dB areas at both frequencies was chosen for both S_1 and S_2, and the spatial average intensity was calculated as a function of the drive signal amplitude at each frequency. For the focused array transducer, the focal spot area at each frequency was chosen for S_n, and the focal-spot average intensity was calculated as a function of the supply voltage to the power MOSFET amplifiers. Here, it was assumed that the intensity distribution at each frequency keeps its shape even in a higher voltage range.

C. Chemicals

Potassium iodide (KI) and chloral hydrate ($CCl_3CH(OH)_2$) were purchased from Kanto Chemical (Japan) for *in vitro* experiments. Sodium salt of erythrosine B (erythrosine), 2′,4′,5′,7′-tetraiodofluorescein disodium salt ($C_{20}H_6I_4Na_2O_5\cdot H_2O$) was purchased from Wako Junyaku Kogyo (Japan) for *in vivo* experiments. All other reagents were commercial products of analytical grade.

D. *In vitro* experiment

To evaluate the effect of second-harmonic superimposition on inducing acoustic cavitation and sonochemical reaction in a progressive wave field, a small thin-film bag was filled with sample aqueous solution, suspended 35 mm away from the surface of the plane-wave transducer in degassed water, and insonated at room temperature. A 30 mm$\times$35 mm

bag of 0.03-mm-thick polyethylene film was filled with a 1.75-ml aliquot of air-saturated phosphate buffered saline (*p*H 7.4) containing 0.1 M KI and 0.66 M chloral hydrate, resulting in an overall thickness of approximately 2 mm. This bag was suspended in such a way that its surface was parallel to that of the transducer and ultrasound propagated perpendicularly through it as shown in Fig. 1. The separation between the plastic frames suspending the bag was chosen to be large enough for their acoustic interference with the ultrasonic beam to be ignored.

When sonochemically effective cavitation takes place in the sample solution, iodide ions are oxidized to form molecular iodine. The amount of iodine produced was measured spectrophotometrically at 350 nm after insonation for 1 min. Acoustic emission from the focal zone was simultaneously detected by the focused hydrophone and the signal was processed by a spectrum analyzer (3588A, Hewlett–Packard). The spectral signal of acoustic emission was captured by a PC/AT via a GP-IB. Insonation experiments were performed three times at each condition and the average was calculated.

The water outside the polyethylene bag was degassed not only before being put in the insonation tank but also continuously during the experiment by a degasser unit (ERC-3502W, Erma Cr., Japan) in order to prevent acoustic cavitation in the water itself during the insonation experiments. The tank water was taken through an inlet pipe, passed through a disposal capsule filter (CCG-1-C1H, Advantec, Toyo Roshi Kaisha, Japan) and the degasser, and was fed back to the tank through an outlet pipe by a roller pump. The dissolved oxygen content in the tank water was thereby kept below 2 ppm, which was verified by measuring it with an oxygen electrode. This level of oxygen content is less than 25% of that contained in air-saturated water.

E. *In vivo* experiment

The *in vivo* effect of second-harmonic superimposition was tested in the progressive wave mode. The *in vivo* specimens were also placed so that the focused ultrasound propagated perpendicularly through them without any other acoustic interference. A xanthene derivative, erythrosine B (erythrosine), was chosen for the chemical agent used in combination with the insonation.[28] The experimental animals were treated according to the guideline proposed by the Science Council of Japan.

After surgical anesthesia with sodium pentobarbital was given to a 5-week-old male ddY mouse (25–30 g), purchased from Sankyo Labo Service (Japan), a liver lobe was mobilized and exteriorized through an upper midline incision. Erythrosine was intravenously administered to the mouse at a dose of 50 mg/kg right before the incision. The mouse was held vertically in degassed saline at 39 °C with its exteriorized liver suspended with a thin suture so that the liver lobe was sticking out through the slit perpendicularly to the abdominal wall. The position of the mouse was adjusted to locate the lobe at the focal spot, and its angle was adjusted to align the lobe parallel to the focal plane. This let the focused ultrasound propagate perpendicularly through the lobe, 2–3 mm in thickness.

The insonation setup is schematically shown in Fig. 2. The water level in the insonation tank was kept below the mouth to allow easy respiration. The saline was also continuously filtered and degassed with a circulating water system in the same way as the tank water in the above described *in vitro* experiment. The outlet of degassed saline in the tank was located so that the focused ultrasound propagated through freshly degassed saline to the mouse liver lobe. Acoustic emission from the liver lobe in the focal zone was also simultaneously detected by the focused hydrophone during insonation and analyzed in the same way as in the *in vitro* experiment.

The two-dimensional intensity threshold producing hemorrhagenic focal tissue damage paired with fractional harmonic acoustic emission was investigated with three insonation experiments at each combination of the fundamental and second-harmonic intensities. The second-harmonic phase Φ_2 at the focus was adjusted to the phase which maximized the sonochemical effects in the *in vitro* experiment. The temperature rise in the liver tissue during insonation was checked by measurements with a 0.25-mm-diam sheathed Chromel–Almel thermocouple (Sukegawa Electric, Japan) inserted into the tissue.

II. RESULTS AND DISCUSSION

A. Field measurement

The results of acoustic field measurement are shown in Figs. 3 and 4 for the plane-wave transducer and the focused array transducer, respectively. The acoustic intensity distri-

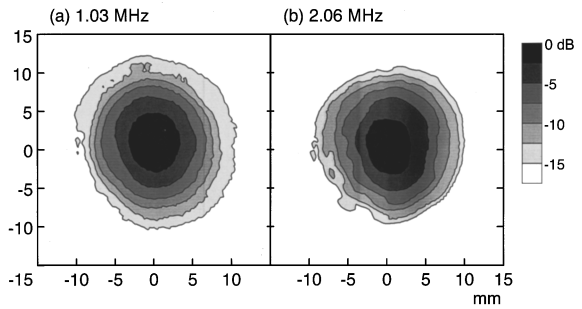


FIG. 3. Plane-wave acoustic field for second-harmonic superimposition at (a) 1.03 MHz and (b) 2.06 MHz. The ultrasonic intensity distributions on the plane 35 mm away from the transducer surface were normalized by the peak intensity at each frequency and contour-plotted.

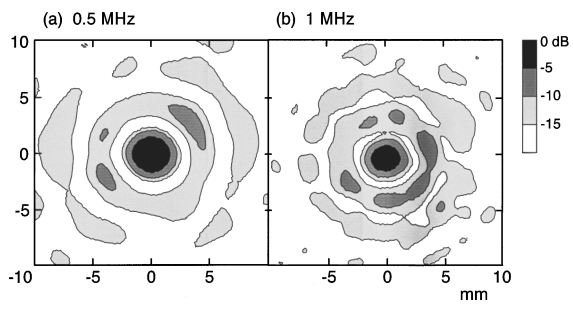


FIG. 4. Focal acoustic field for second-harmonic superimposition at (a) 0.5 MHz and (b) 1.0 MHz. The ultrasonic intensity distributions on the focal plane were normalized by the peak intensity at each frequency and contour-plotted.

butions at 1.03 and 2.06 MHz on a plane 35 mm away from the transducer surface, normalized by the peak intensity at each frequency, were contour-plotted in Fig. 3(a) and (b), respectively. The spatial average intensity I_{S_n} $(n=1,2)$, defined by Eq. (3), was calculated by choosing a circular area 6.3 mm in diameter for both S_1 and S_2. It was approximately 0.7 times the peak intensity at each frequency. The second-harmonic phase $\Phi_2(\mathbf{r})$ was within the average $\pm\pi/8$ throughout more than 99% of the area of 10 mm×10 mm square. The change in $\Phi_2(\mathbf{r})$ was also within $\pm\pi/8$ as the needle-type hydrophone was moved along the axis by 3 mm. The spatial average intensities and the spatial average second-harmonic phase Φ_2 are quoted in descriptions of *in vitro* experimental results.

The acoustic intensity distributions on the focal plane at 0.5 and 1.0 MHz, normalized by the peak intensity at each frequency, were contour-plotted in Fig. 4(a) and (b), respectively. The sidelobe had a level approximately 10 dB lower than the focal spot at each frequency. The focal spot diameter, determined from the −15-dB contours surrounding the peak, was 5.6 mm at 0.5 MHz and 4.8 mm at 1 MHz. The resulting focal spots had similar sizes owing to the design in which the ratios of the track diameter versus the wavelength for the fundamental and the second harmonic were chosen to be approximately equal. They overlay each other well enough to test the effect of second-harmonic superimposition. The focal-spot average intensity I_{S_n} $(n=1,2)$, defined by Eq. (3) was calculated at each frequency. It was 0.26 and 0.22 times the peak intensity at 0.5 and 1 MHz, respectively. Focal-spot average intensities are quoted in descriptions of *in vivo* experimental results. The second-harmonic phase $\Phi_2(\mathbf{r})$ was approximately uniform both throughout the focal spot and throughout the second lobe, respectively, but the phase in the second lobe was opposite to that in the focal spot. The phase $\Phi_2(\mathbf{r})$ in the focal spot changed about $\pi/8$ as the needle-type hydrophone was moved along the axis by 3 mm.

The generation of harmonics due to nonlinear propagation was checked at an actual intensity level used in the *in vitro* and *in vivo* experiments. The second harmonic had the highest spectral strength among the harmonics in the signal, detected at the focus with the needle-type hydrophone in the *in vivo* experimental setup of insonation at a single frequency. However, it was more than 20 dB lower than that of the fundamental even at a focal-spot average intensity of 24 W/cm^2. In the *in vitro* experimental setup of insonation at a single frequency, the second-harmonic level was even lower. These levels of second-harmonic intensity are more than an order of magnitude lower than that of the second harmonic which was deliberately superimposed onto the fundamental in the *in vitro* and *in vivo* experiments.

B. *In vitro* experiment

A preliminary *in vitro* experiment was performed to determine the optimal phase relation between the fundamental and the second harmonic for obtaining the effect of the second-harmonic superimposition. The obtained iodine production rate is plotted in Fig. 5 as a function of the spatial average second-harmonic phase Φ_2. The spatial average intensity I_{S_n} at each frequency was 2.3 W/cm^2 and kept constant. The normalized intensity of simultaneously detected subharmonic emission at 0.515 MHz is also plotted. The intensity was averaged for 1 min and then normalized to the maximum. The acoustic pressure waveforms corresponding to $\Phi_2=\pi/4$, $(3/4)\pi$, $(5/4)\pi$, and $(7/4)\pi$ are schematically shown in Fig. 5, too.

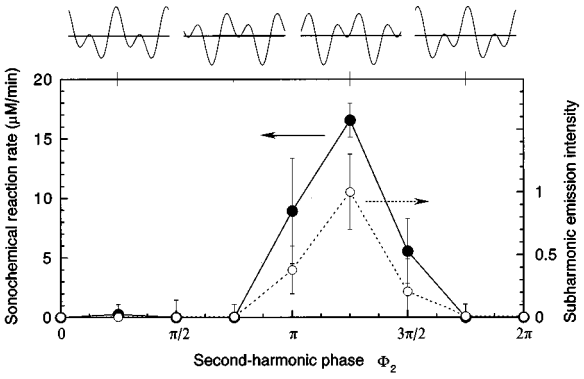


FIG. 5. Iodine production rate and subharmonic emission intensity as a function of the spatial-average second-harmonic phase Φ_2. Here, the acoustic pressure A at point $\mathbf{r}$ at time t is written as $A(\mathbf{r},t)=a_1(\mathbf{r})\cos(\omega t+\Phi_1(\mathbf{r}))+a_2(\mathbf{r})\cos[2(\omega t+\Phi_1(\mathbf{r}))+\Phi_2(\mathbf{r})]$, where a_1 and a_2 were chosen so that the spatial average intensity at each frequency was 2.3 W/cm^2. Each point and vertical bar represent the average ± standard deviation of three insonation experiments. The intensity of subharmonic emission at 0.515 MHz was normalized to the maximum. The acoustic pressure waveforms corresponding to $\Phi_2=\pi/4$, $(3/4)\pi$, $(5/4)\pi$, and $(7/4)\pi$ are schematically shown, too.

The subharmonic emission intensity and the sonochemical reaction rate correlated well with each other and both were maximized at the same second-harmonic phase, $\Phi_2=(5/4)\pi$. This second-harmonic phase was used in the following *in vitro* and *in vivo* experiments. This approximately equals the phase that maximizes the peak rarefaction $(\Phi_2=\pi)$. In more detail, the optimum second-harmonic phase seemed to be advanced from the maximum rarefaction phase by a small angle within $\pi/4$.

The iodine production rate and the normalized intensity of simultaneously detected subharmonic emission at 0.515 MHz are contour-plotted in Fig. 6(a) and (b), as a function of the fundamental and the second-harmonic spatial average intensities, I_{S_1} and I_{S_2}, on the horizontal and the vertical axes, respectively. Three insonation experiments were made every 1.25 W/cm^2 of the fundamental and second-harmonic intensities up to 5 W/cm^2. Subharmonic emission intensity was averaged for 1 min and for three times, and then normalized so that the maximum intensity equals to 100. Synergy between the fundamental and the second harmonic is clearly seen in both plots. This was best demonstrated by the contours at relatively low iodine production rates and subharmonic emission intensities. The iodine production rate of 5 μM/min was obtained at a total acoustic intensity of approximately 2 W/cm^2 with second-harmonic superimposition, while at least 4 W/cm^2 was needed when only one frequency was used.

The sonochemical reaction rate and the acoustic emission strength correlate well not only in the phase dependence (Fig. 5) but also in the intensity dependence (Fig. 6). Similar

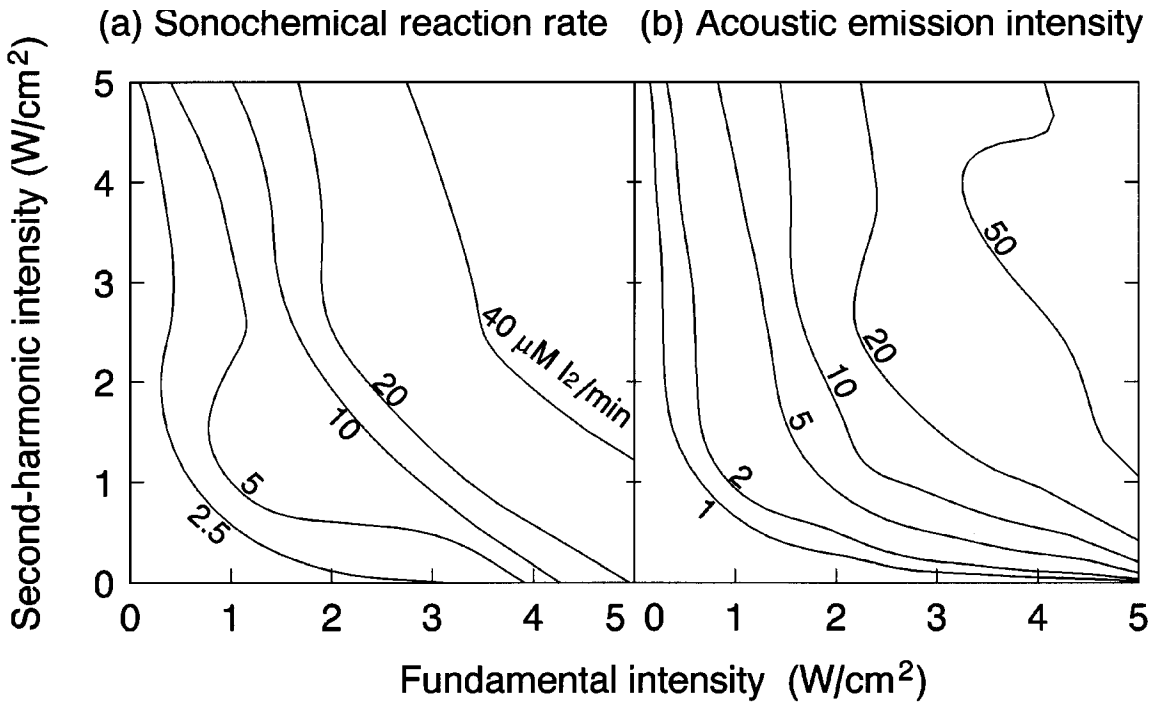


FIG. 6. *In vitro* enhancement of sonochemically active cavitation by second-harmonic superimposition. (a) Iodine production rate and (b) subharmonic emission intensity at 0.515 MHz are contour-plotted as a function of the spatial average intensities at 1.03 and 2.06 MHz, I_{S_1} and I_{S_2}, on the horizontal and the vertical axes, respectively. Focal-spot average intensities at 0.5 and 1 MHz. Subharmonic emission intensity was averaged for 1 min and then normalized so that the maximum intensity equals to 100.

dependency was also observed for acoustic emission at other fractional harmonic frequencies such as 1.585 or 2.615 MHz. These results indicate that the second-harmonic superimposition accelerates sonochemical reactions by enhancing acoustic cavitation. Theoretically,[23,26] the second-harmonic phase maximizing the peak rarefaction is thought to give the highest bubble growth rate due to rectified diffusion when the bubble is much smaller than the resonant size. When it is resonant at the second-harmonic frequency, the optimum phase for bubble growth should be advanced by $\pi/2$ since the bubble response is retarded by $\pi/2$ at the resonant frequency.[15] Since the observed optimum phase was between those of the above two cases, it is reasonable to think that the second-harmonic superimposition mainly enhances the bubble growth rather than its collapse to cause enhancement of cavitation. If its collapse were the process being enhanced, the optimum phase would have been close to the phase maximizing the peak positive pressure.

C. *In vivo* experiment

A typical gross pathology of mouse liver tissue damage produced by focused ultrasound with second-harmonic superimposition is shown with a ruler in millimeters in Fig. 7(a). This tissue sample was excised and fixed in 10% formalin immediately after insonation, and it was later stained with hematoxylin and eosin (H & E) for histological examination. Its histologic section was shown in Fig. 7(b). The mouse liver was insonated at focal-spot average acoustic intensities of 2 W/cm^2 at both 0.5 and 1 MHz after intravenous administration of erythrosine. The hemorrhagenic region, 3–4 mm in diameter, approximately matches the focal spots

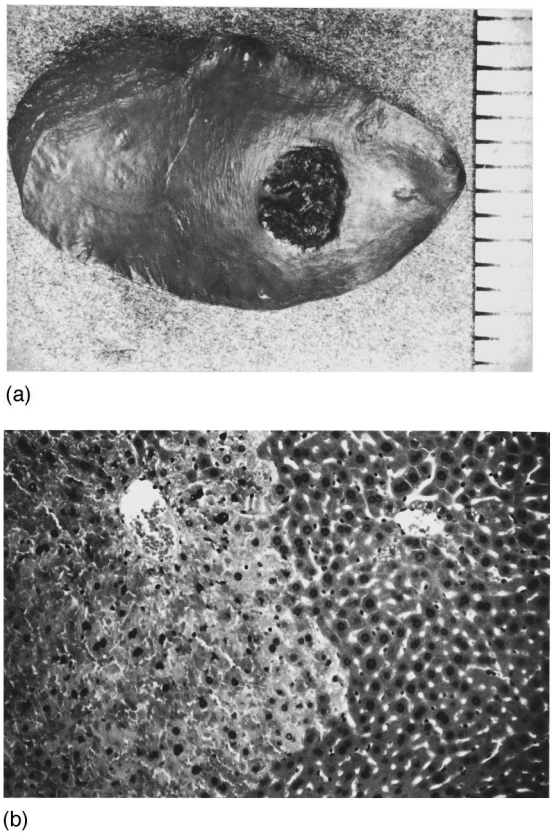


FIG. 7. Typical tissue damage produced by focused ultrasound with second-harmonic superimposition. (a) A gross pathology and (b) a histologic section of a damaged mouse liver tissue are shown. The ruler in (a) is in millimeters. The mouse liver was insonated at focal-spot average acoustic intensities of 2 W/cm^2 at both 0.5 and 1 MHz after intravenous administration of erythrosine. Histologic region on the left seen in (b) corresponds to the hemorrhagenic region.

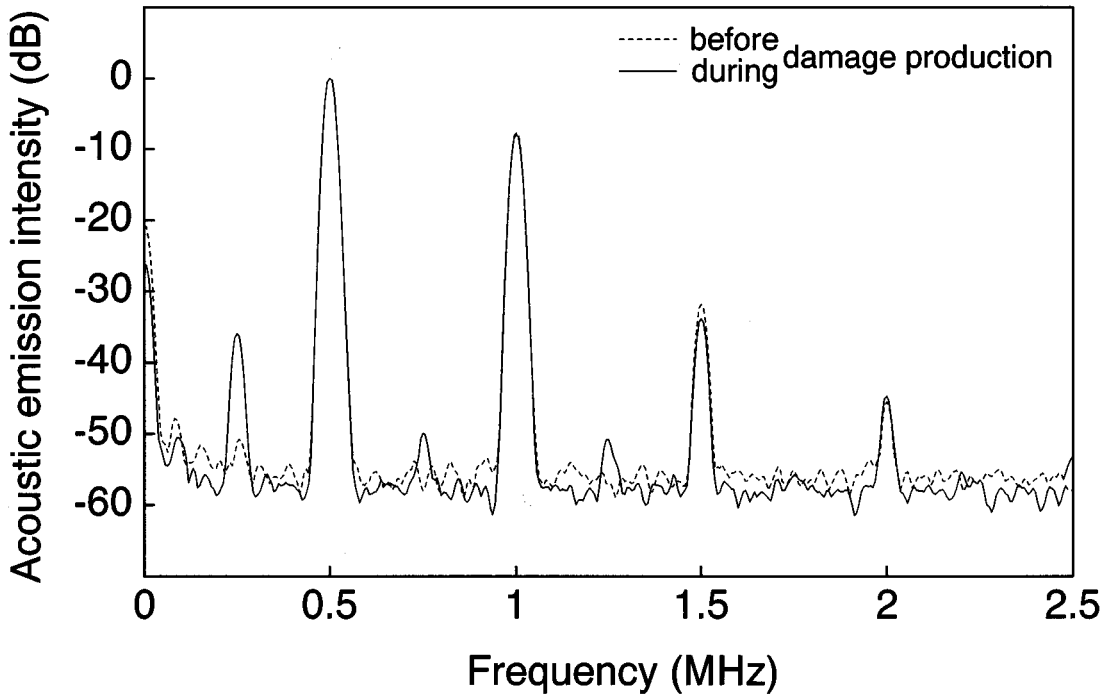


FIG. 8. Acoustic emission paired with hemorrhagic tissue damage produced by second-harmonic superimposition. The mouse liver was insonated at focal-spot average acoustic intensities of 1 W/cm^2 at both 0.5 and 1 MHz after intravenous administration of erythrosine. The power spectra of acoustic emissions, right before and during the damage production, were averaged for 5 s, normalized to the maximum, and plotted with dotted and solid lines, respectively.

of both fundamental and second-harmonic. Histology on both sides of the boundary of the hemorrhagenic region is seen in Fig. 7(b). The histologic region on the left, which corresponds to the hemorrhagenic region, shows a high erythrocyte concentration and a high proportion of hepatocyte necrosis. The region on the right, in contrast, shows no significant change.

Such hemorrhagenic tissue damage was always paired with fractional harmonic emission, which was detected while the damage was being produced but not before its production started. Typical acoustic emission spectra obtained at focal-spot average acoustic intensities of 1 W/cm^2 at both frequencies with erythrosine are shown in Fig. 8. The dotted and solid lines, respectively, denote the 5-s average power spectra of acoustic emissions right before and during the production of focal tissue damage. Fractional harmonic components at 0.25, 0.75, and 1.25 MHz are clearly seen in the spectrum during the production but not in the spectrum before. Since fractional harmonic emission induced by ultrasound is known to be specific to acoustic cavitation,[12] the observed hemorrhagenic focal tissue damage was likely to be primarily due to the cavitational rather than thermal effect of ultrasound. When a liver lobe was not located in the focal zone, no fractional harmonic emission above the noise level was detected up to a total focal-spot intensity of at least 30 W/cm^2.

The temperature rise in the liver tissue during insonation reached a plateau within 1 min and was less than 2 °C even at a focal-spot average intensity of 24 W/cm^2 at 1 MHz with the thermocouple tip located in the middle of the liver lobe thickness and at the focus. It was approximately twice as high for 1 as for 0.5 MHz at the same acoustic intensity with the same thermocouple location. This small temperature rise was primarily due to the degassed saline at constant temperature, cooling the thin mouse liver lobe from both sides. This level of temperature rise is not likely to induce thermal damage of tissues during such a short period of time.[29]

The effect of intensity combinations on producing cavitational tissue damage with and without administration of erythrosine are shown in Fig. 9(a) and (b), respectively. The results are plotted with the fundamental and the second-harmonic focal-spot average intensities, I_{S_1} and I_{S_2}, on the horizontal and vertical axes, respectively. Note that the scales are magnified twice in Fig. 9(a) than in (b). Each point represents the result of three mice. Solid and open circles correspond to intensity combinations with which focal tissue damage was produced in all and none of the three mice, respectively. Circles, two thirds and a third of which are filled, correspond to combinations with which focal tissue damage was produced in two and one of the three mice but not in the others, respectively. Each insonation was continued for a maximum of 3 min until tissue damage was observed. In Fig. 9(a), some results without erythrosine were also plotted with triangles for comparison.

Synergism between the fundamental and the second harmonic in producing cavitational tissue damage is quite distinctive in the presence of erythrosine [Fig. 9(a)]. Cavitational tissue damage was observed when focal-spot average acoustic intensities at both frequencies was 1 W/cm^2 or higher, while a focal-spot average intensity of at least 4 W/cm^2 was needed for inducing such tissue damage when only one frequency was used. In the absence of erythrosine [Fig. 9(b)], the observed synergism between the fundamental and the second harmonic was relatively slight.

Single-frequency insonation at a focal-spot average in-

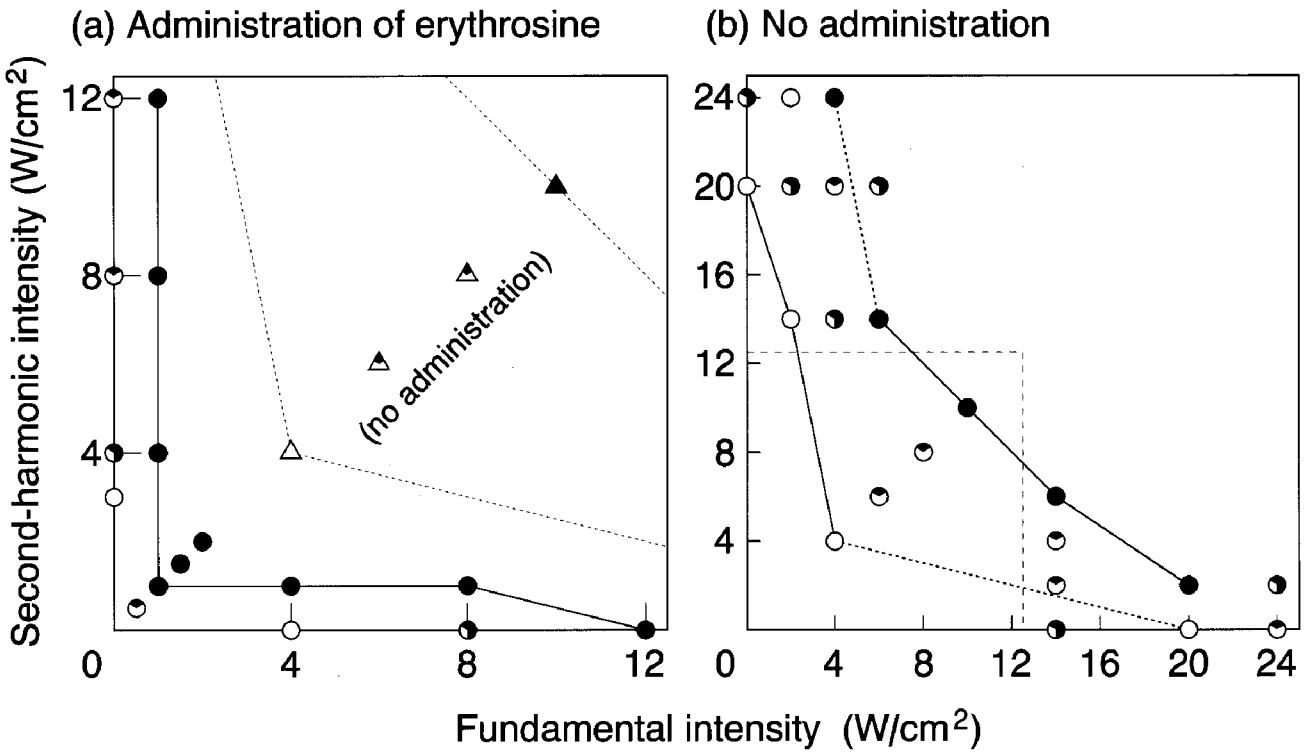


FIG. 9. *In vivo* enhancement of cavitational tissue damage production by second-harmonic superimposition. The results (a) with and (b) without administration of erythrosine (50 mg/kg) are plotted with the fundamental and the second-harmonic focal-spot average intensity on each axis. The scales are magnified twice in (a) relative to (b). Each point represents the result of three mice. Solid and open circles correspond to intensity combinations with which focal tissue damage was produced in all and none of the three mice, respectively. Circles, two thirds and a third of which are filled, correspond to combinations with which focal tissue damage was produced in two and one of the three mice but not in the others, respectively. In (a), some results without erythrosine were also plotted with triangles for comparison.

tensity of 20 W/cm^2 did not produce focal tissue damage without administration of erythrosine [Fig. 9(a)]. This intensity level corresponds to spatial peak intensity of 77 and 92 W/cm^2 at 0.5 and 1 MHz, respectively. These are consistent with earlier reports on focal lesion formation, in which the spatial peak intensity threshold for liver tissues was higher than a few hundred W/cm^2.[30]

Martin *et al.* reported[19] that they observed cavitational damage in mouse liver tissue at 0.8 MHz at a focal-spot average intensity of a few W/cm^2, which was an order of magnitude lower than the threshold observed in our study. They attributed the low intensity threshold to the initiation of cavitation within the saline coupling medium rather than in the tissue. A much higher cavitation threshold in the coupling medium was achieved in our experiments by degassing the saline continuously and by configuring the insonation to restrain standing waves from being generated near the liver lobe. Since erythrosine has been reported to reduce *in vitro* cavitation threshold in aqueous solutions,[28] the observed distinctive difference between the *in vivo* experimental results with and without administration of erythrosine can be explained much more easily by attributing it to cavitational events in the tissue rather than in the saline.

The above described *in vivo* experiment was not intended to determine the optimum second-harmonic phase for inducing cavitational tissue damage. Further studies are needed to find out whether it is similar to the *in vitro* optimum phase. Carstensen and his co-workers investigated the effect of asymmetric, distorted waveform on inducing cavitational bioeffects.[31] Although the fundamental and the second harmonic were the two predominate components of the distorted waveform, only a particular second-harmonic phase arising from waveform distortion due to nonlinear propagation of finite amplitude ultrasound was tested in their experiments. Furthermore, the pulse length they used was about ten wavelengths, which might be too short for bubbles to grow by rectified diffusion. These might help to explain why the distorted pulse did not result in a higher killing ratio than the sinusoidal pulse at the same acoustic intensity in their experiments.

D. Acoustic field design

The transducer used in this paper was designed so that the second harmonic was superimposed on the fundamental selectively in the focal zone as shown in Fig. 2. This design has two advantages. First, the observed synergism can take place only in the focal zone, neither in front of nor beyond it. This longitudinal spatial selectivity can be even better if the *in vivo* effect of second-harmonic superimposition is as phase sensitive as the *in vitro* effect. Second, the effect of distortion due to nonlinear propagation on the superimposed waveform in the focal zone can be minimized thereby if it is significant for therapeutic insonation at higher intensity levels.

These may provide good control for the sonodynamic

treatment employing second-harmonic superimposition and prevent potential side effects which might occur outside the treatment region. Further experiments using thick tissues are needed to confirm the longitudinal spatial selectivity of this approach. The liver lobes used in the *in vivo* experiment in this paper were too thin for this purpose. In the experiments using thick tissues, the temperature rise in the tissue may no longer be small enough to ignore.[32,33]

III. CONCLUSION

It was found that acoustic cavitation can be an order of magnitude enhanced by superimposing the second harmonic on the fundamental. Significant synergistic effects between the fundamental and the second harmonic were observed in both *in vitro* and *in vivo* experiments using a progressive wave field. A significant effect of second-harmonic superimposition was observed on *in vitro* sonochemical reaction and on the fractional harmonic emission simultaneously detected. The intensity threshold for *in vivo* production of focal tissue damage paired with fractional harmonic emission was significantly lowered by second-harmonic superimposition especially with intravenous administration of erythrosine, a sonodynamically active xantine dye. Insonation with second-harmonic superimposition in combination with such a sonodynamically active agent may have potential use for selective tumor treatment.

ACKNOWLEDGMENTS

We thank Dr, K. Yasuda, S. Matsunami, N. Moriya, H. Kawaguchi, and K. Shindo of Hitachi for cooperation in designing and constructing the experimental systems. The useful discussions with Dr. K. Umemura, Dr. R. Nishigaki and Dr. N. Yumita of Toho University are also highly appreciated.

[1] P. P. Lele, ''Local hyperthermia by ultrasound,'' in *Physical Aspect of Hyperthermia*, edited by G. H. Nussbaum (Am. Inst. Phys., New York, 1982), pp. 393–440.
[2] F. J. Fry, G. Kossoff, R. C. Eggleton, and F. Dunn, ''Threshold ultrasonic dosages for structural changes in the mammalian brain,'' J. Acoust. Soc. Am. **48**, 1413–1417 (1970).
[3] L. Chen, I. Rivens, G. ter Harr, S. Riddler, C. R. Hill, and J. P. M. Bensted, ''Histological changes in rat liver tumours treated with high-intensity focused ultrasound,'' Ultrasound Med. Biol. **19**, 67–74 (1993).
[4] F. W. Kremkau, J. S. Kaufmann, M. M. Walker, P. G. Burch, and C. L. Spurr, ''Ultrasonic enhancement of nitrogen mustard cytotoxicity in mouse leukemia,'' Cancer **47**, 1643–1647 (1976).
[5] R. Akimoto, ''An experimental study on enhancement of the effect of anticancer drug by ultrasound,'' J. Jpn. Soc. Cancer Therapy **20**, 562–570 (1985).
[6] N. Yumita, A. Okumura, R. Nishigaki, K. Umemura, and S. Umemura, ''The combination treatment of ultrasound and anti-tumor drugs on Yoshida sarcoma,'' Jpn. J. Hypeerthermic Oncol. **3**, 175–182 (1987).
[7] N. Yumita, R. Nishigaki, K. Umemura, and S. Umemura, ''Synergetic effect of ultrasound and hematoporphyrin on sarcoma 180,'' Jpn. J. Cancer Res. **81**, 304–308 (1990).
[8] G. H. Harrison, E. K. Balcer-Kubiczek, and H. A. Eddy, ''Potentiation of chemotherapy by low-level ultrasound,'' Int. J. Radiat. Biol. **59**, 1453–1466 (1991).
[9] K. Umemura, N. Yumita, R. Nishigaki, and S. Umemura, ''Sonodynamically induced antitumor effect of pheophorbide a,'' Cancer Lett. **102**, 151–157 (1996).
[10] N. Yumita, K. Sasaki, S. Umemura, and R. Nishigaki, ''Sonodynamically induced antitumor effect of a gallium-porphyrin complex, ATX-70,'' Jpn. J. Cancer Res. **87**, 310–316 (1996).
[11] K. S. Suslick, ''Sonochemistry,'' Science **247**, 1439–1445 (1990).
[12] N. Yumita, R. Nishigaki, K. Umemura, and S. Umemura, ''Hematoporphyrin as a sensitizer of cell damaging effect of ultrasound,'' Jpn. J. Cancer Res. **89**, 219–222 (1989).
[13] S. Umemura, N. Yumita, and R. Nishigaki, ''Enhancement of ultrasonically induced cell damage by a gallium-porphyrin complex, ATX70,'' Jpn. J. Cancer Res. **84**, 582–588 (1993).
[14] S. Umemura, N. Yumita, R. Nishigaki, and K. Umemura, ''Sonochemical activation of hematoporphyrin: a potential modality for cancer treatment,'' *Proc. 1989 IEEE Ultrasonics Symposium* (IEEE, New York, 1989), pp. 955–960.
[15] A. Henglein, ''Sonochemistry: historical development and modern aspects,'' Ultrasonics **25**, 6–16 (1987).
[16] P. P. Lele, ''Effects of ultrasound on 'solid' mammalian tissues and tumors *in vivo*,'' in *Ultrasound: Medical Applications, Biological Effects and Hazard Potential*, edited by M. H. Repacholi, M. Grandolfo, and A. Rindi (Plenum, New York, 1986), pp. 275–306.
[17] C. R. Hill, ''Ultrasonic exposure threshold for changes in cells and tissues,'' J. Acoust. Soc. Am. **52**, 667–672 (1971).
[18] K. I. Morton, G. R. ter Haar, I. J. Stratford, and C. R. Hill, ''Subharmonic emission as an indicator of ultrasonically-induced biological damage,'' Ultrasound Med. Biol. **9**, 629–633 (1983).
[19] C. J. Martin, D. W. Gregory, and M. Hodgkiss, ''The effects of ultrasound *in vivo* on mouse liver in contact with an aqueous coupling medium,'' Ultrasound Med. Biol. **7**, 253–265 (1981).
[20] S. Umemura, K. Kawabata, N. Yumita, R. Nishigaki, and K. Umemura, ''Sonodynamic approach to tumor treatment,'' *Proc. 1992 IEEE Ultrasonics Symposium* (IEEE, New York, 1992), pp. 1231–1240.
[21] K. Kawabata and S. Umemura, ''Study of efficient sonochemical reaction by focused ultrasound,'' *Proc. 1992 IEEE Ultrasonics Symposium* (IEEE, New York, 1992), pp. 1281–1285.
[22] S. Umemura and K. Kawabata, ''Enhancement of sonochemical reactions by second-harmonic superimposition,'' *Proc. 1993 IEEE Ultrasonics Symposium* (IEEE, New York, 1993), pp. 917–920.
[23] S. Umemura, K. Kawabata, and K. Sasaki, ''Theoretical analysis of cavitation enhancement by second-harmonic superimposition,'' *Proc. 1994 IEEE Ultrasonics Symposium* (IEEE, New York, 1994), pp. 1843–1846.
[24] S. Umemura, K. Kawabata, and K. Sasaki, ''Reduction of threshold for producing sonodynamic tissue damage by second-harmonic superimposition,'' *Proc. 1995 IEEE Ultrasonics Symposium* (IEEE, New York, 1995), pp. 1567–1570.
[25] K. Kawabata and S. Umemura, ''Effect of second-harmonic superimposition on efficient induction of sonochemical effect,'' Ultrasonics Sonochemistry **3**, 1–5 (1996).
[26] S. Umemura, K. Kawabata, and K. Sasaki, ''Enhancement of sonodynamic tissue damage production by second-harmonic superimposition: Theoretical analysis of its mechanism,'' IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control, (1996) accepted.
[27] S. Umemura and C. A. Cain, ''Acoustical evaluation of a prototype sector-vortex phased-array applicator,'' IEEE Trans. Ultrason. Ferroelectr. Frequency Control **39**, 32–38 (1992).
[28] K. Kawabata, K. Sasaki, and S. Umemura, ''Effect of functional groups of organic compounds for reducing cavitation threshold in vivo,'' *Proc. 5th Meeting European Soc. Sonochemistry* (1996), p. 42.
[29] S. A. Sapareto and W. C. Dewey, ''Thermal dose determination in cancer therapy,'' Int. J. Radiat. Oncol. Biol. Phys. **10**, 787–800 (1984).
[30] L. A. Frizzell, C. A. Linke, E. L. Carstensen, and C. W. Fridd, ''Thresholds for focal ultrasonic lesions in rabbit kidney, liver, and testicle,'' IEEE Trans. Biomed. Eng. **24**, 393–396 (1977).
[31] E. J. Ayme and E. L. Carstensen, ''Cavitation induced by asymmetric, distorted pulses of ultrasound: a biological test,'' Ultrasound Med. Biol. **15**, 61–66 (1989).
[32] F. G. Sommer and D. Pounds, ''Transient cavitation in tissues during ultrasonically induced hyperthermia,'' Med. Phys. **9**, 1–3 (1982).
[33] K. Hynynen, ''The threshold for thermally significant cavitation in dog's thigh muscle *in vivo*,'' Ultrasound Med. Biol. **17**, 157–169 (1991).

Mechanisms of vocal production in budgerigars (*Melopsittacus undulatus*)

Elizabeth F. Brittan-Powell and Robert J. Dooling
Psychology Department, University of Maryland, College Park, Maryland 20742-4411

Ole N. Larsen
Center for Sound Communication, Institute of Biology, Odense University, DK-5230 Odense, Denmark

James T. Heaton
Psychology Department, University of Maryland, College Park, Maryland 20742-4411

(Received 8 May 1996; accepted for publication 27 August 1996)

Songbirds vocalizing in helium show a change in the spectral quality of their vocalizations. This effect is due to an increase in the speed of sound in helium that in turn alters the resonance properties of the vocal tract. Here, this approach is extended to a psittacine, the budgerigar (*Melopsittacus undulatus*), whose syringeal anatomy and innervation differ from that of a songbird. Contact calls from birds vocalizing in heliox (70/30 helium/oxygen environment) showed an overall increase in the amount of energy at frequencies above the fundamental, slight changes in the frequency of the fundamental and harmonics, and some change in the level of harmonics. Calls produced by a syringeally denervated bird showed more dramatic changes. Recordings from live birds were compared with sounds produced by various simple ''artificial'' tracheal and syringeal models. Results suggest that budgerigars produce contact calls using the syringeal membranes as a unitary sound source which produces acoustic energy in a narrow frequency band whose fundamental frequency is matched to the resonant frequency of the trachea. The syrinx is not normally coupled to the tracheal resonator, and resonances probably play only a minor role in shaping the spectrum of contact calls. © *1997 Acoustical Society of America.*
[S0001-4966(97)03101-9]

PACS numbers: 43.80.Ka, 43.70.Aj, 43.72.Ar [FD]

INTRODUCTION

Sound production in birds has historically been viewed as different from that of humans and other mammals. Humans produce speech beginning with an interruption of air flow across the vocal cords (i.e., vibration of the vocal cords). The filtering properties of the vocal tract, which can be modified both by constrictions of the tract and by movement of the tongue and lips, shape the resulting sound spectrum (e.g., Rossing, 1989). The rate of vibration of the vocal cords is determined primarily by their mass and tension and normally undergoes little change during speech production. In humans, the vibrations of the vocal cords act independently of the resonant cavities of the vocal tract. The human vocal tract can be modeled fairly effectively by a single tube extending from the vocal cords to the lips. The resonant frequencies of such a tube, closed at one end by the glottis and open at the lips, are its odd harmonics. The fundamental frequency of a human voice in helium gas does not change significantly because helium does not affect the vibration frequency of the vocal cords (Beil, 1962). However, helium does affect the speed of sound thus altering the bandpass or resonance characteristics of the vocal tract. Speech produced in helium compared to air contains more energy at higher harmonics and less energy at lower harmonics with no change in frequency of the fundamental. This set of results is taken as evidence that in humans the vocal cords (the source) and the trachea (resonator) are uncoupled during speech production. Here, we test whether a similar relation holds for the budgerigar (*Melopsittacus undulatus*) syrinx and trachea.

Birds have syringes that contain specialized tympaniform membranes (TMs) that act as acoustic sources when caused to vibrate by air flow across their surface through the Bernoulli effect (Fletcher, 1992). Whether vocal tract resonances influence the characteristics of bird vocalizations has been a long standing debate in the study of bird song. For instance, analogies have often been drawn between bird phonation and sound production by woodwind instruments. These theories assume that the resonator (the trachea) is coupled to the source (syringeal membranes) as with woodwind musical instruments (e.g., Nowicki, 1987; Nowicki and Marler, 1988). Still other research has suggested a closer parallel between bird phonation and human sound production (Nowicki, 1987). In at least some songbirds, the syrinx and trachea appear to be uncoupled with the resonant properties of the vocal tract affecting the harmonic spectrum of vocalizations but not the fundamental (Nowicki, 1987; Westneat *et al.*, 1993).

Most of the previous work on bird vocalizations has been concerned with how birds produce tonal signals such as the contact calls of budgerigars which are tonal, frequency modulated signals falling in the spectral region of 2–4 kHz (Dooling, 1986). A review of the literature suggests there are essentially three classes of models about how the avian syrinx and trachea produce tonal vocalizations. The first model suggests that pure-tone whistles are obtained from harmonic

signals generated by the syringeal source (this source can be either a whistle or a vibrating membrane) with vocal tract resonances acting to filter out all but a single dominant frequency (i.e., the fundamental or sometimes the second harmonic) (e.g., Nowicki, 1987; Nowicki and Marler, 1988; Westneat *et al.,* 1993). As in human speech production, the source and the resonant tube appear to act in an uncoupled fashion in this model so that tracheal resonances have little or no effect on the behavior of the source. In this model, a bird may be able to influence the output level, but probably not the frequency, of the fundamental by changing its tracheal resonances. It can potentially do this by increasing tracheal length, partially obstructing the trachea with the tongue, or flaring the beak (Podos *et al.,* 1995; Westneat *et al.,* 1993).

A second model suggests that vocal tract resonances directly influence vibrational characteristics of syringeal membranes by constraining them to vibrate in a more nearly sinusoidal fashion and at a particular frequency. In this model, the bird controls the tonal properties of the sounds it produces by actively controlling the resonant properties of the tracheal filter by the same mechanisms described above—all of which can potentially influence the pattern of vibration of the syringeal membranes. The defining characteristics of this model are that the behavior of the syringeal membranes can be modified by tracheal resonances, and the syringeal source must be strongly coupled to the trachea.

Finally, there is a logical third possibility that, like the first model, does not require that the syringeal source and the trachea be strongly coupled. In this model, a bird first changes the spectral quality of the sounds it produces by altering the vibration pattern of the syringeal source through neuromuscular control of membrane location and/or tension. The bird then adjusts its tracheal resonances to actively ''track'' these new vibration frequencies (Nowicki, 1987). In this model, the syringeal source and the trachea are either uncoupled or weakly coupled so that they act in an independent but coordinated fashion.

What makes this issue interesting is that there is tremendous variation in vocal production among species of birds, and it is unlikely that there is one mechanism that accounts for all avian vocalizations or even that only one mechanism operates exclusively in a single species. Songbirds, for instance, have a bronchiosyrinx with one ''syrinx'' in each bronchus (four membranes), which are separately innervated, while psittacines have a tracheosyrinx consisting of two membranes at the bronchotracheal junction and a hemidescussate innervation pattern (Gaunt, 1983). Taken together, the models described above touch on a number of unresolved issues in the study of vocal production in birds including: (1) whether the trachea acts as an acoustic filter that enhances and/or suppresses part of the spectrum of the vocalizations; (2) whether the trachea functions as a tube that has both ends opened (at the syrinx and the beak), both ends closed, or only one end open (beak); (3) whether the resonances of the avian vocal tract are passive characteristics of the tube or actively manipulated by the bird; (4) whether the syringeal membranes (the vibrating sources) are coupled to the trachea (the resonator); and finally, (5) whether the two syringeal membranes act as independently vibrating sources if they are physically coupled to a same trachea.

It is worth noting that recent accounts of sound production in birds have concentrated on the effects of the tracheal tube on sound produced by the vibrating syringeal membranes (see, for example, Fletcher, 1992; Nowicki and Marler, 1988; Suthers, 1994; Westneat *et al.*, 1993). The bandpass or resonant properties of the passive trachea depend on its length and shape. Both the resonant frequency and the width of the resonance filter can affect the shape of the output spectrum. As far as we know, the resonant properties of the passive trachea have not been directly measured in birds but arguments based on indirect evidence, from analyses of songbird vocalizations, strongly point to an important role in vocal production at least in some species (see, for example, Myers, 1917; Nowicki, 1987; Sutherland and McChesney, 1965; Suthers, 1990, 1994).

Much of the previous work on the mechanisms of avian sound production has been conducted on songbirds, and some of the logic that arises from these studies is the following. In songbirds, if the two syringeal membranes can *simultaneously* produce two, harmonically unrelated tones of about the same intensity, then the bandwidth of the tracheal filter must be quite broad or only low pass with a relatively high cutoff frequency (Greenewalt, 1968). The *nonsimultaneous* production of two harmonically unrelated tones, on the other hand, could result from a trachea which acts as either a broad filter, only a low-pass filter, or as a variable filter that is capable of rapid adjustments either in frequency or bandwidth (Nowicki, 1987; Suthers, 1990). Gaunt *et al.* (1982) have shown that active frequency modulations can be entirely source generated from adjustments of the syringeal configuration through the action of the syringeal musculature.

So at least for songbirds then, the data show that some aspects of the mechanism(s) by which the bird produces tonal sounds can be teased apart from experiments with birds vocalizing in a mixture of air and helium (heliox). A songbird whose syringeal membranes and trachea are uncoupled should show no change in the frequency of the fundamental but a decrease in its relative amplitude when vocalizing in heliox. Depending on the harmonic content generated by syringeal vibration, a vocalization produced in heliox may also show an increase in amplitude of specific harmonics and an attenuation of others. Some previous work (see Nowicki and Marler, 1988 for review) has shown that the tonal quality of some bird vocalizations bears a strong resemblance to sounds produced by wind instruments—implying that the characteristics of a bird's vocal tract can affect the fundamental frequency of the song it produces (Nowicki, 1987; Nowicki and Marler, 1988). A wind instrument, such as a trombone, provides a standard model for such effects. The tube component of a trombone (which has certain bandpass or resonance properties) is coupled to the sound source (the lips). When a resonator is strongly coupled to a source, the source is forced to match its vibrations to the harmonic spectrum of the resonator (e.g., Greenewalt, 1968). Because of this coupling, the spectrum of a sound played by a trombone in helium shows an increase in the fundamental frequency in

the same proportion as the increase in the resonances of the corresponding tube. Thus, a bird whose syringeal membranes and trachea are coupled, should show the same effects as those shown by a trombone—changes in tracheal resonances should alter the vibrational pattern of the syringeal membranes thereby increasing the frequency of the fundamental.

Data from songbirds vocalizing in heliox also address the issue of whether vocalizations are solely source generated or modified by the resonances of the vocal tract (Nowicki, 1987; Nowicki and Marler, 1988; Nowicki *et al.*, 1989). Constant frequency song elements produced in air have no overtones, while the same constant frequency song elements produced in heliox show increased energy at harmonic overtones of the fundamental. These results suggest a potential role for vocal tract resonances in avian sound production. Harmonic overtones appearing in vocalizations produced in heliox, but not in air, would obtain if the bird's vocal tract normally acts as a narrow acoustic filter centered on the fundamental frequency (Nowicki, 1987).

Suthers (1994) recently extended work on acoustic resonance in avian sound production to a nonsongbird, the oilbird (*Steatornis caripensis*). In this species, anatomical differences between the left and right bronchus lead to differences in filtering properties that are realized as different formants in the oilbirds' social vocalizations. The effect of structural changes in vocal tract parameters on vocalizations in birds, in general, has not been well studied. Interestingly, others have recently reported a positive correlation between beak gape and sound frequency in the song of two species of sparrows (*Zonotrichia albicollis* and *Melospiza georgiana*) (Westneat *et al.*, 1993). This is an interesting approach since it could reveal the mechanisms by which birds might alter the resonant properties of the vocal tract and affect the spectral quality of vocal output.

In the following experiments, we sought to learn the extent to which resonances of the vocal tract influence vocalizations produced by a psittacine species, the budgerigar. The budgerigar represents an interesting addition to previous studies for several reasons. First, this species has an unusually complex vocal repertoire consisting of both tonal and broadband sounds (Dooling, 1986; Farabaugh *et al.*, 1992). Second, the budgerigars' syringeal anatomy and innervation differ considerably from that of songbirds and other nonpsittacines (see Gaunt, 1983; Heaton *et al.*, 1995; Manogue and Nottebohm, 1982; Nottebohm, 1976). Third, budgerigars learn their contact calls throughout life (Dooling, 1986; Farabaugh *et al.*, 1994). The calls are tonal, frequency modulated vocalizations that average about 150–200 ms in duration, with energy concentrated in the frequency region of 2–4 kHz (Dooling, 1986; Dooling *et al.*, 1987). Contact calls also show no apparent sexual dimorphism and yet have patterns of frequency modulation that can differ dramatically between birds. Each individual bird generally produces one to several call types, but the pattern of frequency modulation within each call type is highly stereotyped from one rendition to another (Brockway, 1969; Dooling, 1986; Wyndham, 1980). Birds typically produce contact calls when they are separated from flock mates, in flight, or preparing for the evening roost (Wyndham, 1980).

To summarize, budgerigars may control the spectral content of contact calls by at least two general mechanisms. They could produce a broadband harmonic signal at the source (syrinx) which the trachea (resonator) subsequently filters to produce a narrowband output. Moreover, the syrinx and the trachea could function in an acoustically coupled or uncoupled fashion. Alternatively, budgerigars could produce a narrow-band acoustic signal at the source which is matched to the resonant properties of the trachea. This system could also operate in an acoustically coupled or uncoupled fashion. In each case, the trachea alone could act as a tube opened at both ends, closed at both ends, or closed at one end (syrinx) and open at the other (beak). These mechanisms lead to different predictions as to the effect of heliox on the production of contact calls.

In the following experiments, we placed birds in an environment consisting of 70% helium and 30% oxygen (heliox) and examined the acoustic characteristics of contact calls in this gas mixture. We also examined the effect of heliox on simple sound producing whistles and tubes loosely modeled after the budgerigar syrinx in order to better understand the possible mechanisms budgerigars use to produce vocalizations.

I. METHODS

A. Subjects

A total of eight adult budgerigars (seven males, one female) served as subjects in this experiment. We housed the birds in an avian vivarium at the University of Maryland and kept them on a photoperiod correlated with the season. As a test of whether the vocal tract resonances are actively controlled, one subject underwent bilateral tracheosyringeal (ts) nerve resections to denervate its syrinx. Briefly, we anesthetized the bird with an intramuscular injection of ketamine hydrochloride (40 mg/kg) and xylazine hydrochloride (10 mg/kg) mixed in a 0.7% saline vehicle. We resected at least 1 cm of its ts nerve bilaterally through a small incision in the plucked skin of the neck and closed the incision with Nexaband S/C liquid.

B. Apparatus

We recorded the birds in a small wire mesh cage ($11\times7\times8.5$ in.^{3}) which was placed within an airtight Plexiglas box ($27\times14\times17.5$ in.^{3}). The Plexiglas box could then be filled with heliox (a gas mixture consisting of 70% helium and 30% oxygen, Matheson Gas, Inc., New Jersey) which was released from a pressurized tank fitted with a regulator. The heliox entered the Plexiglas box through a 1-in.-diam hole in the floor of the box. Due to properties of lighter gases, the heliox rose to the top of the box and in effect pushed the air, the more dense gas, toward the bottom of the tank where it escaped through a small opening.

The interior of the Plexiglas box contained a perch, and the box was insulated on four sides with 1-in. acoustic foam (Sonex) to reduce sound reflection. A window (3 in.^{2}) was cut in the foam on one side to allow the test bird to view other birds. An omnidirectional Realistic electret tie pin mi-

crophone (frequency response 50–15 000 Hz), connected to a reel-to-reel Teac A-3440 four-track tape recorder, was used to record the test birds' calls. This microphone was mounted directly above the wire cage that housed the test bird. We positioned a second microphone and a high-frequency click generator at approximately the same height as the head of the bird and about 10 cm apart, to measure the velocity of sound. This second 1/2-in. microphone was connected by a 3-m extension cable to a precision sound level meter (SLM) (General Radio model 1982), and the ac output of this SLM was displayed on an oscilloscope (Hitachi, V-212). The click generator consisted of a custom built driver and a high-frequency condenser speaker (Machmerth *et al.*, 1975). The click generator produced 50-μs duration clicks delivered at the rate of 200 Hz. To measure propagation time, a trigger pulse from this driver also initiated the oscilloscope sweep trace. We monitored the temperature and calculated the velocity of sound in air according to the following formula:

$$\text{Velocity}_{\text{air}} = 331.4(1 + (t/273))^{1/2}$$

with t = temperature in °C (Weast *et al.*, 1987, p. F-108). The velocity of sound in the heliox mixture was derived from the change in propagation time of the click (a halving of the propagation time means a doubling of sound velocity).

In addition to taking recordings from live birds, we also constructed several ''artificial'' syringes consisting either of toy plastic whistles of different lengths or small diameter plastic tubes of roughly the diameter and length of the budgerigar trachea. For one model, vibrating ''membranes'' constructed from small, stretched pieces of latex cut from a rubber condom were used as the source. The acoustic output from these simple sound-producing instruments were measured in air and heliox and, in the case of the tube plus latex membrane, the vibration of the latex membranes were measured simultaneously. Spectra were calculated by placing the microphone of the sound level meter at the end of the tube and passing the ac output to a Stanford Research Systems FFT spectrum analyzer (model SR760) or a Bruel & Kjaer model 4181 probe microphone, amplifier (B&K type 2609) and signal analyzer (HP 3562A). The vibrations of the latex membranes were measured with a Dantec type 41X60 laser Doppler vibrometer (Klump and Larsen, 1992). These measures were compared to calls produced by live birds to gain insight into the possible mechanisms of sound production in budgerigars.

C. Procedure and analysis

1. Recording

Each bird was placed in the small wire mesh cage after being isolated from all other birds for at least 1–2 h. The Plexiglas box was placed over the smaller cage, and it was sealed with vaseline to reduce leaks. We calculated the speed of sound in air from the ambient temperature. Once a recording session began, a cage containing several other budgerigars was placed within view of the bird in the Plexiglas box. At least 10–20 contact calls were elicited from the test bird in the normal air environment. When the bird produced the selected number of calls, the other birds were placed out of sight of the test bird. The regulator to the heliox tank was then set to release heliox into the Plexiglas box slowly. The click generator was then turned on and changes in the speed of sound with increasing concentrations of helium were calculated from the oscilloscope trace every 15 min. Once sound velocity remained unchanged for at least 15 min, the cage containing the other budgerigars was again placed in view of the test bird, and the bird's calls were recorded in heliox. On average, this steady-state concentration of heliox was reached in 50 min.

Theoretically, the velocity of sound at 0 °C should increase from 331 m/s to 506 m/s when air is replaced by a mixture of 70% helium and 30% oxygen (Weast *et al.*, 1987, p. F-104). This increase in velocity should cause the resonant frequencies of a simple tube to increase by 52.9% at 0 °C.

2. Analysis

Nine calls from each bird in air and heliox were sampled at a rate of 40 960 Hz by a Kay 5500 real time analyzer and stored as files on a microcomputer. The Kay real time analyzer was set to compute spectra from 0 to 16 kHz. In addition, we converted each call to a pattern of spectro-temporal values using a fast Fourier transform (FFT) based commercial software package called SIGNAL (Version 2.20; Beeman, 1992). These calls ($n = 18$) for each subject were analyzed for similarity using a spectrogram cross-correlation technique (Clark *et al.*, 1987) to determine the variance in call production. Power spectra were calculated using a 8192-pt FFT resulting in a spectral line every 5 Hz. To reduce the number of data points for each call, the spectral lines were averaged over bandwidths of 100 Hz to produce final power spectra consisting of 91 intensity values from 0.1 to 10.0 kHz in steps of 100 Hz.

Budgerigars produce contact calls that are quite stereotyped but the intensity can vary from call to call, especially if the bird changes position relative to the microphone. To minimize the effect of intensity variation among calls produced in each condition, the spectrum for each call was normalized to its peak frequency (the fundamental). Using the measured increase in the speed of sound in heliox as an index, the hypothesized shift in fundamental frequency was calculated. The dependent variable in this experiment was taken as the total energy in the $\frac{1}{4}$ octave frequency band around this frequency for each call produced in air and in heliox. The amount of energy in this band for calls produced in air versus calls produced in heliox was compared using a two way repeated measures ANOVA. We expected that an upward shift of tracheal resonances would lead to an increase in energy in the band of frequencies above the fundamental. A paired t test was used to test resonance effects in the one ts nerve resected bird.

Several additional steps were taken to characterize the effect of tracheal resonances further. First, Nowicki and Marler (Nowicki, 1987; Nowicki and Marler, 1988) have shown in songbirds that within the production of a single song, some constant frequency portions of the song will show evidence of harmonics appearing in heliox while other adjacent portions will not. One interpretation of these results is that some birds may be capable of rapidly (i.e., on a note by note

basis) adjusting their tracheal filter. Unfortunately, budgerigar contact calls do not lend themselves well to such an analysis since the entire call is only about 200 ms in duration and they tend to be frequency modulated. Usually there are one to a few ''constant frequency'' segments in each call but these are rarely adjacent. Nonetheless, we extracted these portions from each call and analyzed them by a spectral analysis and by zero crossing analysis (constancy of intonation according to Greenewalt, 1968) to look for overtones.

Second, as a way of separating the effects of tube resonances from other contributions to vocal output, we measured the acoustic output from 2 mm in diameter soft plastic tubes with lengths of 3.5, 5.5, 6.5, and 7.5 cm in air and in heliox. These lengths bracket the length of the typical budgerigar trachea plus buccal cavity. Broadband noise or a swept pure tone was presented through a small speaker (2P20A, 8 Ω) mounted in the large end of an infant otoscopic speculum. The small end of the speculum was inserted in the plastic tube 1.5 cm from one end. We also measured the acoustic output of a 5-cm whistle in air and heliox which was attached to a turkey baster. Squeezing the hollow, pliable rubber bulb of the baster produced rather uniform air flow through the whistle for 300–500 ms as evidenced by a relatively constant pitch.

The acoustic spectrum of the sound emanating from these tubes and whistles was calculated with the spectrum analyzer. The final spectrum consisted of an average of 500 spectra for each condition for each tube length, in air and in heliox, and with these tubes open at both ends or closed at the end nearest to the insertion of the speculum. The resonance spectra were obtained by subtracting the spectrum taken at the speculum (without a tube attached) from the spectrum obtained from the speculum plus the tube. These tests were conducted in a foam-lined Plexiglas box (20.5 $\times 11 \times 13$ in.^{3}) similar but smaller than that used to record the birds. The change in the speed of sound provided a direct measure of the exact concentration of helium.

For the more realistic psittacine syrinx, we also used stretched latex over two holes on opposite surfaces of a 2 mm in diameter plastic tube. The length of the tube from the membranes to the tip was 5.5 cm. Changing the tension on the latex membrane changed the frequency of vibration of the membranes. For a fixed tension that produced a fundamental frequency of vibration around 3 kHz, we concurrently measured the vibration of one membrane with a Dantec laser Doppler vibrometer and the acoustic output at the end of the tube with a Bruel & Kjaer 4181 probe microphone under various conditions. Briefly, we placed small reflecting spheres (Scotchlite, 30–50 μm diameter, weight<1 μg) on the membrane to improve signal-to-noise ratio. The laser was focused on the sphere and vibration measured in the manner described by Klump and Larsen (1992) for tympanum vibration in live birds. We compared acoustic output and membrane vibration for conditions including changes in tube length, flaring of the distal end of the tube (mimicking the effects of beak opening), partial obstructions of the distal end of the tube at different locations (mimicking the effect of different tongue placements), and other manipulations (e.g., loading with rubber cement, tearing, reducing tension) to only the contralateral of the two membranes.

TABLE I. Velocity of sound in air and heliox for each bird.

Bird	Air velocity (m/s)	Heliox velocity (m/s)	% Shift
93-20	343	485	41
93-14	345	495	44
92-02	346	485	40
92-06	347	490	41
91-10	345	449	30
91-12	345	487	41
93-06	346	489	41
90-03*	348	500	44

II. RESULTS

A. Recordings from live birds

Spectrogram cross correlations generated by SIGNAL provided a quantitative check that each bird always produced the same call type in air as in heliox. The nine renditions of dominant call types from the seven normally innervated birds showed an average correlation of 0.80 in air and 0.83 in heliox, with intensities ranging 73–88 A-weighted sound level in dB in air and 60–82 dB in heliox. The overall correlations for the calls produced by the ts nerve resected bird were 0.75 in air and 0.66 in heliox and intensities ranged from 63 to 77 dB in air and were constant at 68 dB in heliox.

Peaks in the total power spectrum of budgerigar contact calls appear at both odd and even harmonics. This suggests that the trachea acts predominantly as a tube open, or closed, at both ends creating resonances appropriate to $\frac{1}{2}$ the tube length, $f_n=(nv)/(2l)$. An adult budgerigar's trachea measures approximately 5 cm from the glottis to the bottom of the syrinx. Using this length in the formula above, the fundamental ($H1$) should theoretically occur at 3310 Hz, the second harmonic ($H2$) at 6620 Hz, and the third ($H3$) at 9930 Hz. The average frequency value of the fundamental and harmonics across birds was 3413 ± 439 Hz for the fundamental, 6842 ± 861 Hz for the second harmonic, and $10\,257\pm1293$ Hz for the third harmonic. This is a difference of 3.1%, 3.1%, and 3.3%, respectively. From the speed of sound and the fundamental frequency of each bird's call, we estimated the effective length of a tube, open at both ends, having a resonance centered at the bird's fundamental frequency to be about 5 cm, which is identical to the length of the actual tracheal taken from a dead budgerigar.

The average sound velocities calculated from the clicks produced in air and heliox for the eight subjects were 345.5 ± 1.5 m/s and 485 ± 14.4 m/s, respectively. This change in sound velocity results in an increase of 40% in the resonant frequencies of a simple tube and corresponds to an average steady-state concentration of helium in the Plexiglas box of 67 ± 2%. This is very close to the maximum attainable concentration of 70%. The velocity of sound in air and heliox and the relative percent increase between air and heliox for each bird are given in Table I.

Figure 1 shows the sonagrams of one token dominant contact call type, from four birds, recorded in air and heliox.

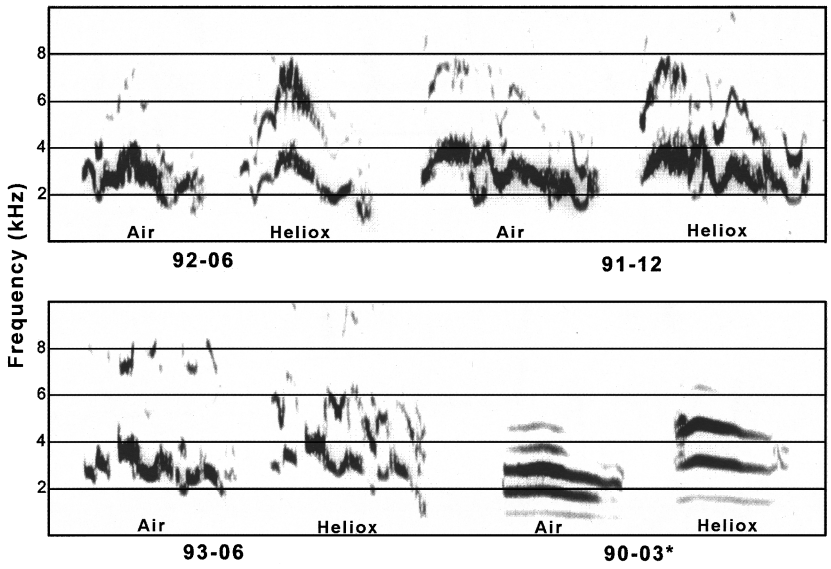


FIG. 1. Sonagrams of contact calls produced in normal air and in heliox (10-kHz analysis range, 300-Hz analysis bandwidth) for three budgerigars with normally innervated syringes and one ts nerve resected budgerigar (90-03*).

Three sonagrams were from normally innervated birds and one from the ts nerve resected bird. In the sonagrams of calls produced by normal birds in heliox, energy appeared at harmonic frequencies around 5–6 kHz that was not present in calls produced in air. The sonagrams of calls produced by the ts nerve resected bird (90-03*) showed more complex effects. For this bird, the fundamental frequency increased significantly as reflected by the greater separation between the harmonics in calls produced in heliox compared with those produced in air (Fig. 1).

Figure 2 shows the average total power spectra of the nine calls for four birds whose contact calls are shown in Fig. 1. There is a shift of energy to frequencies above the fundamental in the three normally innervated birds but little change in the fundamental. There is significantly more energy in the $\frac{1}{4}$ octave band above the fundamental in calls produced in heliox as compared to air [$F(1,6)=10.26$; $p<0.05$]. The average shift in spectra for all calls produced in air compared to heliox by normally innervated birds is shown in Fig. 3. By contrast, the bird that received bilateral ts nerve resection produced calls in heliox showing a large increase in the frequency of the fundamental and its harmonics compared to calls produced in air ($t=-3.73$; $p<0.05$). The fundamental frequency ($H1$) increased 61% and the harmonics ($H2$, $H3$, and $H4$) increased 59%, 61%, and 46%, respectively.

To obtain a more precise estimate of shifts in frequency and amplitude of vocalizations produced in heliox, we selected a relatively constant frequency portion from each contact call produced in air and heliox and analyzed only these pieces. For normally innervated birds, the frequency content of these call portions was analyzed by a zero crossing analysis followed by average frequency plots of the zero crossings for each call. The period (frequency) with the highest count for calls was correlated with the peak in the power spectrum of these constant frequency portions and there was no significant difference in peak frequency of the fundamental for calls produced in air compared to those produced in heliox [$F(1,6)=2.31$; $p>0.05$]. Figure 4 shows the average frequency plots for calls produced in air compared with calls produced in heliox normalized to the predominate frequency. The shape of these distributions confirms what was obtained by computing the power spectra—these call portions are relatively narrow band, show little evidence of side bands, and show an upward shift in frequency in heliox. For comparison, a frequency distribution from a zero crossing analysis of a 200-Hz band of noise is also shown. The intensity of

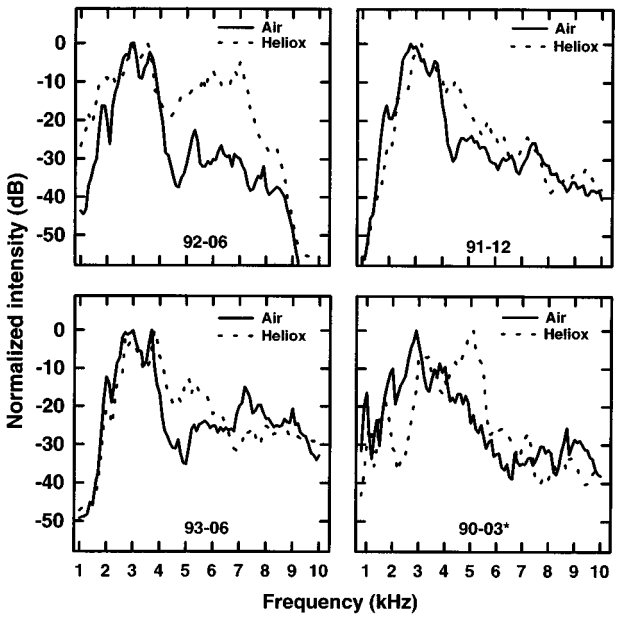


FIG. 2. Average power spectrum (40-kHz sampling rate, 10-kHz analysis range, 8192-pt FFT, yielding a frequency resolution of 5 Hz) for the nine contact calls presented in sonagraphic form in Fig. 1. In heliox, energy is decreased in the 2–4 kHz range and increased in the 4–7 kHz range (dotted line). Intensity was normalized relative to the peak in the power spectrum.

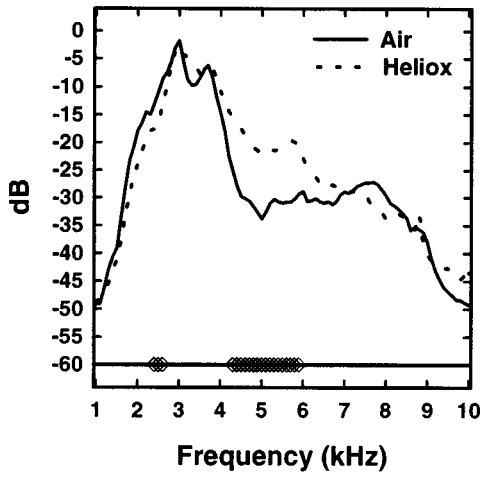


FIG. 3. Paired t-tests were performed on each spectral line throughout these average spectra and the symbols at the bottom show which of these spectral lines were different between calls produced in air and in heliox. On average, calls produced in heliox show less energy at 2.4–2.6 kHz (below the fundamental) and more energy at 4.3–5.9 kHz (above the fundamental) compared to calls produced in air.

the fundamental relative to the intensity of the second harmonic when birds were vocalizing in air compared to heliox were also not significantly different [$F(1,6)=0.09$; $p>0.05$] by a power spectral analysis.

Of the six normally innervated birds, one bird (92-02) was somewhat unusual in showing four relatively constant frequency portions in its contact call. This provided the opportunity to examine several constant frequency portions in the same call in air and in heliox. Figure 5 shows the power spectra from four constant frequency portions within this bird's contact call produced in air and in heliox. Overall these spectra show more energy at higher frequencies in calls produced in heliox, but there were only slight shifts in the frequency and intensity of the fundamental and harmonics, and these shifts are not in a consistent direction.

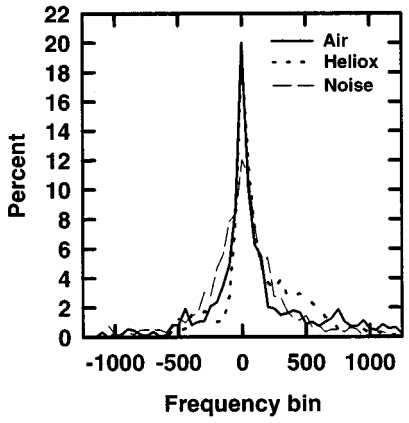


FIG. 4. Frequency plots for calls produced in air (solid line) and heliox (dotted line) along with the frequency distribution resulting from a zero crossing analysis of a 200-Hz band of noise (dashed line) are shown.

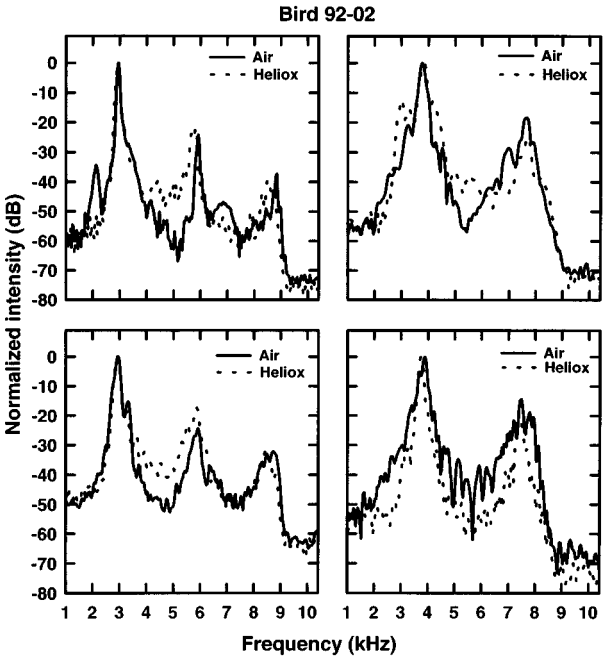


FIG. 5. Total power spectra for four flat segments within a single bird's (92-02) call. Other than an overall shift of energy to higher frequencies, these spectra are noteworthy for their lack of consistent changes from air to heliox. Intensity was normalized relative to the peak in the power spectrum.

B. Recordings from artificial sources

The control experiments with plastic tubes, whistles, and ''artificial'' syringes illustrated the effects of passive tube resonances and several expected effects are shown in Fig. 6(a)–(d). Opening and closing a 5.5-cm tube [Fig. 6(a)] driven by a broadband white noise showed a shift in resonance frequencies from both odd and even harmonics to only odd harmonics. Changing the tube length also shifted resonance frequencies as expected [Fig. 6(b)] and the resonance shifts for the tubes open at both ends more closely matched the peaks expected by the calculations (see Table II) than did those of the tubes closed at one end. The upward shift in the spectra of the noise emanating from a 5.5-cm tube open at both ends in air and in heliox shown in Fig. 6(c) is also predictable from the concentration of helium in the environment. In air, the fundamental resonance of this tube is at 3062 Hz, the second harmonic at 5969 Hz, and the third harmonic at 8500 Hz. In heliox, the three peaks are shifted upward an average of 35% so that the fundamental was located at 4188 Hz, the second harmonic at 7938 Hz, and the third harmonic at 11 563 Hz. Very similar effects occurred when the sound source was changed from a small speaker producing white noise to a 5-cm plastic whistle. The effect of replacing air with heliox on the spectra obtained from the plastic whistle is shown in Fig. 6(d).

In aggregate, the effects shown in Fig. 6 with plastic tubes and whistles can be explained by simple physical principles involving a broadband or harmonic source, changes in the speed of sound in air and heliox, and the resonances of a simple tube. In these simple syringeal models, the source and the resonating tube are clearly not coupled. The more realistic syringeal model constructed from a 5-cm tube and two

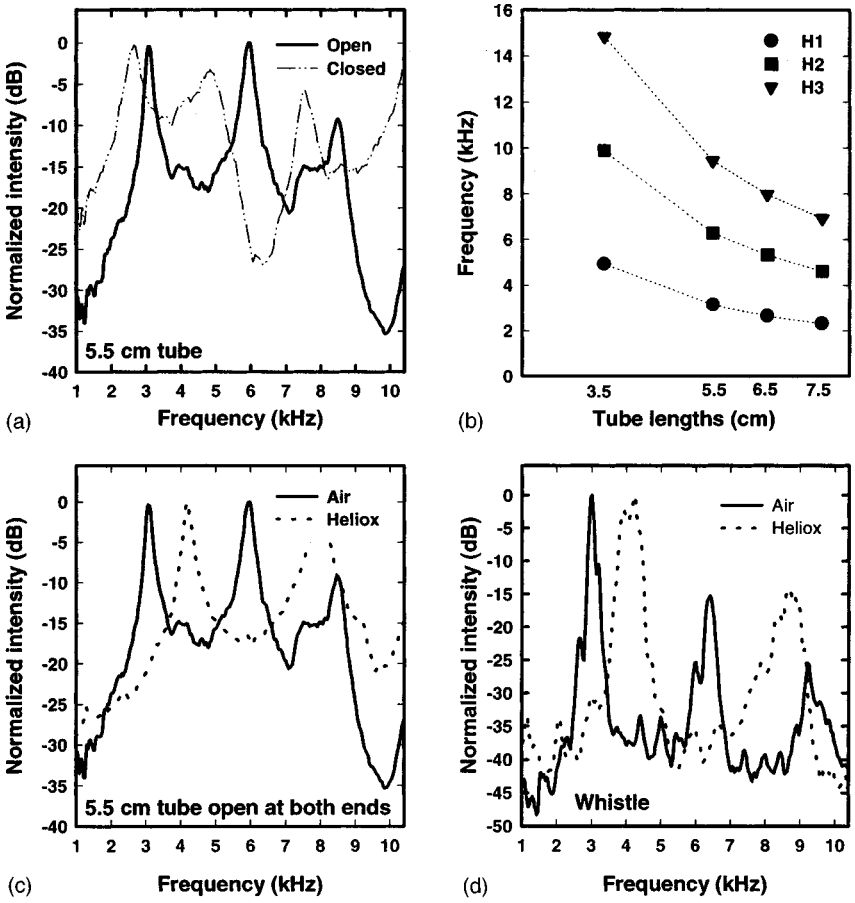


FIG. 6. (a) Total power spectra from a 5.5-cm tube open at both ends (solid line) and open at one end and closed at the other (dashed line). (b) Change in frequency of the fundamental and harmonics as a function of tube length. (c) Total power spectra from a 5.5-cm tube open at both ends in air (solid line) and heliox (dotted line). (d) Total power spectra from a whistle played in air (solid line) and heliox (dotted line).

vibrating latex membranes driven by moving air from a turkey baster (the pressure source) showed different effects. These changes from air to heliox were similar to those seen in the calls of a normally innervated bird. Figure 7(a) shows the effect of replacing air with heliox on the output from a model syrinx with vibrating latex membranes. For comparison, Fig. 7(b) shows the effect of replacing air with heliox on the constant frequency portion of the contact call produced by a live budgerigar. In both cases, there is only a small upward shift in the frequency of the fundamental with little or no change in its amplitude—an effect that could be due to the decreased load on the vibrating membranes from the lighter gas, heliox. In both cases, there is also an overall shift in energy to higher frequencies above the fundamental.

Figure 8 shows the relation between one of the vibrating rubber membranes and the spectrum of the acoustic output at the end of the 5-cm tube. Peaks in the vibration spectrum and the acoustic spectrum are at exactly the same frequency. We repeated these measures with different tube lengths (mimicking putative tracheal changes when a bird stretches its neck), partial occlusion of the tube opening (as might occur from occlusion by the tongue), and flaring of the tube (loosely modeling the effects of opening the beak). None of these manipulations had any effect on the relation between the spectrum of the vibrating membrane and the acoustic spec-

TABLE II. Predicted and actual measures for resonance frequencies from a 5.5-cm tube.

Predicted				Actual			
	Open		Closed		Open		Closed
*H*1	3145	*H*1	1573	*H*1	3063	*H*1	2656
*H*2	6290	*H*3	4718	*H*2	5969	*H*3	4844
*H*3	9435	*H*5	7863	*H*3	8500	*H*5	7531

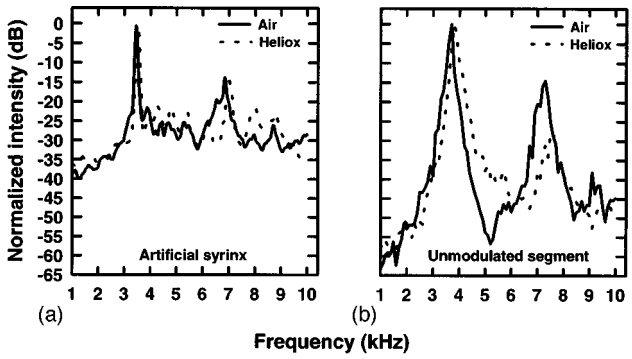


FIG. 7. Example of the total power spectrum for the artificial syrinx compared to the power spectrum of an unmodulated segment of a contact call for bird 93-06. Intensity was normalized relative to the peak in the power spectrum.

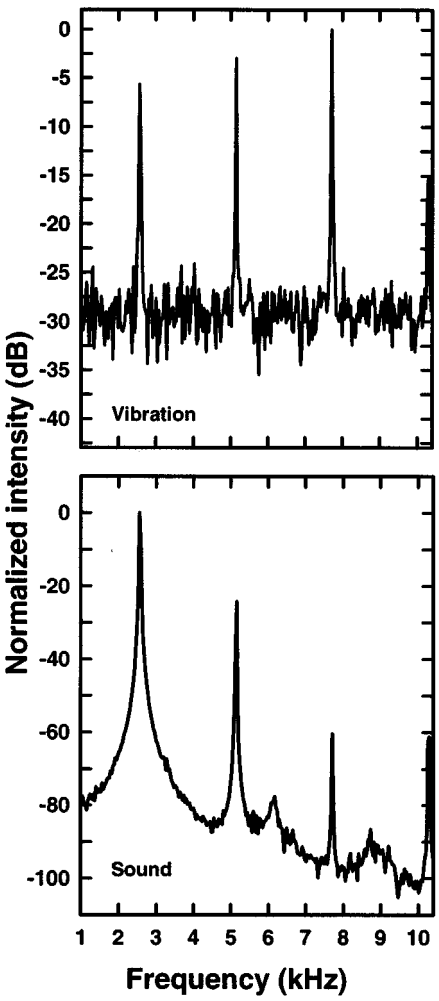


FIG. 8. Power spectra for the vibration of the latex membrane and the acoustic output of the artificial syrinx. Intensity was normalized relative to the peak in the power spectrum.

trum measured at the output of the tube. Decreasing tension on the opposite latex membrane, puncturing it, or loading it with rubber cement resulted in a decrease in fundamental frequency.

III. DISCUSSION

These experiments have attempted to elucidate the mechanical processes involved in the production of vocalizations by budgerigars. To return to the questions posed earlier:

Can the trachea act as an acoustic filter which modifies the spectrum of vocalizations?

The resonant frequencies of the trachea depend on the length, diameter, and stiffness of the tissues. A wealth of correlational evidence argues that the dominant frequencies in the calls of birds are those syringeal membrane frequencies that most nearly approach the tracheal resonant frequencies (see, for example, Sutherland and McChesney, 1965). The strong suggestion here is that the trachea (and probably other air chambers) can act to enhance or dampen certain harmonic overtones that provide timbre to a bird's voice. But there may be other possibilities as well. Zebra finches, for example, can emphasize and suppress particular harmonics in their calls and song syllables, and these patterns can be learned (Williams *et al.*, 1989). Syringeally denervated zebra finches presumably lack such timbre control suggesting that some of the harmonic suppression found in normal zebra finch song is accomplished by the syrinx and not the trachea (Williams *et al.*, 1989).

The tonal, frequency modulated contact calls of budgerigars have fundamental frequencies that fall between 2 and 4 kHz. Comparing the total power spectrum of the contact calls produced in heliox with those produced in air shows an increase at high frequencies, consistent with the notion that energy in frequencies above the fundamental in air are normally attenuated by a tracheal filter centered around 3 kHz. These results suggest that the trachea might be acting as a broad, bandpass acoustic filter that shifts upward in frequency in heliox, thereby slightly attenuating low-frequency spectral components and simultaneously enhancing high-frequency components of the vocalization spectrum.

Interestingly, a closer examination of brief, constant frequency portions of contact calls, however, failed to show either a consistent reduction in amplitude of the fundamental or other changes in amplitudes of the other harmonics of calls produced in heliox as compared with those produced in air. This is in contrast to that reported by Nowicki (1987) for the songbird. In theory, a filter centered over the fundamental (and approximating the width of those shown for our plastic tubes), when shifted upward in heliox, should cause a decrease in amplitude of the fundamental and an increase in amplitude of the second and perhaps the third harmonic over calls produced in air. Instead, the frequency and amplitude of the fundamental changed very little in calls produced in air compared to those produced in heliox, and the amplitude of the second harmonic showed no consistent change relative to the amplitude of the fundamental.

In general, the findings in live birds are similar to results from sounds produced in air and heliox by an artificial syrinx consisting of vibrating membranes attached to a small diameter tube approximating the length of the budgerigar trachea. These results suggest that tracheal resonances do have a slight effect but do not normally play a very large role in determining the spectral content of contact calls.

Does the trachea function as a tube that is open at both ends (syrinx and beak), closed at both ends or closed at one end (syrinx)?

The preponderance of evidence both from the vocalizations of normally innervated birds, the syringeally denervated bird, as well as from simple models of the avian syrinx strongly suggests that the budgerigar syrinx can best be modeled as a tube open at both ends. But others (Nottebohm, 1976; Westneat *et al.*, 1993) have suggested that the open end (glottal end) can be constricted, thus changing the surface area of the opening so that the trachea can act more like a tube that is closed at both ends. Since we could find both odd and even harmonics in budgerigar contact calls, and the fundamental of budgerigar contact calls is more closely approximated by the formula $f_n=(nv)/(2l)$ given a vocal tract

length of 5 cm, we conclude that the trachea is functioning most of the time as a tube either open or closed at both ends rather than a tube open at one end and closed at the other.

Can the resonances of the budgerigar vocal tract be actively manipulated by the bird?

Even though tracheal resonances may play only a minor role in shaping the spectral characteristics of budgerigar vocalizations, recent work has suggested that acoustic resonances can be altered in the avian vocal production system in at least three ways: (1) lengthening or shortening the tube; (2) partially occluding the open end of the resonating tube or changing its size; or (3) flaring the open end of the tube (Nowicki and Marler, 1988; Westneat *et al.*, 1993). There is ample correlational and anecdotal evidence for a relationship between vocal tract length and vocal pitch. Birds with longer and wider tracheas tend to have deeper voices (e.g., Whooping crane and Trumpeter swan, Portmann, 1950) while birds with narrower and shorter tracheas tend to have higher pitched voices (nestling calls in passerines and budgerigars, Popp and Ficken, 1991; Brittan-Powell *et al.*, in press).

Westneat *et al.* (1993) proposed that beak gape can alter the resonance properties of a bird's trachea in at least two ways: by effectively shortening the distal portion of the tube (thereby increasing the resonant frequency) or by altering the impedance at the anterior end of the vocal tract by occluding the open end of the tube. These investigators conclude that beak motion is functionally related to sound production in songbirds but the evidence is largely correlational.

In the case of budgerigars, there are vocal tract resonances that are passive consequences of tube characteristics (shown dramatically by the denervated bird) but it is unlikely that these are actively manipulated by the bird. Casual observation of budgerigars and more formal observation in other psittacines such as the African Grey Parrot (Patterson and Pepperberg, 1994) show that tongue movement and beak opening is correlated with sound production. Both species, however, are also able to produce species-typical vocalizations with the beak nearly completely closed. Moreover, recent work on denervation of multi-craniomotor systems including the tongue, pharynx, and larynx do not alter contact call patterns appreciably (Brauth *et al.*, in press).

The present results from measurements with simple plastic tubes are illuminating. For a tube the length of the budgerigar trachea, the fundamental falls around 3300 Hz and the average peak in the power spectrum of budgerigar contact calls was about 3400 Hz. Moreover, as in the bird with the denervated syrinx, the frequency of the fundamental and the harmonics obtained from a plastic tube shifted upward in frequency in heliox with the levels of the fundamental changing appropriately, effects completely expected with a source coupled to a resonating tube. This effect was distinctly not observed in normal budgerigars. Together with the fact that manipulations of tube length, partial obstructions, and flaring—all suggested as mechanisms for actively manipulating tracheal resonances—have virtually no effect on either the vibration of the membranes in our artificial syrinx or on the acoustic output at end of the tube, we conclude that active changes in tracheal resonances by the budgerigar are probably not required for normal vocal production. Instead, it is likely that normally innervated budgerigars probably override the passive resonance characteristics of their trachea.

Are the vibrating sources in the budgerigar syrinx coupled to the trachea?

Similar arguments can be brought to bear on the question of whether the syringeal membranes of budgerigars are coupled to the trachea. The fact that the fundamental frequency (nor the harmonics for that matter) of contact calls produced in heliox by normally innervated budgerigars does not significantly increase suggests, at best, a very loose coupling between the syringeal membranes and the trachea. In this regard, the source-filter theory of vocal production as proposed for human speech production is probably an appropriate model for the call production in budgerigars.

The findings from the denervated bird provide important support for this conclusion. Bilateral resection of the ts nerves have numerous consequences (e.g., denervation of the extrinsic tracheal and the intrinsic and extrinsic syringeal muscles among other effects) that can affect the mechanical properties of the vocal system in a dramatic way. Contact calls in air and heliox produced by the syringeally denervated bird differed significantly from those produced by normally innervated birds. In this bird, the harmonic frequencies increased by the amount expected from a shift in tracheal resonances due to changes in the speed of sound in heliox. Moreover, the fundamental frequency in this bird's contact calls increased by a proportional amount suggesting the syringeal membranes were now coupled to the trachea. For this bird, a more appropriate model for vocal production might indeed be a woodwind or brass instrument (Benade, 1976) where the fundamental frequency of such an instrument played in helium increases in the same proportion as the resonances of the corresponding tube (Nowicki and Marler, 1988). These are important data since they show that tracheal resonances *can* have a very pronounced effect on the spectral distribution of energy in contact calls exactly as predicted by the resonance properties of tracheal-length tubes in heliox but only when the syringeal musculature is denervated. Such effects could occur with a coupled source-resonator model but not with a uncoupled source-filter model of vocal production.

Do the two syringeal membranes act as independently vibrating sound sources?

As early as 1947, researchers have suggested that birds might have independent control of their two sound sources (Potter *et al.*, 1947). However, it was not until 1968 that Greenewalt, through a zero crossing analysis of the songs of many songbird species, provided the first evidence that birds have ''two voices.'' He concluded that the two sides of the songbird syrinx may operate independently and produce harmonically unrelated tones simultaneously. Later denervation experiments provided more evidence that the two sides of the syrinx can act independently (Nottebohm, 1971; Lemon, 1973; Nottebohm and Nottebohm, 1976; Seller, 1979). The specific elements deleted from songs of birds with unilateral syringeal denervation were presumably those generated by

the disabled side of the organ. The situation is not completely clear cut, however, since Nowicki and Capranica (1986) clearly showed that, at least in the chickadee, the two syringeal sources behave as if they are coupled in some manner. Using grey catbirds (*Dumetella carolinensis*) and brown thrashers (*Toxostoma rufum*), Suthers (1990) provided more direct and refined evidence of a two-voice theory. He showed that the two sides of the syrinx can act in three ways: both sides may contribute simultaneously to a note or syllable, both may generate the same sound, or each side may produce a different sound. Taken together, these studies show that the two sides of the syrinx are not limited to the production of different sounds but can, in fact, produce combinations of identical or very similar sounds. Thus, the original two-voice theory (at least in its simple form) probably underestimates the true capabilities of the avian syrinx.

In the songbird syrinx, the internal tympaniform membranes (ITMs), located on the medial walls of the primary bronchi just caudal to the syringeal lumen, vibrate to generate sound. Since the tension of the right and left ITMs can be independently manipulated and are driven by potentially independent columns of air, the two syringeal halves can produce harmonically unrelated sounds. In the parrot syrinx, on the other hand, the external tympaniform membranes (ETMs) located within the syringeal lumen are driven by a single column of air, and therefore cannot be ''two voiced'' in the same sense as in the songbird syrinx. Moreover, while hypoglossal innervation of the syrinx is ipsilateral in songbirds, it hemidecussates in parrots through an anastomosis zone distal to the syrinx.

In theory, then, either the left or right hypoglossal nucleus could support normal vocalizations in budgerigars; a point supported by the fact that resections of the right or left ts nerve below the anastomosis do not seem to dramatically affect budgerigar contact calls (Heaton *et al.*, 1995). It is interesting that although the syrinx of the Orange-Winged Amazon parrot is structurally similar to that of the budgerigar, unilateral syringeal denervation just caudal to the anastomosis (disrupting control of the ipsilateral half of the syrinx) markedly affects the long call in this species. This suggests, at least for this psittacine, the behavior of one syringeal membrane during vocal production is dependent on the behavior of the other (Nottebohm, 1976). We observed a similar result with our ''artificial'' syrinx consisting of a 5-cm plastic tube and two latex rubber membranes when one was ''deactivated'' by either reducing tension, puncturing the membrane, or loading the membrane with rubber cement—these manipulations all lowered the fundamental frequency of the acoustic output of the tube. Neither spectral analyses nor zero crossing analyses of constant frequency portions of contact calls in air and in heliox showed any evidence of two sources (''voices'') operating independently when budgerigars produced contact calls.

IV. CONCLUSION

Taken together, these results from budgerigars producing contact calls in air and heliox and from various ''artificial'' syringes and tracheas suggest that budgerigars use their syringeal membranes as a unitary sound source that produces acoustic energy in a relatively narrow frequency band with fairly shallow skirts especially on the low-frequency side of the filter. The fundamental frequency is matched to the resonant frequency of the trachea. The syrinx is not normally coupled to the tracheal resonator, and tube resonances play only a minor role in shaping the overall spectral profile of contact calls.

ACKNOWLEDGMENTS

This work was supported by Grants Nos. DCD00198 and MH00982 to R.J.D. and the Danish National Research Foundation to O.N.L. We thank Lee Miller for lending us his high-frequency speaker and Franz Goller for comments on an earlier draft.

Beeman, K. (**1992**). *Signal Technology*, V2.20 (Engineering Design, Massachusetts).

Beil, R. (**1962**). ''Frequency analysis of vowels produced in a helium-rich environment,'' J. Acoust. Soc. Am. **34,** 347–349.

Benade, A. H. (**1976**). *Fundamentals of Musical Acoustics* (Oxford U.P., New York).

Brittan-Powell, E. F., Dooling, R. J., and Farabaugh, S. M., (**in press**). ''Vocal development in the budgerigar (*Melopsittacus undulatus*): Contact call learning,'' J. Comp. Psychol.

Brockway, B. F. (**1969**). ''Roles of budgerigar vocalization in the integration of breeding behaviour,'' in *Bird Vocalizations*, edited by R. A. Hinde (Cambridge U.P., Cambridge), pp. 131–158.

Brauth, S. E., Heaton, J. T., Shea, S. D., Durand, S. E., and Hall, W. S. (**in press**). ''Functional anatomy of forebrain vocal control pathways in the budgerigar (*Melopsittacus undulatus*),'' Proc. N. Y. Acad. Sci.

Clark, C. W., Marler, P., and Beeman, K. (**1987**). ''Quantitative analysis of avian vocal phonology: An application to Swamp Sparrow song,'' Ethology **76**, 101–115.

Dooling, R. J. (**1986**). ''Perception of vocal signals by the budgerigar (*Melopsittacus undulatus*),'' Exp. Biol. **45**, 195–218.

Dooling, R. J., Gephart, B. F., Price, P. H., McHale, C., and Brauth, S. E. (**1987**). ''Effects of deafening on the contact calls of the budgerigar (*Melopsittacus undulatus*),'' Anim. Behav. **35**, 1264–1266.

Farabaugh, S. M., Brown, E. D., and Dooling, R. J. (**1992**). ''Analysis of warble song of the budgerigar (*Melopsittacus undulatus*),'' Bioacoustics **4**, 111–130.

Farabaugh, S. M., Linzenbold, A., and Dooling, R. J. (**1994**). ''Vocal plasticity in budgerigars (*Melopsittacus undulatus*): Evidence for social factors in the learning of contact calls,'' J. Comp. Psych. **108**, 81–92.

Fletcher, N. A. (**1992**). *Acoustic Systems in Biology* (Oxford U.P., New York).

Gaunt, A. S. (**1983**). ''A hypothesis concerning the relationship of syringeal structure to vocal abilities,'' The Auk **100**, 853–862.

Gaunt, A. S., Gaunt, S. L. L., and Casey, R. M. (**1982**). ''Syringeal mechanics reassessed: evidence from *Streptopelia*,'' The Auk **99**, 474–494.

Greenewalt, C. H. (**1968**). *Bird Song: Acoustics and Physiology* (Smithsonian Institution, Washington, DC).

Heaton, J. T., Farabaugh, S. M., and Brauth, S. E. (**1995**). ''Effect of syringeal denervation in the budgerigar (*Melopsittacus undulatus*): The role of the syrinx in call production,'' Neurobiol. Learn. Mem. **64**, 68–82.

Klump, G. M., and Larsen, O. N. (**1992**). ''Azimuthal sound localization in the European Starling (*Sturnus vulgaris*): I. Physical binaural cues,'' J. Comp. Physiol. **170,** 243–251.

Lemon, R. E. (**1973**). ''Nervous control of the syrinx in white-throated sparrow (*Zontrichia albicollis*),'' J. Zool. **171**, 131–140.

Manogue, K. R., and Nottebohm, F. (**1982**). ''Relation of medullary motor nuclei to nerves supplying the vocal tract of the budgerigar (*Melopsittacus undulatus*),'' J. Comp. Neurol. **204,** 384–391.

Machmerth, V. H., Theiss, D., and Schnitzler, H.-U. (**1975**). ''Konstruktion eines Luftultraschallgebers mit konstantem Frequenzgang im Bereich von 15 kHz bis 130 kHz,'' Acustica **34,** 81–85.

Myers, J. A. (**1917**). ''Studies of the syrinx of *Gallus domesticus*,'' J. Morphol. **29**, 165–214.

Nottebohm, F. (**1971**). ''Neural lateralization of vocal control in a passerine bird: I. Song,'' J. Exp. Zool. **177**, 229–262.
Nottebohm, F. (**1976**). ''Phonation in the Orange-winged Amazon Parrot, *Amazona amazonica*,'' J. Comp. Physiol. **108**, 157–170.
Nottebohm, F., and Nottebohm, M. E. (**1976**). ''Left hypoglossal dominance in the control of canary and white-crowned sparrow song,'' J. Comp. Physiol. **108**, 171–192.
Nowicki, S. (**1987**). ''Vocal tract resonances in oscine bird sound production: evidence from birdsong in a helium atmosphere,'' Nature (London) **325,** 53–55.
Nowicki, S., and Capranica, R. R. (**1986**). ''Bilateral syringeal interaction in the production of oscine bird sound,'' Science **231,** 1297–1299.
Nowicki, S., and Marler, P. (**1988**). ''How do birds sing?,'' Music Percept. **5**, 391–426.
Nowicki, S., Mitani, J., Nelson, D., and Marler, P. (**1989**). ''The communicative significance of tonality in birdsong: responses to songs produced in helium,'' Bioacoustics **2**, 35–46.
Patterson, D. K., and Pepperberg, I. M. (**1994**). ''A comparative study of human and parrot phonation: acoustic and articulatory correlates of vowels,'' J. Acoust. Soc. Am. **96,** 634–638.
Podos, J., Sherer, J. R., Peters, S., and Nowicki, S. (**1995**). ''Ontogeny of vocal tract movements during song production in song sparrows,'' Anim. Behav. **50**, 1287–1296.
Popp, J., and Ficken, M. S. (**1991**). ''Comparative analysis of acoustic structure of passerine and woodpecker nestling calls,'' Bioacoustics **3**, 255–274.
Portmann, A. (**1950**). ''Le developpement postembryonnaire,'' in *Traite de Zoologie Tome XV Oiseaux*, edited by P. Grasse (Masson et Cie, Paris).
Potter, R. K., Kopp, G. A., and Green, H. C. (**1947**). *Visible Speech* (Van Nostrand, New York).
Rossing, T. D. (**1989**). *The Science of Sound* (Addison-Wesley, New York).
Seller, T. J. (**1979**). ''Unilateral nervous control of the syrinx in Java sparrows (*Padda oryzivora*),'' J. Comp. Physiol. **129**, 281–288.
Sutherland, C. A., and McChesney, D. S. (**1965**). ''Sound production in two species of geese,'' Living Bird **4**, 99–106.
Suthers, R. A. (**1990**). ''Contributions to birdsong from the left and right sides of the intact syrinx,'' Nature (London) **347,** 473–477.
Suthers, R. A. (**1994**). ''Variable asymmetry and resonance in the avian vocal tract: a structural bias for individually distinct vocalizations,'' J. Comp. Physiol. **175**, 457–466.
Weast, R. C., Astle, M. J., and Beyer, W. H. (Eds.) (**1987**). *CRC Handbook of Chemistry and Physics* (CRC, Boca Raton, FL).
Westneat, M. W., Long, J. H., Hoese, W., and Nowicki, S. (**1993**). ''Kinematics of birdsong: Functional correlation of cranial movements and acoustic features in sparrows,'' J. Exp. Biol. **182**, 147–171.
Wiliams, H., Cynx, J., and Nottebohm, F. (**1989**). ''Timbre control in zebra finch (*Taeniopygia guttata*) song syllables,'' J. Comp. Psychol. **103**, 366–380.
Wyndham, E. (**1980**). ''Environment and food of the budgerigar (*Melopsittacus undulatus*),'' Austral. J. Ecol. **5**, 47–61.

Hearing deficits measured in some *Tursiops truncatus*, and discovery of a deaf/mute dolphin

Sam H. Ridgway and Donald A. Carder
Biosciences Division, Naval Command, Control and Ocean Surveillance Center, RDT&E Division, Code D3503B, 49620 Beluga Road, Room 200, San Diego, California 92152-6266

(Received 22 April 1996; accepted for publication 14 June 1996)

Eight bottlenose dolphins *Tursiops truncatus* (four male, four female) were trained to respond to 100-ms tones. Three male dolphins (ages 23, 26, and 34) exhibited hearing disability at four higher frequencies—70, 80, 100, and 120 kHz even at 111–135 dB *re*:1 μPa. Two females (ages 32 and 35) responded to all frequencies as did a male (age 7) and a female (age 11). One female (age 33) responded to all tones at 80 kHz and below; however, she failed to respond at 100 or 120 kHz. One young female dolphin (age 9) exhibited no perception of sound to behavioral or electrophysiological tests. This young female was not only deaf, but mute. The dolphin was monitored periodically by hydrophone and daily by trainers (by ear in air) for 7 years until she was age 16. The animal never whistled or made echolocation pulses or made burst pulse sounds as other dolphins do. [S0001-4966(97)02812-9]

PACS numbers: 43.80.Lb, 43.80.Ka, 43.80.Jz [FD]

INTRODUCTION

Audiograms have been done on several species of the cetacean superfamily Delphinoidea (Au, 1993; Richardson, 1995). Most of these species are represented by only one or two young animals. All of these animals, with the exception of one killer whale, *Orcinus orca* (Hall and Johnson, 1971), had good sensitivity from 60–120 kHz. The first detailed audiogram of the bottlenose dolphin, *Tursiops truncatus*, yielded a threshold of 42 dB *re*: 1 μPa (10^{-14} W m^2) at 60 kHz with about a 20-dB increase at 120 kHz and a very steep increase thereafter, to a maximum of 150 kHz (Johnson, 1967). Johnson's animal was 9 years old.

During an acoustic response time task (Ridgway *et al.*, 1991), we tested the hearing of eight *Tursiops* (four males, four females) at levels that were expected to be 60–80 dB above threshold, based on earlier delphinoid audiograms mentioned above. One of our experimental dolphins, a male aged 26, had been tested 13 years earlier by Ljungblad *et al.* (1982). The animal had been shown to have good hearing at this earlier date. Although this dolphin (IAY), at age 13 in the early 1980s, had thresholds 5–10 dB higher than the male age 9 used by Johnson (1967), Au (1993) has pointed out that this difference could possibly be accounted for, in part at least, by the differences in test methodology.

Until we first presented this at the Denver meeting of the Acoustical Society (Ridgway and Carder, 1993a), no tests of hearing had been done with older (>25 years) dolphins of either sex. During the past 33 years with the Navy marine mammal program, we have observed sound production and some related behavior in about 200 bottlenose dolphins (cf. Ridgway, 1983). Recently, we had the opportunity for the first time to observe a dolphin that was both deaf and mute.

I. MATERIALS AND METHODS

Age and sex of each of the experimental dolphins are given in Fig. 1. The oldest male was age 34 at the time of the test and had been with our laboratory since 1962. During the 1960s and 1970s, he had demonstrated apparent good hearing and echolocation ability, although an audiogram had never been done. Health and medication records were kept on all the dolphins since their initial acquisition or birth. Among the animals we tested, records on animal MAU, for example, go back to 1962. The potential for ototoxicity has always been a consideration for dolphin medication, however, two of the animals had received aminoglycosides (Anon, 1994) for infections. Animal MAY was given gentamycin (600 mg twice daily) for seven days in 1980, six years prior to the hearing tests. Dolphin SLA was given one injection of penicillin/streptomycin in 1968 and a single injection of amakacin and penicillin G in 1992.

The dolphins listed in Fig. 1 were trained to whistle or burst pulse when a stimulus tone (St) was delivered through an underwater hydrophone located 1 m in front of the animal. This training was similar to that reported previously (Ridgway and Carder, 1988; Ridgway *et al.*, 1991). We noticed that when a dolphin whistled, there was a characteristic movement along the left posterior margin of the nasal plug of the closed blowhole. Burst pulse sounds generally resulted in a somewhat different movement, more to the right side of the dolphin's blowhole. Our trainers quickly induced dolphins to repeat vocalizations by tapping with a finger or manipulating the area of the blowhole where movement or any escaping air concurrent with sound had been detected. After whistles or burst pulse sounds were reliably elicited in this manner, the signal was transferred slowly to a simple stroke to the dolphin's melon. Then, with the dolphin underwater in front of the trainer, the melon stroke was paired with a tone until the animal reliably gave the vocalization each time the tone was presented through the hydrophone.

The animals were trained to station on a plastic bite plate 1.0 m underwater and remain stationary until an underwater buzzer (bridge or S2 signal that informs the animal that a fish reward will soon follow) was sounded. Initially,

HIGH-FREQUENCY RESPONSE OF EIGHT TURSIOPS
TO A 100 ms TONE STIMULUS OF 111 dB re 1 μPa

Animal	Sex	Age	Stimulus Frequency (kHz) 40	60	80	100	120
MAU	♂	34	▲	△	△	△	△
MKA	♂	23	▲	◭	△	△	△
IAY	♂	26	▲	◭	△	△	△
NAY	♂	7	▲	▲	▲	▲	▲
TOD	♀	35	▲	▲	▲	▲	▲
SLA	♀	33	▲	▲	◭	△	△
BRT	♀	32	▲	▲	▲	▲	▲
SAY	♀	11	▲	▲	▲	▲	▲

▲ > 90% Responses
◭ > 75% Responses
◭ > 50% Responses
△ < 5% Responses

FIG. 1. The animal identifier, sex, age, and indication of correct responses to 100-ms, 111-dB tones for eight bottlenose dolphins *Tursiops truncatus* employed in this study.

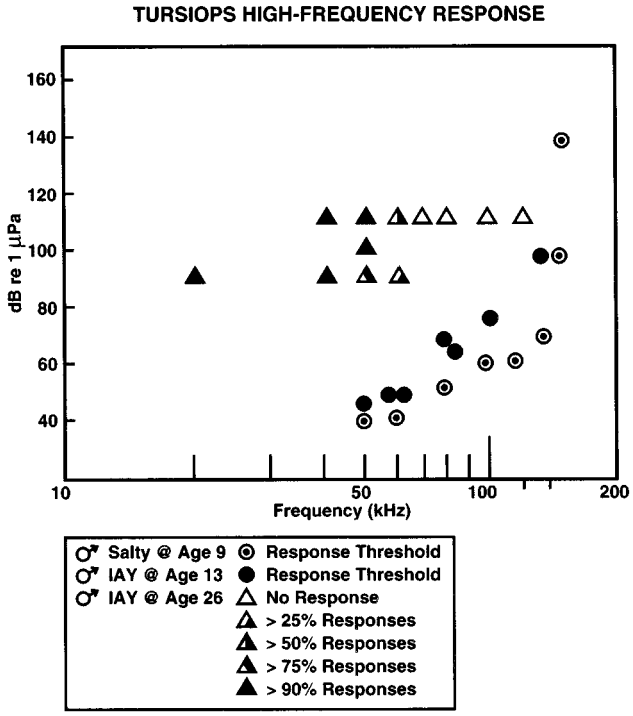


FIG. 2. Responses of two bottlenose dolphins to high-frequency tones. Points (circles with dot) for Salty at age 9 from Johnson (1967), points (filled circles) for IAY at age 13 from Ljungblad *et al.* (1982). All triangles from present study.

the dolphin was given an S2 and rewarded each time it vocalized after a tone. Gradually the reward schedule was reduced until the animal made up to 20 responses in a row. The S2 was given immediately after the last correct vocal response in the series. The S2 was followed by a reward of one to several fish when the dolphin returned to the surface to breathe. The longest period the animal was required to remain on the underwater station was two minutes; however, both the time the animal was required to remain on the underwater station, and the number of tones presented during this time were varied in a random fashion. For catch (no stimulus) trials, the dolphins were sent down to the station but no tones were presented. After the dolphin had remained stationary and silent for periods varying between 30 and 120 s, the S2 was given and the animal surfaced for reward. Improper responses, i.e., leaving the station before the S2, vocalizing prior to or in the absence of the stimulus, or giving the wrong vocalization, were not reinforced with fish.

A trial series or testing dive (TD) was started when the trainer signaled the animal to go down to the plastic bite plate 1.0 m under the surface (Fig. 2), and 1.0 m from the stimulus hydrophone (an F42B for frequencies of 5–70 kHz; an LC-10 for frequencies of 80–120 kHz). During the earlier stages of training, 20% of the TDs were catch trials which were inserted randomly in the series of TDs. When the false alarm rate decreased to 5% or less of the correct response level, catch trials were reduced to 10% of TDs.

Tone stimulus (St) duration was 100 ms with a 2-ms gradual rise in intensity at onset and decline on termination. The findings of Johnson (1968) suggested to us that this duration was adequate. With three of the older males, some tests were done with 300- and 450-ms tones. Frequencies were 5, 10, 20, 40, 50, 60, 70, 80, 100, and 120 kHz. Stimuli were 111 dB, increasing in 6-dB steps to 135 dB in cases where the animal did not respond to the baseline level. With the dolphin at 1.0-m depth and 1 m from the St hydrophone (Fig. 2), the trainer waited a variable period then pushed a switch starting a randomly variable St block. The computer selected Sts from a file in random initial delay and interval (1.1–2.1 s in 0.1-s steps) and offered Sts via a St generator as long as the trainer held the switch button down. Thus, the trainer could give several Sts in a row in the randomly variable sequence from the computer file, then let up on the switch and interpose a period without Sts before pressing the switch again for more Sts. Randomness in St delivery was maintained both by the computer program and by the trainer's switch press out of sight of the dolphin. Animal responses (ARs=whistle or burst pulse) were received by another hydrophone, digitized, and stored for confirmation of correct response. Each AR file with 20–200 Sts was edited on a CRT display of a 700-ms St window. No-AR trials, noisy trials, and wrong ARs were identified, and a database was constructed. The baseline stimulus of 111 dB *re*: 1 μPa generally exceeded background noise in San Diego Bay by about 50–80 dB in the 60–120 kHz range (also see Au *et al.*, 1985).

In addition to attempts at applying the above procedures, the apparently deaf dolphin SIB was trained to respond to a 45-kHz underwater locating beacon[1] (model DK355L), a ''pinger'' that was lowered into the water. The source level of the pinger was 160 dB *re*: 1 μPa and it produced one 10-ms pulse each second. After the animal had learned to take fish from the trainer's hand, the pinger was dipped into the water and the animal was rewarded for approaching it.

Gradually, the animal came to the pinger whenever it was put into the water.

Further, hearing of SIB was tested by evoked potential audiometry (Ridgway *et al.*, 1981). Tones and clicks at various intensities, repetition rates, and durations were presented via the same hydrophones mentioned above and positioned both 1 m in front of the animal and adjacent to the lower jaw, or attached by suction cup to the lower jaw (Moore *et al.*, 1995).

II. RESULTS

A. Responses of eight hearing dolphins of various ages

Results were obtained from the eight dolphins at various frequencies between 5 and 120 kHz (Fig. 1). At the baseline level of 111 dB re: 1 μPa, all dolphins responded at better than 90% correct responses to frequencies of 5, 10, 20, 40, and 50 kHz, with the exception of one old male, MAU, that dropped to just over 50% at 50 kHz, 111 dB. The results at frequencies of 60, 70, 80, 100, and 120 kHz varied considerably between the different animals. One female and three male dolphins under age 20 at the time of testing and two females over the age of 30 demonstrated a capability for responding to all the frequencies at a correct response rate over 90%, and most over 95%. All of the males over age 23 showed varying degrees of inability to respond to tones of 60 kHz, and above.

The degree of hearing deficit with respect to frequency varied somewhat in the three old males and the one old female that demonstrated a hearing deficit. One male, IAY, responded consistently to tones of 60 kHz but responded to no tones of 70 kHz, and higher even when St duration was increased to 450 ms. The single old female that demonstrated a hearing deficit, dolphin SLA, also had a sharp hearing cutoff but at a higher frequency of 100 kHz. Two older males had a more gradual or incomplete hearing deficit. At 70 kHz, MKA responded to $>$75% at 135 dB and $>$50% at 129 dB but was $<$5% at 111 dB. At 80, 100, and 120 kHz, his correct response level dropped to less than 5% (near false alarm rate) at all intensities under 135 dB re: 1 μPa, and at this level his correct performance was just under 25%. Correct response level was not increased significantly when tone duration was extended to 300 ms.

Figure 2 shows thresholds at the higher frequencies of a male dolphin age 9 (Salty) studied by Johnson (1967) compared with IAY at age 13 (Ljungblad *et al.*, 1982), and our findings on IAY at age 26 when the dolphin failed to respond to tones 40–50 dB above his threshold established by Ljungblad *et al.* (1982) 13 years earlier.

B. Behavioral observations of the deaf dolphin (SIB)

The first unusual behavior was noticed soon after SIB was brought to our facility in San Diego Bay. We noticed that when SIB was apparently asleep, she adopted a posture that was different from any dolphin we had ever observed. We called this a ''spar buoy'' posture since the dolphin's rostrum was pointed straight overhead, and its tail hung straight down as the animal bobbed in the water.

Most dolphins in our program are trained to respond to a pinger or other acoustic device. This facilitates movement of animals around our dolphin pod complexes, and the pinger is used as a recall device when the animals are released in the bay or in the open sea. During initial training, soon after the dolphin was collected in the Mississippi Sound in 1984, SIB along with six other dolphins in her group appeared to respond normally when the pinger was dipped into the water. After the task was moved into the open bay, when SIB was away from other dolphins, and, especially as the distance over which the dolphin was required to respond was increased, trainers began to suspect that SIB was relying on vision instead of hearing the sound of the 45-kHz pinger. When SIB was separated from other dolphins in the group, and the pinger was inserted in such a way that the dolphin could not see the action, she did not respond.

C. Other tests of hearing and sound production for SIB

Next, our trainers tried to elicit sound from SIB by the methods mentioned above. Neither whistles or burst pulse sounds could be elicited. The only sounds made by SIB were low Bronx cheer like sounds as the nasal plug fluttered during forced exhalations through a partially open blowhole.

We had noticed that when dolphins are separated from their group, they sometimes increase the rate of vocalization, especially the production of whistles. Twice, SIB was placed in a portable netting enclosure 5×4×3 m and slowly moved away from the group in San Diego Bay. Sound was monitored continuously by hydrophones (B&K 8103 with a B&K charge amplifier, and a Racal tape recorder with a frequency response at least as high as 150 kHz) for 3 h during each period of separation. No whistles, burst pulses, or echolocation pulses were recorded.

Finally, we attempted the electrophysiological approach which we have applied in the past to screen hearing in more than a dozen dolphins (Seeley *et al.*, 1976; Ridgway, 1980; Ridgway *et al.*, 1981). With both tone and click presentations from 1–120 kHz from hydrophones attached to the lower jaw, near the lower jaw, or in the water in front of the dolphin, no auditory evoked potentials were obtained, even to stimuli as high as 141 dB *re*: 1 μPa.

III. DISCUSSION AND CONCLUSIONS

Humans underwater can hear very high frequency tones by bone conduction (Deatherage *et al.*, 1954; MacKay, 1984), but there is no pitch discrimination above 15 or 20 kHz or above that person's hearing range. It would be interesting to know if the two older male dolphins, MAU and MKA, that showed some responses to the highest intensity tones (135 dB), retained any pitch discrimination at the frequencies from 60–120 kHz.

Although two out of four of our dolphins with hearing deficits had been treated with aminoglycosides for infections during their many years with our program, the short course of treatment, and the presence of normal kidney function as

indicated by clinical screens, suggest to us that such treatment did not cause the high-frequency hearing loss we observed.

Because a high percentage of the human population (males more than females) show hearing loss with age (Ries, 1982), it should not be surprising that other mammals share this deficit. Although our older dolphins with high-frequency hearing loss produce echolocation pulses, we have not studied them in echolocation tasks. We suspect that echolocation requiring fine discrimination in the presence of noise would be degraded. Au (1993) has shown that in Kaneohe Bay, where background noise in the 20–100 kHz range is dominated by snapping shrimp, dolphins shift their echolocation click peak frequency above 100 kHz. Our old dolphins with high-frequency hearing deficits would likely be at a disadvantage in such an environment.

For our old dolphins, survival is not dependent on the use of echolocation in the sea. *Tursiops* do survive in the wild to advanced ages. One extreme example of a female estimated to be age 52 has been reported (Scott *et al.*, 1996). We suspect that high-frequency hearing loss may well be a consequence of dolphin aging in the wild as well.

Although dolphin hearing and echolocation characteristics have received much more attention than other sensory abilities (cf. Au, 1993), *Tursiops* has good vision, some chemoreception, and good tactile senses (Nachtigall, 1986). The sense of touch is especially well developed (Ridgway and Carder, 1993b). After we determined that SIB was deaf, it became apparent from observing the dolphin's behavior that she had become adept at employing the other dolphins in the group to derive information that the others all received by the acoustic sense. For example, when the recall pinger was placed in the water, she probably became immediately aware of it by observing the behavior of other dolphins. Only when SIB was removed from the immediate presence of other dolphins, and the pinger insertion hidden from view, did we determine that the dolphin could not hear the pings.

When SIB was collected from the Mississippi Sound (Cat Island near Gulfport, MS) in 1984, she was a robust and apparently healthy animal within the weight range expected for the population (Ridgway and Fenner, 1982), We suspect SIB was able to survive, and maintain good nutrition not only by using senses other than audition, but by observing other dolphins. The mutual survival benefits of dolphin schools have been discussed by several authors (Norris and Dohl, 1980; Connor and Norris, 1982; Bradbury, 1986; Würsig, 1986).

We showed that dolphin calves produce echolocation pulses by about 60 days of age (Carder and Ridgway, 1984); however, we have recorded shrill whistles from calves within ten minutes after birth. Because SIB produced none of the usual dolphin sounds, we suspect that she may have developed deafness near or even before birth.

We do not know whether the unique ''spar buoy'' resting and sleeping posture of SIB was related to deafness or vestibular dysfunction. We noted this unusual behavior at the outset; however, we did not immediately suspect deafness. We now surmise that the unusual posture may have been related to the deaf and mute condition. Among the possible causes for her condition are infections. Severe infections can damage the vestibular system as well as the cochlea. A generalized infection affecting the cranium and nasal sinuses such as meningitis could result in such damage. When these dolphins with hearing loss die, histologic examination may shed light on the cause of this deafness.

ACKNOWLEDGMENTS

We thank Tricia Kamolnick who was in charge of all the dolphin training for our high-frequency response experiments. William Root of San Diego State University designed our computer programs and Michelle Reddy assisted with manuscript preparation. Whitlow Au, Ted Cranford, Bernd Würsig, and Patrick Moore made helpful suggestions concerning the manuscript. This work was supported in part by the Office of Naval Research.

[1]Dukane Corporation SeaCom Division, 2900 Dukane Drive, St. Charles, IL 60174.

Au, W. W. L. (**1993**). *The Sonar of Dolphins* (Springer-Verlag, New York).

Au, W. W. L., Carder, D. A., Penner, R. H., and Scronce, B. L. (**1985**). ''Demonstration of adaptation in beluga whale echolocation signals,'' J. Acoust. Soc. Am. **77**, 726–730.

Anon (**1994**). ''Aminoglycosides,'' in *Drug Information*, edited by G. K. McEvoy (American Society of Hospital Pharmacists, Bethesda, MD), pp. 55–70.

Bradbury, J. W. (**1986**). ''Social complexity and cooperative behavior in delphinids,'' in *Dolphin Cognition and Behavior: A Comparative Approach*, edited by R. Schusterman, J. Thomas, and F. Wood (Erlbaum, Hillsdale, NJ), pp. 361–372.

Carder, D. A., and Ridgway, S. H. (**1983**). ''Apparent echolocation by a sixty-day-old dolphin, *Tursiops truncatus*,'' J. Acoust. Soc. Am. **74**, S74.

Connor, R. C., and Norris, K. S. (**1982**). ''Are dolphins reciprocal altruists?'' Am. Naturalist **119**, 358–374.

Deatherage, B. H., Jeffres, L. A., and Blodgett, H. C. (**1954**). ''A note on the audibility of intense ultrasonic sound,'' J. Acoust. Soc. Am. **25**, 582.

Hall, J. D., and Johnson, C. S. (**1971**). ''Auditory thresholds of a killer whale,'' J. Acoust. Soc. Am. **51**, 515–517.

Johnson, C. S. (**1967**). ''Sound detection thresholds in marine mammals,'' in *Marine Bioacoustics*, edited by W. Tavolga (Pergamon, New York), pp. 247–260.

Johnson, C. S. (**1968**). ''Relation between absolute threshold and duration of tone pulse in the bottlenosed porpoise,'' J. Acoust. Soc. Am. **44**, 965–967.

Ljungblad, D. K., Scoggins, P. D., and Gilmartin, W. G. (**1982**). ''Auditory thresholds of a captive Eastern Pacific bottle-nosed dolphin, *Tursiops* spp,'' J. Acoust. Soc. Am. **72**, 1726–1729.

Mackay, R. S. (**1984**). *Medical Images and Displays* (Wiley, New York).

Moore, P. W. B., Pawloski, D. A., and Dankiewicz, L. (**1995**). ''Interaural time and intensity difference thresholds in the Bottlenose dolphin (*Tursiops truncatus*),'' in *Sensory Systems of Aquatic Mammals*, edited by R. A. Kastelein, J. A. Thomas, and P. E. Nachtigall (De Spil, Woerden, The Netherlands), pp. 11–23.

Nachtigall, P. (**1986**). ''Vision, audition and chemoreception,'' in *Dolphin Cognition and Behavior: A Comparative Approach*, edited by R. Schusterman, J. Thomas, and F. Wood (Erlbaum, Hillsdale, NJ), pp. 79–113.

Norris, K. S., and Dohl, T. P. (**1980**). ''The structure and function of cetacean schools,'' in *Cetacean Behavior: Mechanisms and Processes*, edited by L. M. Herman (Wiley–Interscience, New York), pp. 211–261.

Richardson, J. W. (**1995**). ''Marine mammal hearing,'' in *Marine Mammals and Noise*, edited by J. W. Richardson, C. R. Greene, Jr., C. I. Malme, and D. H. Thompson (Academic, San Diego), pp. 205–240.

Ridgway, S. H. (**1980**). ''Electrophysiological experiments on hearing in odontocetes,'' in *Animal Sonar Systems*, edited by R. G. Busnel and J. F. Fish (Plenum, New York), pp. 483–493.

Ridgway, S. H. (**1983**). ''Dolphin hearing and sound production in health and illness,'' in *Hearing and Other Senses: Presentations in Honor of E. G. Wever*, edited by R. R. Fay and G. Gourevitch (The Amphora Press), Vol. 16, pp. 247–296.

Ridgway, S. H., and Carder, D. A. (**1993a**). ''High-frequency hearing loss in old (25+ years-old) male dolphins,'' J. Acoust. Soc. Am. **94**, 1830.

Ridgway, S. H., and Carder, D. A. (**1993b**). ''Features of dolphin skin with potential hydrodynamic importance,'' IEEE Eng. Med. Biol. **12**, 83–88.

Ridgway, S. H., and Carder, D. A. (**1988**). ''Nasal pressure and sound production in an echolocating white whale,'' *Delphinapterus leucas*, in *Animal Sonar: Processes and Performance*, edited by P. E. Nachtigall and P. W. B. Moore (Plenum, New York), pp. 53–60.

Ridgway, S. H., and Fenner, C. A. (**1982**). ''Weight-length relationships of wild-caught and captive Atlantic bottlenose dolphins,'' J. Am. Vet. Med. Assoc. **181**, 1310–1315.

Ridgway, S. H., Carder, D. A., Kamolnick, P. L., Skaar, D. J., and Root, W. A. (**1991**). ''Acoustic response times (RTs) for *Tursiops truncatus*,'' J. Acoust. Soc. Am. **89** 1967–1968.

Ridgway, S. H., Bullock, T. H., Carder, D. A., Seeley, R. L., Woods, D., and Galambos, R. (**1981**). ''Auditory brainstem response in dolphins,'' Proc. Natl. Acad. Sci. (USA) **78**, 1943–1947.

Ries, P. W. (**1982**). ''Hearing ability of persons by sociodemographic and health characteristics in the United States (Series 10, No. 140),'' National Center for Health Statistics (U. S. Government Printing Office, Washington, DC).

Scott, M. D., Wells, R. S., and Irvine, A. B. (**1996**). ''Long-term studies of bottlenose dolphins in Florida,'' IBI Rep. **6**, 73–81.

Seeley, R. L., Flanigan, Jr., W. F., and Ridgway, S. H. (**1976**). ''A technique for rapidly assessing the hearing of the bottlenosed porpoise, *Tursiops truncatus*,'' Naval Undersea Center, NUCTP **522**, 1–15.

Würsig, B. (**1986**). ''Delphinid foraging strategies,'' in *Dolphin Cognition and Behavior: A Comparative Approach*, edited by R. Schusterman, J. Thomas, and F. Wood (Erlbaum, Hillsdale, NJ), pp. 347–359.

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.

Acoustic scattering from a rectangular aperture in a thick hard screen

Hyun H. Park and Hyo J. Eom
Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, 373-1, Kusong Dong, Yusung Gu, Taejon, Korea

(Received 29 January 1996; accepted for publication 13 September 1996)

Acoustic scattering from a rectangular aperture in a thick hard screen is examined. The Fourier transform is used to represent the scattered wave in the spectral domain and the boundary conditions are enforced to represent a solution in closed form. Numerical computations are performed to illustrate the behavior of the scattered wave from a thick rectangular aperture. The solution is represented in series which are numerically efficient. © *1997 Acoustical Society of America.* [S0001-4966(97)04801-7]

PACS numbers: 43.20.Fn [ANN]

INTRODUCTION

Acoustic wave scattering from a rectangular aperture in a thick hard screen has been studied with an approximate technique[1] using the radiation impedance concept. Although the approximate solution in Ref. 1 fairly well agrees with the measurement data in the low-frequency regime, it is also of interest to obtain a more rigorous exact solution. The motivation of the present study is to develop such a solution by using the Fourier transform and the mode matching. The method of the Fourier transform and the mode matching has been used in Refs. 2–4 to study electromagnetic wave scattering from an infinitely long rectangular aperture or a channel in a conducting plane. The solution presented in this paper is a simple, convergent series so that it is not only exact but also computationally very efficient. In the next section, we represent the scattered wave in the spectral domain and enforce the appropriate boundary conditions. In Sec. II, we perform the numerical calculations for the scattered field and the transmission coefficient. A brief summary is given in Sec. III.

I. FIELD ANALYSIS

Consider a rectangular aperture in a thick hard screen in Fig. 1. In region (I) ($z>0$), an incident field (velocity potential) Φ^i impinges on a rectangular aperture. The wave number is $k(=2\pi/\lambda,\ \lambda$: wavelength) and $e^{-i\omega t}$ time-harmonic convention is suppressed. In region (I) the total fields consist of the incident, reflected, and scattered as

$$\Phi^i(x,y,z)=e^{ik_x x+ik_y y-ik_z z}, \tag{1}$$

$$\Phi^r(x,y,z)=e^{ik_x x+ik_y y+ik_z z}, \tag{2}$$

$$\Phi^s(x,y,z)=\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\tilde{\Phi}^s(\zeta,\eta)e^{-i\zeta x-i\eta y+i\kappa z}d\zeta\ d\eta, \tag{3}$$

where $k_x=k\cos\phi\sin\theta$, $k_y=k\sin\phi\sin\theta$, $k_z=k\cos\theta$, $\kappa=\sqrt{k^2-\zeta^2-\eta^2}$, and $\tilde{\Phi}^s(\zeta,\eta)$ is the Fourier transform of $\Phi^s(x,y,0)$. In region (II) ($-d<z<0, |x|\leqslant a, |y|\leqslant b$) the total field is

$$\Phi^d(x,y,z)=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}[c_{mn}\cos\xi_{mn}(z+d)+d_{mn}\sin\xi_{mn}(z+d)]\times\cos a_m(x+a)\cos b_n(y+b), \tag{4}$$

where

$$a_m=\frac{m\pi}{2a},\quad b_n=\frac{n\pi}{2b},\quad \text{and}\quad \xi_{mn}=\sqrt{k^2-a_m^2-b_n^2}.$$

In region (III) ($z<-d$) the transmitted field is

$$\Phi^t(x,y,z)=\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\tilde{\Phi}^t(\zeta,\eta)e^{-i\zeta x-i\eta y-i\kappa(z+d)}\ d\zeta\ d\eta. \tag{5}$$

To determine the unknown coefficients c_{mn} and d_{mn}, we enforce the boundary condition on the field continuities.

First, we enforce the velocity continuity condition on the upper boundary at $z=0$:

 0001-4966/97/101(1)/595/4/$10.00

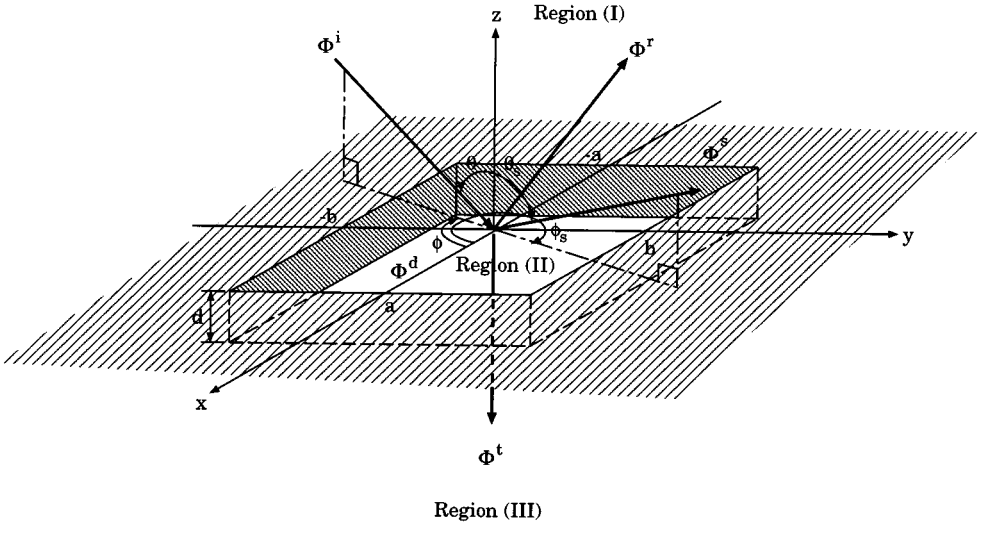


FIG. 1. Three-dimensional acoustic scattering and transmission by rectangular aperture in a thick hard screen.

$$\frac{\partial[\Phi^i(x,y,z)+\Phi^r(x,y,z)+\Phi^s(x,y,z)]}{\partial z}\Big|_{z=0}$$

$$=\begin{cases}\left.\dfrac{\partial\Phi^d(x,y,z)}{\partial z}\right|_{z=0} & \text{for } |x|<a,\ |y|<b,\\ 0, \quad \text{otherwise .}\end{cases} \tag{6}$$

Applying the Fourier transform to Eq. (6), we obtain

$$\tilde{\Phi}^s(\zeta,\eta)=-i(ab)^2\frac{\zeta\eta}{\kappa}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\xi_{mn}[c_{mn}\sin(\xi_{mn}d) - d_{mn}\cos(\xi_{mn}d)]F_m(\zeta a)F_n(\eta b), \tag{7}$$

where

$$F_m(u)=\frac{(-1)^m e^{iu}-e^{-iu}}{(u)^2-(m\pi/2)^2}. \tag{8}$$

Second, we enforce the pressure continuity condition on the upper boundary at $z=0$:

$$\Phi^i(x,y,0)+\Phi^r(x,y,0)+\Phi^s(x,y,0)=\Phi^d(x,y,0) \quad \text{for } |x|<a,\ |y|<b. \tag{9}$$

Substituting Eqs. (1) through (4) into (9), and multiplying by $\cos a_p(x+a)\cos b_q(y+b)dx\,dy$ $(p,q=0,1,2,3,\ldots)$, and performing integration yields

$$\gamma_{pq}+\frac{i(ab)^3}{(2\pi)^2}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\xi_{mn}[c_{mn}\sin(\xi_{mn}d) - d_{mn}\cos(\xi_{mn}d)]I_{mnpq} =\varepsilon_p\varepsilon_q[c_{pq}\cos(\xi_{pq}d)+d_{pq}\sin(\xi_{pq}d)], \tag{10}$$

where $\varepsilon_0=2,\ \varepsilon_1=\varepsilon_2=\cdots=1$, and

$$\gamma_{pq}=-2abk_xk_yF_p(k_xa)F_q(k_yb), \tag{11}$$

$$I_{mnpq}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{(\zeta\eta)^2}{\kappa}F_m(\zeta a)F_p(-\zeta a)F_n(\eta b) \times F_q(-\eta b)d\zeta\,d\eta. \tag{12}$$

An approximate, analytic evaluation of I_{mnpq} is formulated in Refs. 5 and 6; we, however, present a more efficient numerical scheme for I_{mnpq}. Thus

$$I_{mnpq}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{(\zeta\eta)^2}{\kappa}F_m(\zeta a)F_p(-\zeta a)F_n(\eta b) \times F_q(-\eta b)d\zeta\,d\eta =\int_0^{\infty}\rho d\rho\int_0^{\pi/2}d\alpha\frac{P(\zeta,\eta)}{\sqrt{1-\rho^2}}, \tag{13}$$

where $\zeta/k=\rho\cos\alpha$, $\eta/k=\rho\sin\alpha$, and

$$P(\zeta,\eta)=\frac{16k(\zeta\eta)^2}{(ab)^4}\frac{[1-(-1)^m\cos(2\zeta a)]}{(\zeta^2-a_m^2)(\zeta^2-a_p^2)} \times\frac{[1-(-1)^n\cos(2\eta b)]}{(\eta^2-b_n^2)(\eta^2-b_q^2)}.$$

This integral is approximated as[7]

$$I_{mnpq}\approx\sum_{i=1}^{M}\sum_{j=1}^{M}\frac{P(\zeta_{ij},\eta_{ij})}{\sqrt{1+\rho_{ij}}}\Delta\alpha\int_{\rho_{ij}-\Delta\rho/2}^{\rho_{ij}+\Delta\rho/2}\frac{\rho d\rho}{\sqrt{1-\rho}}, \tag{14}$$

where (ζ_{ij},η_{ij}) represents a center of the subdivided region $|\rho-\rho_{ij}|\leqslant\Delta\rho/2$, $|\alpha-\alpha_{ij}|\leqslant\Delta\alpha/2$, and

$$\int_{\rho_1}^{\rho_2}\frac{\rho d\rho}{\sqrt{1-\rho}}=\begin{cases}-\frac{2}{3}(\rho+2)\sqrt{1-\rho}\Big|_{\rho_1}^{\rho_2} & \text{for } \rho_2\leqslant 1,\\ -i\,\frac{2}{3}(\rho+2)\sqrt{\rho-1}\Big|_{\rho_1}^{\rho_2} & \text{for } \rho_1\geqslant 1.\end{cases} \tag{15}$$

Good convergence is obtained by truncation at $\rho=20$. The boundary conditions at $z=-d$ imply, similarly,

$$\frac{i(ab)^3}{(2\pi)^2}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\xi_{mn}d_{mn}I_{mnpq}=\varepsilon_p\varepsilon_qc_{pq}. \tag{16}$$

From (10) and (16) we obtain a matrix equation for c_{mn} and d_{mn},

$$\begin{bmatrix}\Psi_1 & \Psi_2\\ \Psi_3 & \Psi_4\end{bmatrix}\begin{bmatrix}C\\ D\end{bmatrix}=\begin{bmatrix}\Gamma\\ 0\end{bmatrix}, \tag{17}$$

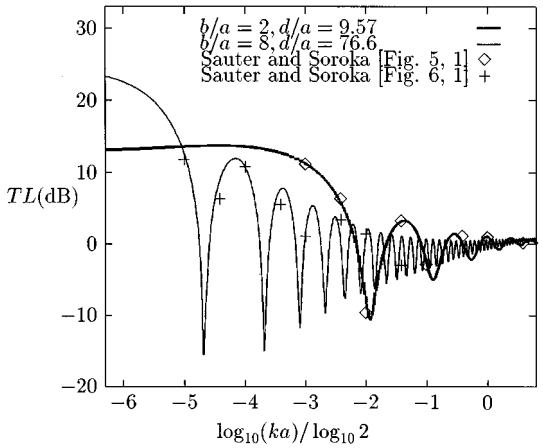


FIG. 2. Transmission loss TL ($=-10\log_{10}\tau$) versus normalized frequency $ka:\theta=\phi=0°$.

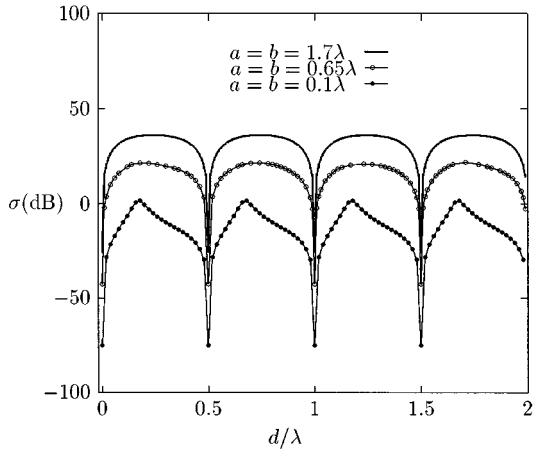


FIG. 4. Far-field pattern σ in the pit versus d/λ : $\theta=\phi=0°$.

where C, D, and Γ are column vectors of c_{mn}, d_{mn}, and γ_{pq},

$$\psi_{1,mnpq}=\varepsilon_p\varepsilon_q\cos(\xi_{pq}d)\,\delta_{mp}\delta_{nq}-\frac{i(ab)^3}{4\pi^2}\xi_{mn}I_{mnpq}\sin(\xi_{mn}d), \tag{18}$$

$$\psi_{2,mnpq}=\varepsilon_p\varepsilon_q\sin(\xi_{pq}d)\,\delta_{mp}\delta_{nq}+\frac{i(ab)^3}{4\pi^2}\xi_{mn}I_{mnpq}\cos(\xi_{mn}d), \tag{19}$$

$$\psi_{3,mnpq}=-\varepsilon_p\varepsilon_q\delta_{mp}\delta_{nq}, \tag{20}$$

$$\psi_{4,mnpq}=\frac{i(ab)^3}{4\pi^2}\xi_{mn}I_{mnpq}, \tag{21}$$

and δ_{mp} is the Kronecker delta. When $d\to\infty$, the solution is $C=(\Psi_1-i\Psi_2)^{-1}\Gamma$, $D=-iC$. When region (III) is filled with a hard screen (a rectangular pit), $C=\Psi_1^{-1}\Gamma$, $D=0$.

The reflection coefficient ρ and the transmission coefficient τ are shown to be

$$\rho=\frac{1}{4k_z}\,\mathrm{Im}\left\{\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\varepsilon_p\varepsilon_q\xi_{pq}^*[|c_{pq}|^2\sin^*(\xi_{pq}d)\times\cos(\xi_{pq}d)+c_{pq}^*d_{pq}|\sin(\xi_{pq}d)|^2-c_{pq}d_{pq}^*|\cos(\xi_{pq}d)|^2-|d_{pq}|^2\sin(\xi_{pq}d)\cos^*(\xi_{pq}d)]\right\}, \tag{22}$$

$$\tau=\frac{1}{4k_z}\,\mathrm{Im}\left\{\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\varepsilon_p\varepsilon_q\xi_{pq}^*c_{pq}d_{pq}^*\right\}, \tag{23}$$

where ()* and Im( ) denote the complex conjugate and the imaginary part of (). The far-field pattern[8] at the scattered angles θ_s and ϕ_s is given as

$$\sigma=\frac{k^2\cos^2\theta_s}{\pi}|\tilde{\Phi}^s(k\cos\phi_s\sin\theta_s,k\sin\phi_s\sin\theta_s)|^2. \tag{24}$$

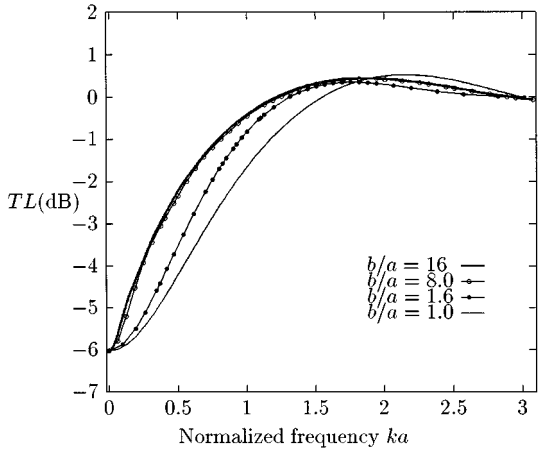


FIG. 3. Transmission loss TL ($=-10\log_{10}\tau$) versus normalized frequency $ka:d\to\infty$, $\theta=\phi=0°$.

II. NUMERICAL COMPUTATION

In order to check the accuracy of our formulation, we compute the transmission loss for a rectangular aperture in a thick hard screen using (23). Figure 2 shows the comparison between our results and Ref. 1, indicating favorable agreement. In order to achieve numerical convergence, the solution (17) should contain all propagation modes plus one or two evanescent modes. Figure 3 shows the behavior of the transmission loss for an infinite thickness aperture ($d\to\infty$) versus the normalized frequency. When $b/a>8$, the transmission loss remains almost insensitive to a change in b/a. Unlike Fig. 2 for a finite thickness aperture, the behavior of the transmission loss shows no oscillatory behavior versus the normalized frequency. Figure 4 shows the behavior of far-field pattern σ from a square pit with a bottom surface of a hard screen. It also shows that backscattering at normal incidence becomes maxima and minima alternatively with $d=\lambda/2$ periodicity.

III. CONCLUSION

Acoustic scattering from a thick rectangular aperture is studied using the Fourier transform and the mode matching technique. Our approach of the mode matching is restricted to problems of aperture shapes with separable geometries such as a rectangle. The solution is represented in series form which is numerically very efficient. The series solutions for scattering from a rectangular pit or semi-infinite pipe are obtained and their theoretical behaviors are investigated.

[1] A. Sauter, Jr. and W. W. Soroka, ''Sound transmission through rectangular slots of finite depth between reverberant rooms,'' J. Acoust. Soc. Am. **47**, 5–11 (1970).

[2] T. J. Park, S. H. Kang, and H. J. Eom, ''TE scattering from a slit in a thick conducting screen: Revisited,'' IEEE Trans. Antennas Propag. **42**, 112–114 (1994).

[3] S. H. Kang, H. J. Eom, and T. J. Park, ''TM-scattering from a slit in a thick conducting screen: Revisited,'' IEEE Trans. Microwave Theory Tech. **41**, 895–899 (1993).

[4] T. J. Park, H. J. Eom, and K. Yoshitomi, ''An analysis of transverse-electric scattering from a rectangular channel in a conducting plane,'' Radio Sci. **28**, 663–673 (1993).

[5] H. G. Davies, ''Low frequency random excitation of water-loaded rectangular plates,'' J. Sound Vib. **15**, 107–126 (1971).

[6] L. D. Pope and R. C. Leibowitz, ''Inter-modal coupling coefficients for a fluid-loaded rectangular plate,'' J. Acoust. Soc. Am. **56**, 408–415 (1974).

[7] K. Yoshitomi and H. R. Sharobim, ''Radiation from a rectangular waveguide with a lossy flange,'' IEEE Trans. Antennas Propag. **41**, 1398–1404 (1994).

[8] D. Feit, ''Sound radiation from orthotropic plates,'' J. Acoust. Soc. Am. **47**, 388–389 (1970).

High-frequency scattering from an acoustic cavity

Joel Garrelick
Cambridge Acoustical Associates, Inc., 200 Boston Avenue, Medford, Massachusetts 02155

(Received 29 February 1996; accepted for publication 9 August 1996)

High-frequency reradiation, or scattering, from a plate covered acoustic cavity is analyzed by assuming that the cavity geometry is sufficiently complex to produce a diffuse field. The plate is taken to be effectively infinite and the cavity baffled. For an illustrative example the cavity is shown to smear the return relative to the backscattered directivity pattern for the equivalent rigid plate alone, that is reduce the peak and fill in the nulls. Also, the effect of the plating above its flexural coincidence frequency is shown to redirect the peak return from normal incidence at lower frequencies to the vicinity of the coincidence angle. © *1997 Acoustical Society of America.* [S0001-4966(97)00901-6]

PACS numbers: 43.20.Fn, 43.30.Gv [ANN]

INTRODUCTION

The problem posed is sketched in Fig. 1. A plane wave is incident upon, and scatters from, an acoustic cavity that is covered with a plate and embedded in an effectively rigid infinite baffle. The cavity volume is geometrically complex, perhaps with obstacles that present additional scattering surfaces. However we limit our attention to what may be called the ''diffuse scattered field'' in the sense that the specular returns from these surfaces/structures are ignored. This picture may represent an idealization of scattering from a submerged fluid-filled tank structure. As an alternative to developing a detailed, e.g., finite element, model to capture this complexity, we present a less precise high-frequency asymptotic approach based on the assumption that the acoustic field within the volume is diffuse and therefore amenable to room acoustics techniques.[1] By high frequency we require that (1) $kL_c \gg 1$ and $kL \gg 1$, the characteristic dimensions of the cavity (L_c) and plate (L) are large in terms of acoustic wavelength ($\lambda = 2\pi/k$) for a diffuse acoustic field and to minimize the influence of edge diffraction and (2) $k_f L \gg 1$, the lateral plate dimensions are large in terms of its flexural wavelength ($\lambda_f = 2\pi/k_f$) to minimize the influence of individual resonances and boundary conditions.

Our mathematical model consists of a plane acoustic wave incident on a baffled, lossless, thin plate which serves as one surface of an otherwise lossless acoustic volume [Fig. 1(a)]. A similar geometry, namely a membrane-covered cylindrical cavity, is analyzed in Ref. 2 for small values of the ratio of fluid to membrane densities. Also, the problem of radiation from a sound source within a rectangular cavity is considered in Ref. 3 using the geometrical theory of diffraction.

Acoustic power is transmitted across the plating producing a diffuse field in the cavity. This field impinges on the underside of the plate as a spatially uniform distribution of plane waves. All of the power that enters the cavity is transmitted back across the plate and radiates. This reradiation is our scattered field.

II. ANALYSIS

A. Open cavity (transparent plating)

We first analyze the problem absent the plating. The input power is taken to be

$$P_i = S\overline{p_i^2}\cos(\theta_i)/2\rho c, \tag{1}$$

where $\overline{p_i}$ is the amplitude of the incident pressure wave, S the exposed surface area, and θ_i the incident angle measured from the surface normal.

Allowing for no energy loss in the volume from dissipation and/or coupling through other wall surfaces, all of the incident power enters the cavity and is subsequently reradiated. This assumption that radiation damping dominates provides an upper limit estimate for the scattered field. The effect of diffusion within the cavity is to redirect or scatter the return over the half-space. The projection of the diffuse field on the exposed surface, taken to be in the x-y plane, is envisioned as an incoherent set of plane waves each of form

$$p_d = \overline{p_d}\exp[ik(x\sin(\theta)\cos(\phi) + y\sin(\theta)\sin(\phi))] \tag{2}$$

uniformly distributed over the spherical coordinates θ and ϕ. For each such wave the power flowing across the surface is given by

$$P_d = S\overline{p_d^2}\cos(\theta)/2\rho c \tag{3}$$

and the total power is

$$P_{\text{rad}} = \int\int P_d \sin(\theta) d\theta\, d\phi = \pi S\overline{p_d^2}/2\rho c. \tag{4}$$

Setting the input and radiated powers equal, that is $P_i = P_{\text{rad.}}$, we obtain

$$\overline{p_d} = \overline{p_i}[\cos(\theta_i)/\pi]^{1/2}. \tag{5}$$

Corresponding to each wave is the velocity component normal to the surface S given by

$$v_{\text{rad}} = (\overline{p_d}/\rho c)\cos(\theta)\exp[i(k_x x + k_y y)], \tag{6}$$

with

$$k_x = k\sin(\theta)\cos(\phi)$$

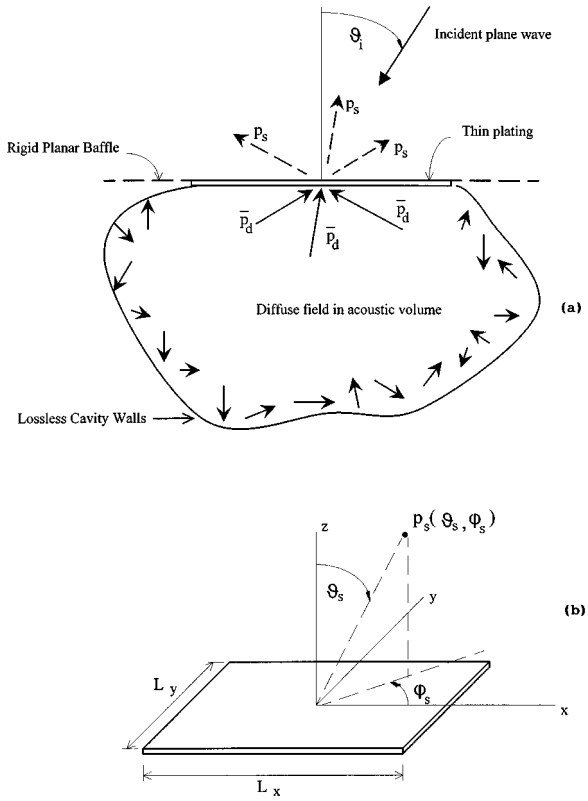


FIG. 1. (a) Insonification of, and scattering from, an acoustic volume with a diffuse field. (b) Plate geometry for scattered field calculation.

and

$$k_y = k\,\sin(\theta)\sin(\phi).$$

The associated far-field pressure radiated/scattered to (θ_s, ϕ_s) by this velocity field confined to the baffled surface S is given by [Fig. 1(b)]

$$|p_s(\theta_s, \phi_s; \theta, \phi)| = (kL_xL_y/2\pi R)|\rho c\, v_{\text{rad}} j_0(\beta_x) j_0(\beta_y)|, \tag{7}$$

with

$$\beta_x = (kL_x/2)[\sin(\theta)\cos(\phi) - \sin(\theta_s)\cos(\phi_s)],$$

$$\beta_y = (kL_y/2)[\sin(\theta)\sin(\phi) - \sin(\theta_s)\sin(\phi_s)],$$

and where R is the far-field range and $j_0(z)$ is the spherical Bessel function of argument z and order zero.[4] Equation (7) peaks at $\theta = \theta_s$ and $\phi = \phi_s$, although all wave components contribute to the pressure radiated to (θ_s, ϕ_s). The total squared pressure becomes

$$|p_{s(\text{tot})}(\theta_s, \phi_s)|^2 = \int\int |p_s(\theta_s, \phi_s; \theta, \phi)|^2 \sin(\theta)\, d\theta\, d\phi \tag{8}$$

or

$$|p_{s(\text{tot})}(\theta_s, \phi_s) R/\overline{p_i}|^2 = [A^2/\pi]\cos(\theta_i) I, \tag{9}$$

with

$$I = \int\int |\cos(\theta) j_0(\beta_x) j_0(\beta_y)|^2 \sin(\theta)\, d\theta\, d\phi \tag{10}$$

and where

$$A^2 = (kL_xL_y/2\pi)^2 \tag{11}$$

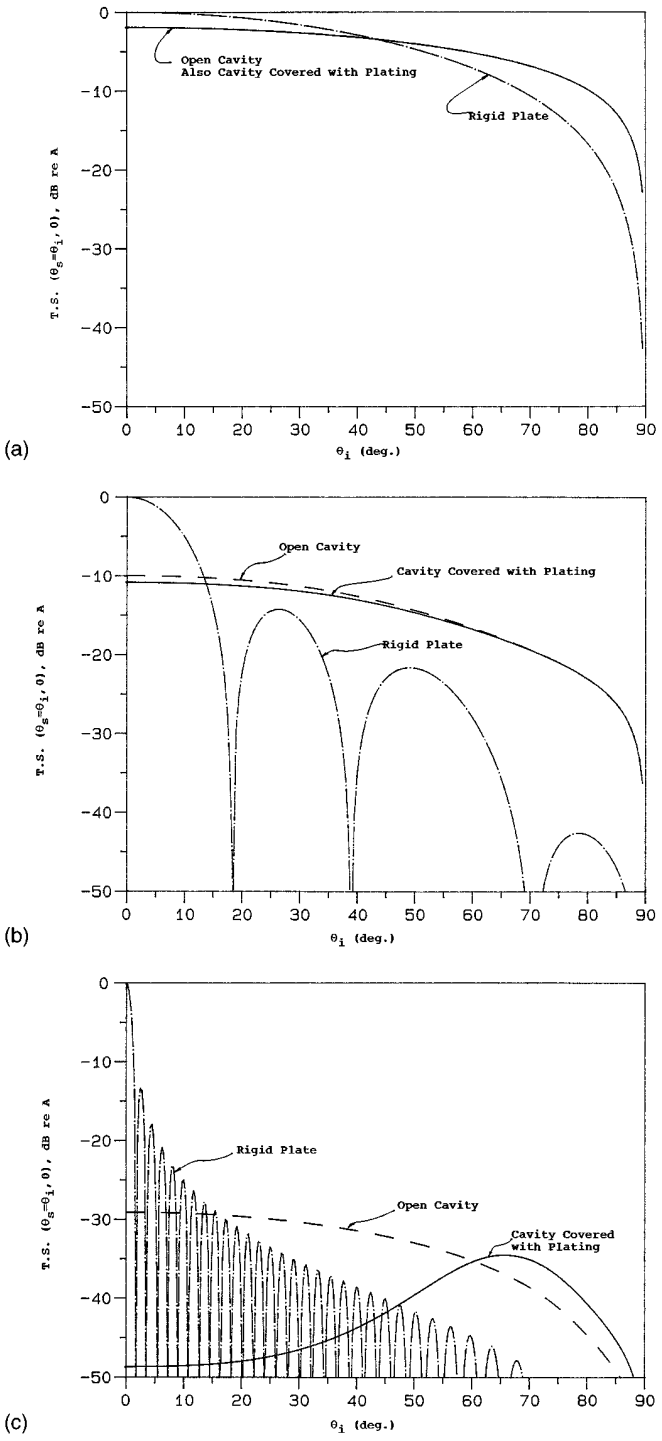


FIG. 2. (a) Backscattered target strength versus polar incidence angle: $kL = 1$. (b) Backscattered target strength versus polar incidence angle: $kL = 10$. (c) Backscattered target strength versus polar incidence angle: $kL = 10^2$.

is recognized[5] as the geometric acoustics, or Kirchhoff, expression for the specular return from the planar surface of area $S = L_xL_y$.

It must be noted that since this simplified approach invokes the Kirchhoff or physical acoustics high frequency approximation it is subject to its inconsistencies. Specifically, acoustic diffraction is ignored in connection with the incident field but is exhibited by the scattered field, to order $(kL)^{-2}$ generally or $(kL)^{-1}$ for scattering in the orthogonal

planes containing the plate edges. (Another manifestation of this approximation is that acoustic reciprocity is not strictly satisfied.[6]) This emphasizes the $kL \gg 1$ limitation on the analysis.

B. Cavity covered with thin elastic plating

We now allow for thin plating to separate the cavity from the acoustic half-space. For a plane-wave incident on submerged thin plating taken to be of infinite extent, the ratio of the transmitted to incident wave amplitudes is given by

$$t(\theta_i) = |\overline{p_t}/\overline{p_i}| = |1 - i\gamma[1 - (\omega/\omega_c)^2 \sin^4(\theta_i)]|^{-1}, \quad (12)$$

with

$$\gamma = \omega \rho_p h \cos(\theta_i)/2\rho c,$$

where ρ_p is the specific gravity of the plating, c_p is its sound speed and h its thickness, and $\omega_c = \sqrt{12}c^2/hc_p$ is the flexural coincidence frequency. With this transmission loss the power input to the cavity Eq. (1) now becomes

$$P_{i(\mathrm{plt})} = t^2(\theta_i) S \overline{p_i^2} \cos(\theta_i)/2\rho c. \quad (13)$$

The plating also affects the reradiated waves. No losses in the cavity have been assumed other than radiation through the surface S. Therefore to counter the lower radiated power implied by a finite transmission loss across the plating the pressure within the cavity must build up. Assuming the field remains perfectly diffuse, this requires that the enhanced pressure of the waves within the cavity, $\overline{p_{d(\mathrm{plt})}}$, satisfies the equation

$$P_{i(\mathrm{plt})} = [S/2\rho c] \int \int \overline{p_{d(\mathrm{plt})}^2} t^2(\theta) \cos(\theta) \sin(\theta) d\theta \, d\phi \quad (14a)$$

or, substituting Eq. (13),

$$(\overline{p_{d(\mathrm{plt})}}/\overline{p_i})^2 = \zeta^2(\theta_i) = t^2(\theta_i) \cos(\theta_i)/\chi \quad (14b)$$

with

$$\chi = \int \int t^2(\theta) \cos(\theta) \sin(\theta) d\theta \, d\phi.$$

The associated total scattered pressure is now

$$|p_{s(\mathrm{tot})}(\theta_s, \phi_s) R/\overline{p_i}|^2 = \zeta^2(\theta_i) A^2 I_t, \quad (15)$$

with

$$I_t = \int \int |t(\theta) \cos(\theta) j_0(\beta_x) j_0(\beta_y)|^2 \sin(\theta_s) d\theta \, d\phi. \quad (16)$$

Finally, using either Eqs. (9) or (15), we define a normalized (dB *re*: A) target strength as

$$\mathrm{T.S.}(\theta_s, \phi_s) = 10 \log |p_{s(\mathrm{tot})}(\theta_i, \phi_i) R/A\overline{p_i}|^2, \quad (17)$$

with the backscattered return given by $\theta_s = \theta_i$ and $\phi_s = \phi_i$.

III. ILLUSTRATIVE EXAMPLE

As an illustration of the above analysis we consider a square cavity, $L_x = L_y = L$, and a plating thickness such that $h/L = 5 \times 10^{-3}$. The ratios of plating to acoustic medium density and sound speed are for those of steel in water, $\rho_p/\rho = 7.8$ and $c_p/c = 3.5$, respectively. The propagation vector of the incident wave is taken to be in the x,z plane, that is $\phi_i = 0$.

Computed backscattered T.S. levels are plotted versus polar angle $\theta_s = \theta_i$ in Fig. 2 for three nondimensionalized frequencies $kL = 1$, 10, and 10^2 with the lowest value clearly stretching our high-frequency assumptions. As a reference, the flexural coincidence frequency of our plate is $k_c L = \sqrt{12}(c/c_p)/(h/L) \simeq 99$. Thus for example, $kL = 10$ ($k_f L = 31.5$) may refer to a frequency of 1 kHz, a cavity measuring $L = 2.4$ m on a side and a plate of thickness $h = 0.012$ m. At sample frequencies and incidence angles the total scattered power was computed using a one degree resolution to calculate the required integrals. In all cases it was found to be within 1 dB of the incident power and well within this value in most cases. This is a numerical check only since, by construction, they have been set equal.

In each figure three curves are presented showing the return with and without the plate and, for perspective, the equivalent return from the baffled plate alone if taken to be rigid. The latter is given by $20 \log |A j_0(\beta_x) j_0(\beta_y)|$ with $\theta = \theta_s$ and $\phi = \phi_s$. In Fig. 2(a) and (b) $\gamma = 0.039$ and 0.39, respectively, and the plate is effectively transparent.

It is concluded that the cavity moderates the directivity of the equivalent rigid plate both by reducing the main lobe peak and by filling in the nulls. This is also the case above coincidence [Fig. 2(c)]. However now the effect of the plating is strong as it redirects the peak in the scattered field to the vicinity of the coincidence angle $[\theta_c = \sin^{-1}(k_c/k)^{1/2} \cong 84°]$ which is preferred because of low transmission loss. At normal incidence the inertial impedance of the plating is considerable ($\gamma = 3.9$) and the T.S. is reduced by approximately 20 dB from that for the open cavity.

ACKNOWLEDGMENTS

The author gratefully acknowledges the reviewers for their helpful comments and the Naval Ship Weapons Center, Carderock, MD for its support.

[1] A. D. Pierce, *Acoustics* (McGraw-Hill, New York, 1981), p. 255.
[2] W. P. Baker, G. A. Kriegsmann, and E. L. Reiss, ''Acoustic scattering by baffled cavity-backed membranes,'' J. Acoust. Soc. Am. **83**, 423–432 (1988).
[3] P. Saha and A. D. Pierce, ''Geometrical theory of diffraction by an open rectangular box,'' J. Acoust. Soc. Am. **75**, 46–50 (1984).
[4] M. C. Junger and D. Feit, *Sound, Structures and their Interaction* (MIT, Cambridge, 1986), 2nd ed., p. 334.
[5] R. J. Urick, *Principles of Underwater Sound for Engineers* (McGraw-Hill, New York, 1967), p. 244.
[6] M. C. Junger, ''Formulation of short wavelength bistatically scattered fields in terms of monostatic returns,'' J. Acoust. Soc. Am. **95**, 3055–3058 (1994).

Estimation of temperature gradient effects on the normalized surface impedance of soils

Richard Raspet, James M. Sabatier, and W. Patrick Arnott[a)]
Department of Physics and Astronomy and National Center for Physical Acoustics, University of Mississippi, University, Mississippi 38677

(Received 13 November 1995; revised 14 August 1996; accepted 26 August 1996)

Ground impedance measurements are used for sound propagation predictions and to determine soil properties. Solar heating of the ground leads to significant temperature swings and gradients in the near surface soil. The equations of thermoacoustics are applied to estimate the magnitude of the temperature effects on the impedance and develop an approximate equation for the adjustment of measured impedance. Ambient temperature effects are shown to be significant; temperature gradient effects in soils are negligible. The theory is applicable to noise control applications where larger gradients may occur in sound absorbing materials. © *1997 Acoustical Society of America.* [S0001-4966(97)03912-X]

PACS numbers: 43.28.Fp, 43.55.Ev [LCS]

INTRODUCTION

Measurement of the surface impedance of the ground is important for the prediction of sound propagation over the ground.[1] Surface impedance measurements are also used to determine soil properties of agricultural grounds.[2,3]

Soils outdoors undergo wide temperature variations. The temperature profiles in the ground are governed by the heat input to the surface and the thermal properties of the soil. In many cases the temperature profiles are approximately linear down to the damping depth d in the soil (the $1/e$ length for the daily cycle). Surface temperatures can vary by as much as 30 K in a day.[4] Damping depths depend on the soil type and moisture content but generally range[5] from 7 to 15 cm.

Temperature gradients can have large effects on sound propagation in the boundary layer of solid surfaces. In resonance tubes, large temperature gradients can produce amplification of sound waves. The study of this effect is called ''thermoacoustics.''[6] Arnott *et al.*[7] explicitly demonstrated the connection between the literature and notation of sound propagation through porous media to thermoacoustics. In a subsequent paper, Arnott *et al.* measured the changes in impedance of a thermoacoustic stack as the temperature gradient was varied.[8] Significant, measurable changes in the impedance occurred for a 20-K change in temperature across a 4.0-cm-long porous stack.

In this paper, the theory of thermoacoustics is employed to investigate the possible effects of temperature and temperature gradients on the normalized surface impedance of soils. Section I presents the thermoacoustics equations and describes the adaptation of these equations to calculate the surface impedance. Section II develops an ambient temperature normalization factor for the fitted flow resistivity, then develops a calculation of the effect of temperature gradients on the surface impedance of ordinary soils. Section III develops an approximation for the effect of temperature gradients on ground impedance. Section IV contains a discussion of the results and conclusions.

I. APPLICATION OF THE THERMOACOUSTIC EQUATIONS TO GROUND IMPEDANCE

A thorough review of thermoacoustic research is contained in a paper by Swift.[6] Arnott *et al.*[7] demonstrated that thermoacoustics could be formulated in terms of previous porous media research. The notation of Ref. 7 will be followed in this paper.

Temperature gradients in porous materials have two principal effects; the gas properties are functions of temperature and the complex compressibility is modified due to the change in temperature of the porous media with respect to the gas as the gas is displaced by the sound wave. The relevant equations from Ref. 7 give a second-order differential equation for the complex acoustic pressure amplitude in a pore:

$$\frac{\rho(z)}{F(\lambda)}\frac{d}{dz}\left(\frac{F(\lambda)}{\rho(z)}\frac{d\hat{p}(z)}{dz}\right)+2\alpha(\lambda,\lambda_T)\frac{d\hat{p}(z)}{dz}+k_0^2(\lambda,\lambda_T)\hat{p}(z)=0, \tag{1}$$

where

$$\alpha(\lambda,\lambda_T)=\beta\frac{dT}{dz}\frac{1}{2}\left(\frac{F(\lambda_T)/F(\lambda)-1}{1-N_{\mathrm{pr}}}\right), \tag{2}$$

and

$$k_0^2(\lambda,\lambda_T)=\frac{\omega^2}{c^2}\frac{1}{F(\lambda)}[\gamma-(\gamma-1)F(\lambda_T)]; \tag{3}$$

and the complex average velocity amplitude in the pore in terms of the pressure amplitude gradient is

$$\hat{v}(z)=\frac{F(\lambda)}{i\omega\rho(z)}\frac{d\hat{p}(z)}{dz}. \tag{4}$$

Here, $\alpha(\lambda,\lambda_T)$ is a complex damping or gain parameter and $k_0(\lambda,\lambda_T)$ is the complex acoustic wave number. For an ideal gas $\beta=1/T$; this is appropriate for air filled pores. The pores

[a)]W. Patrick Arnott is at the Desert Research Institute, Atmospheric Sciences Center, P.O. Box 60220, Reno, NV 89506, and is an Adjunct Assistant Professor in the Department of Physics and Astronomy, University of Mississippi.

are assumed to have a circular cross section of radius R so that the shear wave number (λ) and thermal wave number (λ_T) are given by

$$\lambda = R\sqrt{\rho\omega/\eta}, \quad \lambda_T = R\sqrt{\rho\omega c_p/\kappa}, \tag{5}$$

where $\rho(z)$ is the gas density, ω the radial frequency of the sound wave, η the viscosity, κ the thermal conductivity, and c_p the heat capacity at constant pressure. Here, λ_T and λ are related by the square root of the Prandtl number $N_{\rm pr}$, $\lambda_T = \sqrt{N_{\rm pr}}\lambda$. The thermoviscous function for circular pores is

$$F(\lambda) = 1 - (2/\sqrt{i}\lambda)[J_1(\sqrt{i}\lambda)/J_0(\sqrt{i}\lambda)]. \tag{6}$$

Figure 5 of Ref. 7 shows that the choice of pore shape makes little gross difference in the value of $F(\lambda)$. These small differences are significant when optimizing the performance of thermoacoustic devices, but should be negligible in estimates of temperature effects on ground impedance. The relative independence of the wave number and impedance of uniform porous materials on pore shape factor is discussed in detail by Stinson and Champoux.[9]

The temperature dependence of ρ, λ, λ_T, and speed of sound, c, in terms of reference values at T_0 are[10]

$$\rho = \rho_0 T_0/T, \quad c = c_0(T/T_0)^{0.5}, \tag{7a}$$

$$\lambda = \lambda_0(T_0/T)^{0.9245}, \quad \lambda_T = \lambda_{T_0}(T_0/T)^{0.9245}. \tag{7b}$$

Values (7a) are directly from Ref. 10, pages 29–30, and (7b) are from a power law fit to Eq. (10-1.16a) of Ref. 10. The fact that the Prandtl number for air is approximately independent of temperature was also utilized.

In thermoacoustic studies where the temperature gradients can be as large as 2000 K/m, Eqs. (1) and (4) are usually solved by numerical integration. Atchley *et al.*[11] achieved good agreement with data by solving Eq. (1) as a wave equation with constant $F(\lambda)/\rho$, $\alpha(\lambda,\lambda_T)$, and $k_0(\lambda,\lambda_T)$ evaluated at the center of the stack. The change in the material property $F(\lambda)/\rho(z)$ with depth is large for the small pores typical of the ground, so the method of Ref. 11 is extended. Equation (1) can be rewritten for ideal gases as

$$\frac{d^2\hat{p}'(z)}{dz^2} + 2\alpha'(\lambda,\lambda_T)\frac{d\hat{p}(z)}{dz} + k_0^2(\lambda,\lambda_T)\hat{p}(z) = 0, \tag{8}$$

where

$$\alpha'(\lambda,\lambda_T) = \frac{1}{F(\lambda)}\left\{\frac{dF(\lambda)}{d\lambda}\frac{d\lambda}{dT} + \frac{1}{T}\left[\frac{F(\lambda_T) - N_{\rm pr}F(\lambda)}{(1-N_{\rm pr})F(\lambda)}\right]\right\}\frac{dT}{dz}, \tag{9}$$

includes the effect of the derivative of $F(\lambda)/\rho(z)$.

In this paper, soil is modeled as a homogenous semi-infinite half-space with an imposed linear temperature gradient to the damping depth d. Although the daily and yearly soil temperature variation is quite complicated, the temperature profiles can often be approximated by a linear gradient to fixed depth d. The solution to Eq. (8) is then approximated by

$$\hat{p}(z) = \hat{A}e^{ikz}, \tag{10}$$

where $\hat{A}$ is a complex constant and k is a complex constant evaluated at the surface. Substitution of this form into Eq. (8) yields

$$-k^2 + ik2\alpha'(\lambda,\lambda_T) + k_0^2(\lambda,\lambda_T) = 0 \tag{11}$$

with solutions

$$k^+ = i\alpha' + \sqrt{k_0^2 - \alpha'^2}, \quad k^- = i\alpha' - \sqrt{k_0^2 - \alpha'^2}. \tag{12}$$

Here, α' and k_0 are the gain parameter and complex wave number evaluated at the surface.

The treatment herein differs from Ref. 11 in three ways:

(i) The complex roots are evaluated from Eq. (11) directly yielding unequal wave numbers, k^+ and k^-.

(ii) The wave numbers are evaluated at the surface, not at $d/2$, since the surface impedance is most affected by surface properties for sound absorbing soils.

(iii) The effect of the changing material property $F(\lambda)/\rho(z)$ is included in Eqs. (8) and (9).

If the temperature gradient is zero, the treatment above recovers the usual porous media wave numbers. The velocity contributions corresponding to the two wave numbers can be calculated from Eq. (4):

$$\hat{v}^{\pm}(z) = [F(\lambda)/\omega\rho(z)]k^{\pm}\hat{A}e^{ik^{\pm}z}. \tag{13}$$

The boundary condition at depth d is that the surface impedance is that of a uniform semi-infinite porous media. To account for additional pore length due to the pores not being normal to the surface or not being straight, we introduce the tortuosity factor q, which yields a pore length of qd at depth d. This reduces the temperature and pressure gradients by $1/q$. Impedance matching then determines the ratios of amplitudes $\hat{A}^-/\hat{A}^+$ within the gradient layer:

$$\frac{i\omega\rho(d)}{F(\lambda)}\frac{[e^{ik^+qd} + \hat{A}^-/\hat{A}^+ e^{ik^-qd}]}{[ik^+e^{ik^+qd} + \hat{A}^-/\hat{A}^+ ik^-e^{ik^-qd}]} = \frac{\omega\rho(d)}{F(\lambda)k(d)} \equiv Z(d), \tag{14}$$

where both $F(\lambda)$ are evaluated at damping depth d. The solution to Eq. (14) is

$$\frac{\hat{A}^-}{\hat{A}^+} = e^{i(k^+ - k^-)qd}\frac{[1 - k^+/k(d)]}{[1 - k^-/k(d)]}. \tag{15}$$

The surface impedance of the media is then the ratio of $\hat{p}(0)/\hat{v}(0)$ modified to account for the porosity and tortuosity of the pores. Porosity, Ω, is the ratio of open pore area to total area. The tortuosity introduces an additional factor of q to the surface impedance. With these average medium modifications, the normalized surface impedance is given by

$$Z(0) = \frac{\omega q}{c(0)F(\lambda)\Omega}\frac{(1 + \hat{A}^-/\hat{A}^+)}{[k^+ + (\hat{A}^-/\hat{A}^+)k^-]}. \tag{16}$$

Here, $F(\lambda)$ is evaluated at the surface.

Equation (16) is derived for cylindrical pores with tortuosity q and porosity Ω. In ground impedance studies[9,12,13] it is usual to approximate $F(\lambda)$ and to evaluate the impedance in terms of the dc flow resistivity:

$$\sigma = 8q^2\omega\rho/\Omega\lambda^2. \tag{17}$$

TABLE I. Porous media properties.

	Pore radius (R)	Porosity (Ω)	Tortuosity factor (q)	dc flow resistivity at 293 K
Low	9×10^{-5} m	0.5	1.4	72×10^{3} N s m^{-4}
Medium	6×10^{-5} m	0.3	1.4	270×10^{3} N s m^{-4}
High	4×10^{-5} m	0.3	1.4	600×10^{3} N s m^{-4}

II. PREDICTION OF THE EFFECT OF TEMPERATURE GRADIENTS ON NORMALIZED GROUND IMPEDANCE

First, the effect of a uniform temperature change on the normalized surface impedance should be understood. For a uniform temperature the gradient terms in Eq. (1) drop out. The only temperature variation in the formula for the normalized surface impedance is in the gas properties contained in λ and λ_T. Note that the measured flow resistance must be adjusted by $\sigma=\sigma(T_0)(T/T_0)^{0.849}$ as developed from Eq. (17) with Eqs. (7a) and (7b), and c by Eq. (7a) to predict the temperature dependence of Eq. (16). A temperature increase leads to an increase in flow resistivity. A 30-K temperature difference over a year leads to a 10% change in flow resistivity. Surface impedance data of homogeneous soils taken over a wide range of surface temperatures should be normalized to a standard surface temperature.

Next, the dependence of impedance on the temperature gradient in the soil is investigated. The depth dependence of the temperature in the ground as the surface is heated by the sun is well understood.[5] Although the form of the variation is quite complex, the temperature profile near the surface can be approximated by a linear gradient down to the daily damping depth d which is on the order of 7–15 cm. The total daily temperature variation may be as large as 20 K over this distance for cultivated soils. As an estimate of the maximum effect, the impedance for three media corresponding to low, medium, and high impedance soils for a temperature change of 20 K in 10 cm was calculated using Eq. (16). The properties of the soil models are listed in Table I.

Figure 1 displays the dependence of the normalized ground impedance on the temperature gradient for low flow resistivity ground. The imaginary part of the impedance in Fig. 1 is multiplied by -1 for display purposes. The surface temperature is 300 K for all cases. A 200-K/m gradient corresponds to 300 K at the surface and 320 K at 10-cm depth. The largest effect is on the imaginary part of the impedance at very low frequencies. The maximum change in impedance due to the temperature gradient is on the order of 6%. The effects of realistic temperature gradients on the impedance of ordinary soils is negligible. The effect for the medium and high impedance soils is smaller than that displayed in Fig. 1.

Also shown in Fig. 1 is the effect of depth of the gradient layer on the impedance. Doubling the layer depth to 20 cm with the same gradient leads to a small increase in the imaginary part at very low frequencies, but it should be noted that a 40-K total change between the surface and the bottom of the layer is unrealistic. The small change in impedance when the depth is doubled indicates that a semi-infinite approximation may be used for analytic analysis of

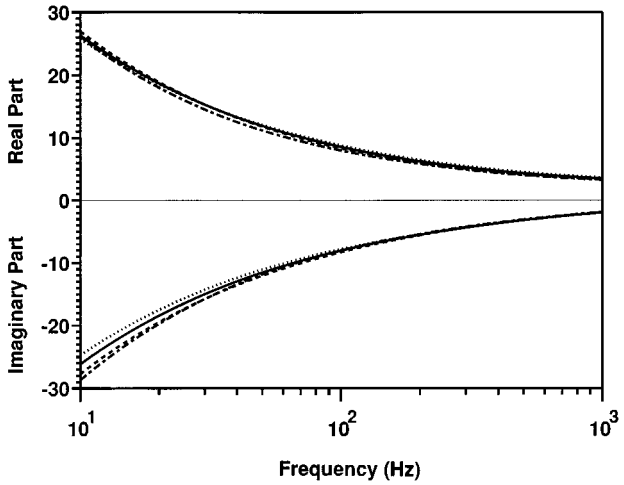


FIG. 1. Normalized surface impedance of the ground with no temperature gradient ———, a temperature gradient of 200 K/m for 10 cm – – –, a temperature gradient of -200 K/m for 10 cm ·····, a temperature gradient of 200 K/m for 20 cm — · —.

temperature gradient effects on impedance. In addition, the similarity of the impedance change to the effect of exponentially varying porosity[14] leads one to suspect that an equivalence between the temperature gradient and exponentially varying porosity can be established and an analytic approximation developed to determine if temperature gradient effects on impedance measurements are significant. This calculation is developed in Sec. III below.

III. APPROXIMATE FORMULATION FOR LOW FREQUENCIES

In this section an approximate expression is developed for the effect of a temperature gradient on the surface impedance of an homogeneous soil layer for low frequencies and for a semi-infinite temperature gradient. Approximations for $F(\lambda)$ have been used to develop expressions for the ground impedance of homogenous media for small thermal and viscous wave numbers.[12] For cylindrical pores

$$F(\lambda)=\left[\frac{4}{3}+\frac{i8}{\lambda^2}\right]^{-1} \tag{18}$$

and

$$k_0^2=\gamma\left(\frac{\omega}{c}\right)^2\left[\frac{4}{3}+\frac{\gamma-1}{\gamma}N_{\rm pr}+i\,\frac{8}{\lambda^2}\right]. \tag{19}$$

Further, $F(\lambda_T)$ can be written in terms of $F(\lambda)$ by noting $1-\sqrt{N_{\rm pr}}$ is small for air filled pores;

$$F(\lambda_T)=F(\sqrt{N_{\rm pr}}\lambda)$$
$$\cong F(\lambda)-\frac{dF(\lambda)}{d\lambda}(1-\sqrt{N_{\rm pr}})\lambda. \tag{20}$$

Substituting Eqs. (9), (19), and (20) in Eq. (8), using $\lambda=\lambda_0(T_0/T)^{\epsilon}$, and evaluating $dF(\lambda)/d\lambda$ in terms of $F(\lambda)$ results in

$$\frac{d^2\hat{p}}{dz^2}+\left[1-\left(\epsilon+\frac{1}{1+\sqrt{N_{\rm pr}}}\right)\frac{2}{[1-i(\lambda^2/6)]}\right]\frac{1}{T}\frac{dT}{dz}\frac{d\hat{p}}{dz}$$
$$+\gamma\left(\frac{\omega}{c_0}\right)^2\frac{T_0}{T}\left[\left(\frac{4}{3}-\frac{\gamma-1}{\gamma}N_{\rm pr}\right)+i\frac{8}{\lambda_0{}^2}\left(\frac{T}{T_0}\right)^{2\epsilon}\right]\hat{p}=0. \tag{21}$$

For low frequencies, λ is small and the terms $i\lambda^2/6$ and $\frac{4}{3}-(\gamma-1/\gamma)N_{\rm pr}$ can be dropped leaving:

$$\frac{d^2\hat{p}}{dz^2}+\left[1-2\left(\epsilon+\frac{1}{1+\sqrt{N_{\rm pr}}}\right)\right]\frac{1}{T}\frac{dT}{dz}\frac{d\hat{p}}{dz}$$
$$+i\gamma\left(\frac{\omega}{c_0}\right)^2\frac{8}{\lambda_0{}^2}\left(\frac{T}{T_0}\right)^{2\epsilon-1}\hat{p}=0. \tag{22}$$

Using Eq. (7), $1-2[\epsilon+1/(1+\sqrt{N_{\rm pr}})]=-1.925$ and $2\epsilon-1=0.849$. Note that the term containing $N_{\rm pr}$ is due to the thermoacoustic modification of the complex compressibility.

The substitutions $T=T_0(1+az)$ and $\zeta=(1+az)/a$, where a is the normalized temperature gradient in m^{-1}, lead to the differential equation and corresponding solution

$$\frac{d^2\hat{p}}{d\zeta^2}-\frac{1.925}{\zeta}\frac{d\hat{p}}{d\zeta}+k^2(0)a^{0.849}\zeta^{0.849}\hat{p}=0, \tag{23}$$

$$\hat{p}(z)=C\zeta^{1.46}H^{(1)}_{1.025}(\mathscr{K}\zeta^{1.425}), \tag{24}$$

where $\mathscr{K}=k(0)a^{0.4245}/1.425$ and C is a complex constant. Here, $H^{(1)}_{1.025}$ is chosen since we have assumed a semi-infinite gradient and radiation boundary conditions apply, i.e., $H^{(2)}_{1.025}$ becomes infinite as the depth increases. With the substitution $y=\mathscr{K}\zeta^{1.425}$, Eq. (24) is similar to Eq. (15) of Raspet *et al.*[14] and the normalized surface impedance can be derived from $\hat{p}(z)$ by following the procedure outlined in Ref. 14 yielding

$$z(0)=z_0\left[1+\frac{i1.5}{2k(0)}\frac{a}{q}\right]\cong z_0+\frac{i}{\gamma\Omega}\frac{3}{4}\frac{c_0}{\omega}\frac{a}{q}. \tag{25}$$

The reduction in gradient along the pore due to tortuosity is introduced in Eq. (25). z_0 is the surface impedance of an isothermal, uniform medium at the surface temperature. Equation (25) employs the low-frequency approximation that $z_0/k(0)=c_0/\gamma\Omega\omega$. Comparison with the result for exponential pores [Eq. (22), Ref. 14] shows that the semi-infinite linear normalized temperature gradient along the pores a/q is equivalent to a normal exponential porosity gradient α with a grain shape factor $n'=1$. Comparison of Eq. (25) with the exact results for a 40-K change in 20 cm from Eq. (16) showed good agreement.

The equivalence of the normalized temperature gradient and the exponential pore gradient is useful in establishing the minimum significant gradient at audio frequencies. Reference 14 demonstrates that the exponential porosity gradient corrections are significant at 100 Hz only when $\alpha\geqslant 10$, a factor of 20 larger than the gradient assumed in this paper. The equations above apply to any porous material (including those used for noise control) with viscous wave number, λ, less than one [see Eq. (5)]. Equation (25) shows that the critical variable for significance of the temperature gradients in such materials is the fractional temperature change per wavelength $c_0/\omega(a)$ compared to the normalized impedance z_0.

IV. CONCLUSIONS

The effect of uniform temperature changes in the ground can be calculated as modifying the dc flow resistivity and density [see Eq. (7) and Sec. II]. Realistic temperature gradients in homogeneous grounds lead to negligibly small changes in only the imaginary part of surface impedance. Temperature increases in the ground are analogous to decreasing exponential porosity grounds, while temperature decreases are analogous to increasing porosity grounds. An approximate formula for the effect of a linear temperature gradient on the surface impedance of the ground has been developed and is used to demonstrate that the effects are small for soils heated by the sun. Temperature gradient effects are measurable for low impedance surfaces as indicated by Eq. (25) and by the results of Reference 8. However, agricultural soils have too high a surface impedance for temperature gradient effects to be measurable.

ACKNOWLEDGMENTS

This work is supported in part by the U.S. Department of Agriculture National Sedimentation Laboratory. This work is an extension of research on thermoacoustic refrigeration supported by the Office of Naval Research.

[1] J. E. Piercy, T. F. W. Embleton, and L. C. Sutherland, ''Review of noise propagation in the atmosphere,'' J. Acoust. Soc. Am. **61**, 1403–1418 (1977).
[2] J. M. Sabatier, H. Hess, W. P. Arnott, K. Attenborough, M. J. M. Römkens, and E. H. Grissinger, ''*In situ* measurements of soil physical properties by acoustical techniques,'' Soil Sci. Soc. Am. J. **54**, 658–672 (1990).
[3] H. M. Hess, K. Attenborough, and N. W. Heap, ''Ground characterization by short-range propagation measurements,'' J. Acoust. Soc. Am. **87**, 1975–1986 (1990).
[4] D. Kirkham and W. L. Powers, *Advanced Soil Physics* (Wiley–Interscience, New York, 1972), Chap. 10.
[5] D. Hillel, *Fundamentals of Soil Physics* (Academic, San Diego, CA, 1980), Chap. 12.
[6] G. W. Swift, ''Thermoacoustic engines,'' J. Acoust. Soc. Am. **84**, 1145–1180 (1988).
[7] W. P. Arnott, H. E. Bass, and R. Raspet, ''General formulation of thermoacoustics for stacks having arbitrarily shaped pore cross sections,'' J. Acoust. Soc. Am. **90**, 3228–3237 (1991).
[8] W. P. Arnott, H. E. Bass, and R. Raspet, ''Specific acoustic impedance measurements of an air-filled thermoacoustic prime mover,'' J. Acoust. Soc. Am. **92**, 3432–3434 (1992).
[9] M. R. Stinson and Y. Champoux, ''Propagation of sound and the assignment of shape factors in model porous materials having simple pore geometries,'' J. Acoust. Soc. Am. **91**, 685–695 (1992).
[10] A. D. Pierce, *Acoustics, An Introduction to Its Physical Principles and Applications* (Acoustical Society of America, Woodbury, NY, 1989).
[11] A. A. Atchley, H. E. Bass, T. J. Hofler, and H. T. Lin, ''Study of a thermoacoustic prime mover below onset of self-oscillation,'' J. Acoust. Soc. Am. **91**, 734–743 (1992).
[12] K. Attenborough, ''Acoustical characteristics of rigid fibrous absorbents and granular materials,'' J. Acoust. Soc. Am. **73**, 785–789 (1983).
[13] K. Attenborough, ''Ground parameter information for propagation modeling,'' J. Acoust. Soc. Am. **92**, 418–427 (1992).
[14] R. Raspet and J. M. Sabatier, ''The surface impedance of grounds with exponential porosity profiles,'' J. Acoust. Soc. Am. **99**, 147–152 (1996).

An adiabatic normal mode solution for the benchmark wedge

C. T. Tindle
Physics Department, University of Auckland, Auckland, New Zealand

Z. Y. Zhang
Defence Scientific Establishment, HMNZ Naval Base, Devonport, Auckland, New Zealand

(Received 31 October 1995; accepted for publication 24 July 1996)

An adiabatic normal mode solution for the benchmark wedge is described for both fluid and solid attenuating ocean bottoms. The continuous mode contribution is treated as a sum of leaky modes. Each trapped mode changes smoothly into a leaky mode as the water depth decreases. The Pekeris branch line integral can be ignored if an approximate mode normalization is used when the normal mode pole approaches the branch point. © *1997 Acoustical Society of America.* [S0001-4966(97)01912-7]

PACS numbers: 43.30.Bp [MBP]

INTRODUCTION

The benchmark problems[1] in underwater acoustics were defined by the research community in 1986 as test cases for the comparison of numerical codes because there are no exact solutions for nontrivial range-dependent problems. A number of different, high-accuracy, numerical solutions of the benchmark wedge problem have now been published.[1–6] In the present work we show that an adiabatic normal mode approach yields a useful approximate solution which gives physical insight into the propagation process.

An important feature of upslope propagation in shallow water is the disappearance of modes as the water depth decreases. This means that the energy associated with a discrete mode passes into the continuum and is rapidly lost from the water column. The simple adiabatic mode theory breaks down at this point because it assumes the energy of a given mode remains in that mode at all times.

In the present paper the disappearance of normal modes from the water column (usually called cutoff) is handled smoothly by treating the normal mode continuum in terms of leaky modes as described by Arvelo and Uberall.[7] The use of leaky modes enables the extension of the adiabatic solution to include attenuation and the possibility of a solid bottom.

I. NORMAL MODE SOLUTIONS

For a horizontally stratified and cylindrically symmetric ocean environment the normal mode solution for the acoustic pressure amplitude $p(r,z)$ as a function of range r and depth z for a harmonic point source at $r=0$ and $z=z_0$ can be written

$$p(r,z)=\left(\frac{2\pi}{r}\right)^{1/2} e^{i\pi/4}\sum_n U_n(z)U_n(z_0)k_n^{-1/2}\exp(ik_n r) + \text{BLI}, \tag{1}$$

where the pressure has unit amplitude one meter from the source. The function $U_n(z)$ is the normal mode function and the k_n are the eigenvalues of k, the wave number for horizontal propagation.

The branch line integral (BLI) must be carefully considered for upslope propagation. In the present work we will use the Pekeris[8] branch cut and an approximate mode normalization to allow omission of the BLI. With the Pekeris branch cut, the normal mode sum in Eq. (1) includes some leaky modes. Leaky modes correspond to rays at grazing angles greater than critical which are partly reflected and partly transmitted into the ocean bottom. A full discussion of leaky modes and choice of branch cut is given in Ref. 9.

To apply normal mode theory to a sloping bottom we assume the usual adiabatic approximation. The normal modes and eigenvalues become range dependent and are determined by the local water depth. The acoustic pressure is given by

$$p(r,z)=(2\pi)^{1/2}e^{i\pi/4}\sum_n U_n(0,z_0)U_n(r,z) \times\exp\left(i\int_0^r k_n(r')dr'\right)\Big/ D. \tag{2}$$

Various forms of the denominator D for the adiabatic approximation have been used but Porter[10] has recently shown that the most appropriate form is

$$D=\left[k_n(0)k_n(r)\int_0^r\left(\frac{1}{k(r')}\right)dr'\right]^{1/2}. \tag{3}$$

We now assume a two-layer model in which a homogeneous layer of water of density ρ_w and sound speed c_w lies over a homogeneous solid bottom of density ρ_b, compressional wave speed c_b, and shear wave speed c_s. This model has been described by Ellis and Chapman[11] and is readily extended to include attenuation of both compressional and shear waves.[12]

The normal mode functions can be written[11]

$$U_n(z)=N_n\sin(\gamma_n z) \tag{4}$$

and the normalization constant N_n is given in Ref. 11 [Eq. (B13)]. The parameter γ_n is the usual vertical wave number.

A. Mode normalization

In the present work we consider propagation in decreasing water depth and a given mode moves smoothly from being a trapped mode to becoming a leaky mode as it passes through cutoff, i.e., when $k_n=\omega/c_b$. We expect that the mode amplitude will change slowly as this occurs but unfor-

tunately the normalization constant N_n in Eq. (4) changes rather rapidly near cutoff (and would go to zero in the absence of attenuation). At the same time as this occurs, the Pekeris branch line integral has a maximum value because the pole associated with the normal mode is very close to the branch point. The combined contribution of the normal mode and the branch line integral has the expected smooth behavior as the mode passes through cutoff.

Kamel and Felsen[13] developed an adiabatic form of normal mode theory including the branch line integral which allows a smooth transition through mode cutoff. However, the method is not easily extended to more general problems. In the present paper we find it more convenient to introduce an approximate mode normalization which has the effect of giving smooth behavior and allowing us to neglect the contribution of the Pekeris branch line integral.

The approximate normalization is readily obtained from the idea of effective depth and the details can be found in Ref. 12. Reflection at the ocean bottom is simulated by placing a perfect reflector displaced into the bottom a distance Δh. The resulting isovelocity waveguide has depth $h+\Delta h$ and the mode normalization is given by

$$N_n=[2/(h+\Delta h)]^{1/2}. \tag{5}$$

As will be shown in the next section this normalization is smooth and slowly varying as a mode passes through cutoff.

II. APPLICATIONS TO THE BENCHMARK WEDGE

A. Fluid bottom

The shallow-water wedge known as ''Benchmark III'' was defined in Ref. 2. The wedge parameters are: $\rho_w=1\text{g/cm}^3$, $c_w=1500$ m/s, $\rho_b=1.5$ g/cm^3, $c_b=1700$ m/s, $\alpha_b=0.5$ dB/λ.

A point harmonic source of frequency 25 Hz is located at 100-m depth in water 200 m deep. The bottom slopes at 2.86° and the depth decreases to zero at the 4-km range. The bottom is a fluid in this case. The solid bottom expressions can be used for a fluid bottom by setting c_s small. The expressions in Ref. 11 are all well behaved and in the limit give the correct expressions for the fluid case. For a fluid bottom we set $c_s=0.1$ m/s.

As the water depth decreases the values of the normal mode eigenvalues change systematically and their loci in the complex wave-number plane are shown in Fig. 1. A logarithmic scale has been used on the imaginary axis in order to show the values more clearly. The dotted line is the Pekeris branch cut and it begins at $k_b=\omega/c_b+i\alpha_b=0.09240+0.00085i$.

The tracks for the first six modes are shown in the figure and their starting points are identified by mode number. The curves begin for a water depth of 200 m at range zero. The loci of the modes are shown as dashed lines after they pass under the branch cut and onto the second Riemann sheet. The leaky modes decay rapidly with range and their contribution to the pressure field decays rapidly after the pole moves above the branch point.

There are three modes trapped in the initial water depth and three of the mode paths begin to the right of the Pekeris

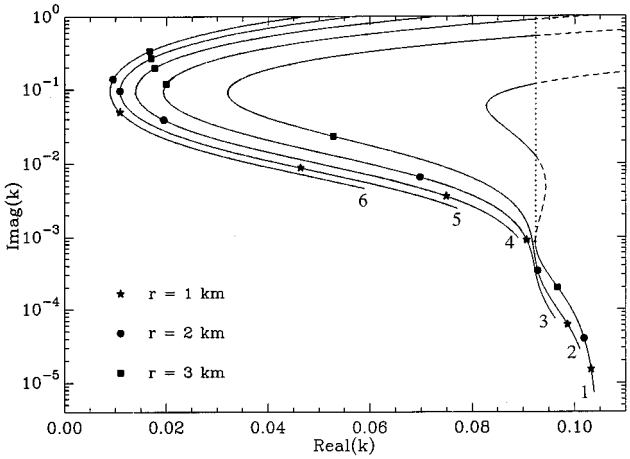


FIG. 1. Pole trajectories in the complex wave-number plane for the first six normal modes for the benchmark wedge.[1] The positions at zero range are identified by mode number and the symbols indicate positions at other ranges. The dotted line is the Pekeris branch cut.

cut. The symbols show the positions of the eigenvalues at various ranges. Mode 3 becomes a leaky mode at a range of about 0.8 km. Similarly modes 2 and 1 become leaky at ranges of 2.1 and 3.4 km, respectively. In the absence of attenuation the trapped mode eigenvalues would all be on the real axis and as range increases they would all move left along the real axis to the branch point. After passing the branch point they would become leaky modes and would move away from the real axis as they continued to move to the left.

The propagation loss calculated using the above adiabatic normal mode theory is shown in Fig. 2 for a receiver depth of 30 m. The three sections of the figure correspond to different ways of treating the mode normalization. The crosses are the same in each part and represent the coupled mode reference solution.[2] The data were kindly supplied by Finn Jensen.

In all three graphs the solid curves are found using the mode sum of Eq. (2). The first six modes are included and at the source position, three are trapped and three are leaky. The curves stop at a range of 3.4 km because the receiver meets the bottom and our expressions no longer apply.

The dashed curves in Fig. 2 are identified by mode number and show the contribution of individual normal modes to the total field. Each mode has been dropped from the mode sum when its amplitude is less than 1% of the amplitude of mode 1. The mode contributions successively become negligible until after a range of 2.8 km only mode 1 remains.

Figure 2(a) shows the result obtained using the exact mode normalization. The solid curve is an excellent fit to the reference solution up to a range of about 3.1 km. The abrupt deviation at this range is due to the rapid change of the normalization factor of mode 1 as the mode passes through cutoff. The dashed curves for the contributions of modes 2 and 3 show similar smaller sudden deviations at ranges of about 2.1 km and 0.8 km, respectively. In the absence of attenuation these deviations would be very large because the exact mode normalization would then go to zero at cutoff.

Figure 2(b) shows the result obtained using the approxi-

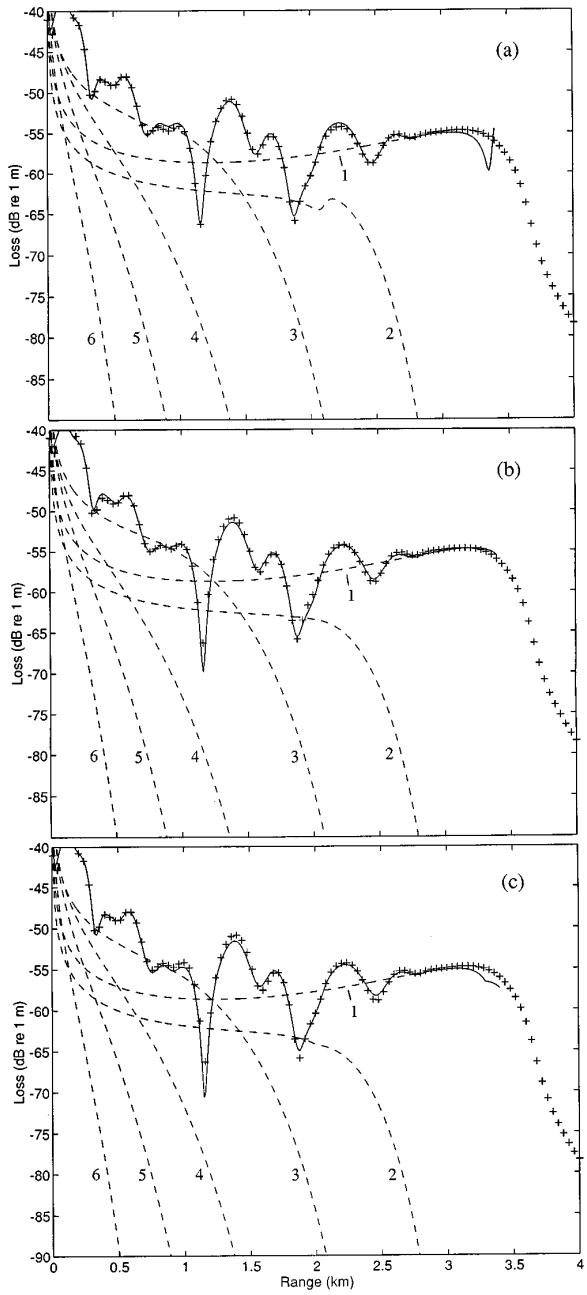


FIG. 2. Propagation loss for the benchmark 3 wedge.[1] The crosses show the reference solution. The solid curves show the adiabatic normal mode result. The dashed curves are identified by mode number and show the propagation loss for individual modes. (a) Exact mode normalization, (b) approximate normalization, (c) exact normalization changing to approximate normalization as the pole nears the branch point.

mate normalization of Eq. (5). It is an excellent fit to the reference solution. The sudden deviations associated with mode cutoff have been eliminated by the use of the approximate normalization.

In order to treat fluid and solid bottoms identically it is convenient to consider also the solution shown in Fig. 2(c). The curves in Fig. 2(c) were calculated using the exact normalization until the normal mode approaches cutoff and then switching over to the approximate normalization. The changeover is made when the distance of the normal mode eigenvalue from the branch point is some small fraction

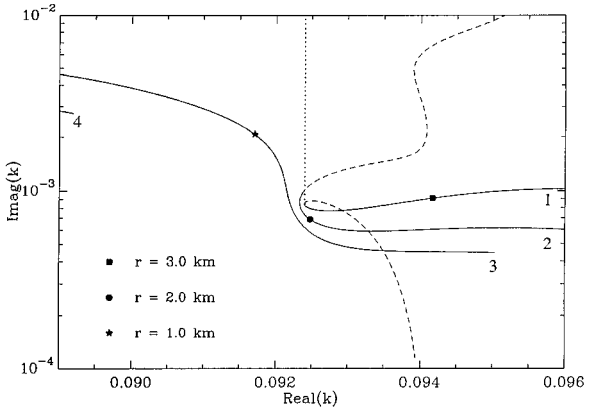


FIG. 3. Detail of pole trajectories near the branch point for an isovelocity wedge over a solid attenuating bottom.

(e.g., 5%) of the distance between the branch point and $k_w(=\omega/c_w)$ the wave number in the water. For Fig. 2(c) the changeover point was taken as when $|k_n-k_b|<0.05|k_w-k_b|$ and the mode contribution was scaled to ensure smooth matching.

The solid curve in Fig. 2(c) is in good agreement with the reference solution and shows that the field in the benchmark wedge can be adequately described in terms of adiabatic normal modes.

B. Solid bottom

The benchmark wedge was extended to include a solid bottom at the Parabolic Equation Workshop II.[14] The extra parameters are: $c_s=800$ m/s, $\alpha_s=0.5$ dB/λ. The paths of the normal mode poles in the complex plane are similar to those of Fig. 1. However, there are some differences of detail near the branch point and these are shown in Fig. 3 which shows a small part of the complex plane.

We note, in passing, that for a solid bottom with $c_s<c_w$ there is always energy loss to the bottom as shear waves and, strictly speaking, all modes are now leaky. Nevertheless, it is still convenient to regard modes with poles to the right of the Pekeris cut as trapped modes because their compressional wave amplitude decays exponentially in the bottom.

We note also that the Scholte wave is neglected in the present work. The contribution of the Scholte wave is small unless both source and receiver are near the interface.

The propagation loss as a function of range corresponding to Fig. 2 but for a solid bottom is shown in Fig. 4. As for Fig. 2 the three sections of the graph correspond to different ways of treating the mode normalization. The crosses in Fig. 4 are the same in all three sections of the figure and show the reference solution generated using a finite element model.[14] The numerical data were kindly provided by Stanley Chin-Bing.

The solid curve is the sum of the first six modes obtained using the mode sum of Eq. (2). As for the fluid bottom case, the first three modes are trapped in the water depth at the source but it is necessary to include some of the leaky modes to get a good approximation to the near field. The individual mode contributions are shown as the dashed curves and are identified by the corresponding mode number.

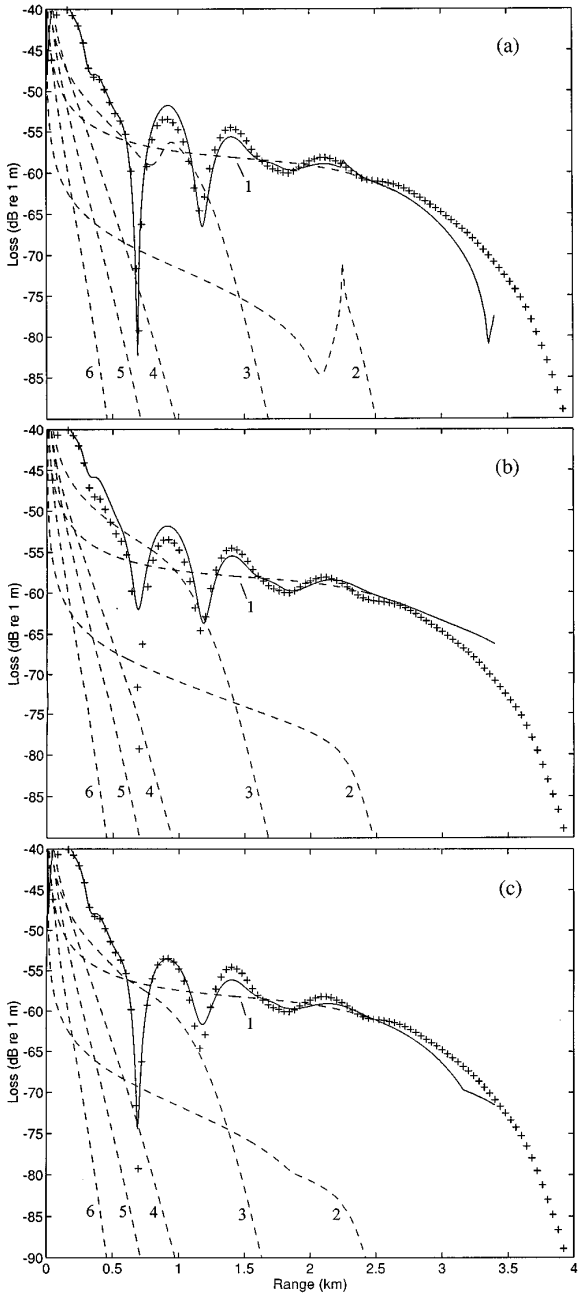


FIG. 4. Propagation loss for an isovelocity wedge over a solid attenuating bottom. The crosses show the reference solution. The solid curves show the adiabatic normal mode result. The dashed curves are identified by mode number and show the propagation loss for individual modes. (a) Exact mode normalization, (b) approximate normalization, (c) exact normalization changing to approximate normalization as the pole nears the branch point.

Figure 4(a) shows the result obtained using the exact normalization. The solid curve is in reasonable agreement with the reference solution out to a range of about 2.6 km. The dashed curves for modes 1–3 all show rapid changes in amplitude in the vicinity of their cutoff. For modes 1 and 2 there is an abrupt change in the mode amplitude as the normal mode pole passes close to the branch point and a corresponding abrupt change in the mode sum. This disagreement between the mode sum and the reference solution is due to the neglect of the branch line integral. We note again that if there were no attenuation the discontinuities in the mode amplitudes would be much greater because the exact normalization of would go to zero at cutoff.

Figure 4(b) shows the results obtained using the approximate mode normalization of Eq. (5). The mode amplitudes are now all smooth and the mode sum is shown by the solid curve.

A comparison of the solid curves in Fig. 4(a) and (b) shows that the exact normalization gives a better result in the range 0.3 to 1.3 km and is much better near the deep null at 0.7 km. Therefore it is an advantage to use a combination solution and the result is given in Fig. 4(c). The changeover point is as for Fig. 2(c).

III. DISCUSSION AND CONCLUSIONS

The above results show that the acoustic field in the benchmark problems can be described to good accuracy by adiabatic propagation of a small number of normal modes. For upslope propagation a mode which is initially trapped passes smoothly through cutoff and becomes a leaky mode. Even though the adiabatic normal mode description of the propagation is not as accurate as other techniques it reveals the simple physics of the situation.

ACKNOWLEDGMENTS

We are grateful to Finn Jensen and Stanley Chin-Bing for providing the reference solution data for Figs. 2 and 4, respectively.

[1] L. B. Felsen, ''Benchmarks: An opportunity for quality assessment,'' J. Acoust. Soc. Am. **87**, 1497–1498 (1990).
[2] F. B. Jensen and C. M. Ferla, ''Numerical solutions of range-dependent benchmark problems in ocean acoustics,'' J. Acoust. Soc. Am. **87**, 1499–1510 (1990).
[3] D. J. Thomson, ''Wide-angle parabolic equation solutions to two range-dependent benchmark problems,'' J. Acoust. Soc. Am. **87**, 1514–1520 (1990).
[4] M. D. Collins, ''Benchmark calculations for higher-order parabolic equations,'' J. Acoust. Soc. Am. **87**, 1535–1538 (1990).
[5] R. A. Stephen, ''Solutions to range-dependent benchmark problems by the finite difference method,'' J. Acoust. Soc. Am. **87**, 1527–1534 (1990).
[6] E. K. Westwood, ''Ray model solutions to the benchmark wedge problems,'' J. Acoust. Soc. Am. **87**, 1539–1545 (1990).
[7] J. I. Arvelo and H. Uberall, ''Adiabatic normal mode theory of sound propagation including shear waves in a range-dependent ocean floor,'' J. Acoust. Soc. Am. **88**, 2316–2325 (1990).
[8] C. L. Pekeris, ''Theory of propagation of underwater sound in shallow water,'' Geol. Soc. Am. Memoir **27**, 1–117 (1948).
[9] C. T. Tindle and Z. Y. Zhang, ''Continuous modes and shallow water sound propagation,'' *Oceans '93, Engineering in Harmony with the Ocean* (IEEE, New York, 1993).
[10] M. B. Porter, ''Adiabatic modes for a point source in a plane-geometry ocean,'' J. Acoust. Soc. Am. **96**, 1918–1921 (1994).
[11] D. D. Ellis and D. M. F. Chapman, ''A simple shallow water model including shear wave effects,'' J. Acoust. Soc. Am. **78**, 2087–2095 (1985).
[12] Z. Y. Zhang and C. T. Tindle, ''Complex effective depth of the ocean bottom,'' J. Acoust. Soc. Am. **93**, 205–213 (1993).
[13] A. Kamel and L. B. Felsen, ''Spectral theory of sound propagation in an ocean channel with weakly sloping bottom,'' J. Acoust. Soc. Am. **73**, 1120–1130 (1983).
[14] ''Test case 3,'' PE Workshop II, *Proceedings of the Second Parabolic Equation Workshop*, edited by S. A. Chin-Bing, D. B. King, J. A. Davis, and R. B. Evans (Naval Research Laboratory, Washington, DC, 1993), pp. 78–86.

The static-dynamic superposition method applied to uniaxial waves in a bi-periodic stack

Michael El Raheb
Dow Chemical Company, Midland, Michigan 48674

(Received 3 January 1996; accepted for publication 31 July 1996)

The method of El-Raheb [J. Acoust. Soc. Am. **94**, 172–184 (1993)] that treats transient uniaxial waves in a weakly coupled bi-periodic stack is modified in the way external force is applied. An inhomogeneous static solution and a homogeneous dynamic solution are superimposed. The modified method accurately predicts stress in the excited layer and converges rapidly with number of terms in the eigenfunction expansion. © *1997 Acoustical Society of America.* [S0001-4966(97)05811-0]

PACS numbers: 43.40.At, 43.20.Bi [CBB]

Reference 1 treats transient uniaxial waves in a finite stack of layered elastic media. Axial displacement and stress are expanded in eigenfunctions $\{\Phi_j(X)\}$ of the traction-free stack. The time-dependent traction applied along one boundary is approximated by a body force with a delta function dependence on the global axis X:

$$\sigma_b(X,t)=\delta(X)\sigma_0 f(t) \tag{1}$$

where σ_0 is the magnitude of the traction and $f(t)$ is its normalized time dependence. This representation is deficient in two ways:

(a) by definition, a body force acts on the volume of material and therefore stress vanishes on the boundary
(b) the expansion of $\delta(X)$ in terms of $\{\Phi_j(X)\}$ converges slowly producing error in stress close to the excited boundary.

An alternative relies on expressing the state vector of displacement $u(X,t)$ and stress $\sigma(X,t)$ as a sum of an inhomogeneous static solution and a homogeneous dynamic solution.[2] The static problem is forced by the magnitude σ_0 of the applied traction in (1). For a free stack, static equilibrium is maintained by including a constant body force opposing the external traction.

Consider a stack of N layers having (E_i,ρ_i,h_i) for moduli of elasticity, density, and thickness. Let h_T be total thickness of the stack:

$$h_T=\sum_{i=1}^{N} h_i. \tag{2}$$

For static equilibrium, the body force is related to σ_0 by

$$b_f=-\sigma_0/h_T. \tag{3}$$

Let x_i be the local axial coordinate of the ith layer. Static equilibrium requires

$$\frac{\partial\sigma_{si}}{\partial x_i}\equiv E_i\frac{\partial^2 u_{si}}{\partial x_i^2}=b_f \tag{4}$$

with solutions

$$\sigma_{si}(x_i)=b_f x_i+\sigma_{si}(0),\quad 0\leqslant x_i\leqslant h_i, \tag{5a}$$

$$u_{si}(x_i)=\frac{b_f}{E_i}\frac{x_i^2}{2}+\frac{\sigma_i(0)}{E_i}x_i+u_{si}(0). \tag{5b}$$

Enforcing boundary conditions

$$\sigma_{s1}(0)=\sigma_0,\quad \sigma_{sN}(h_N)=0 \tag{6}$$

on (5a) determines $\sigma_{si}(0)$ as

$$\sigma_{si}(0)=\sigma_0\left(1-\frac{1}{h_T}\sum_{k=1}^{i-1}h_k\right)\quad (i=2,N). \tag{7}$$

Noting that $u_{si}(h_i)=u_{s(i+1)}(0)$ in (5b) gives

$$u_{s(i+1)}(0)-u_{si}(0)=\frac{h_i}{E_i}\left[b_f\frac{h_i}{2}+\sigma_{si}(0)\right],\quad i=1,N. \tag{8}$$

Equation (8) supplies N equations in the $(N+1)$ unknowns $u_{si}(0)$; $i=1,N+1$, where for convenience, $u_{s(N+1)}(0)$ is used to refer to $u_{s(N)}(h_N)$. The missing equation comes from the constraint that the center of mass is motionless:

$$u_{scG}=0. \tag{9}$$

Taking moments about $X=0$ yields

$$u_{scG}\equiv\frac{1}{M_T}\sum_{i=1}^{N}\rho_i\int_0^{h_i}u_{si}(\xi)d\xi=0,$$

$$M_T=\sum_{i=1}^{N}\rho_i h_i. \tag{10}$$

Substituting (5b) in (10) produces

$$\sum_{i=1}^{N}\rho_i\left\{u_{si}(0)h_i+\frac{h_i^2}{2E_i}\left[b_f\frac{h_i}{3}+\sigma_{si}(0)\right]\right\}=0. \tag{11}$$

Any other constraint applied to (8) leads to ficticious rigid body motion while the forcing function acts.

In terms of static and dynamic variables, the general solution is

$$u(X,t)=u_s(X)f(t)+u_d(X,t), \tag{12a}$$

$$u_d(X,t)=\sum_j a_j(t)\Phi_j(X), \tag{12b}$$

$$\Phi_j(X)=\{\Phi_j^{(i)}(x_i)\}, \tag{12c}$$

where $f(t)$ is the function in (1) and $a_j(t)$ are generalized coordinates of the homogeneous dynamic problem satisfying

$$\ddot{a}_j(t)+\omega_j^2 a_j(t)=F_j(t),$$

$$F_j(t)=-\frac{1}{N_j}[N_{aj}f(t)+N_{sj}\dot{f}(t)], \tag{13}$$

$$N_{aj}\equiv\langle\Phi_j|1|b_f\rangle=b_f\sum_{i=1}^{N}\int_0^{h_i}\Phi_j^{(i)}(\xi)d\xi, \tag{13a}$$

$$N_{sj}\equiv\langle\Phi_j|\rho|u_s\rangle=\sum_{i=1}^{N}\rho_i\int_0^{h_i}\Phi_j^{(i)}(\xi)u_{si}(\xi)d\xi, \tag{13b}$$

$$N_j\equiv\langle\Phi_j|\rho|\Phi_j\rangle=\sum_{i=1}^{N}\rho_i\int_0^{h_i}[\Phi_j^{(i)}(\xi)]^2\,d\xi. \tag{13c}$$

Suppressing the common subscript j, the solution of Eqs. (13) takes the form

$$a(t)=\frac{1}{\omega}\dot{a}(0)\sin\omega t+a(0)\cos\omega t+\frac{1}{\omega}\int_0^t F(\tau)\times\times\sin\omega(t-\tau)d\tau, \tag{14}$$

$$\{a(0),\dot{a}(0)\}=-\frac{N_s}{N}\{f(0),\dot{f}(0)\}. \tag{14a}$$

To determine the integrals in (13a) and (13b), recall the expression for $\Phi^{(i)}(x_i)$ developed in Ref. 1:

$$\Phi^{(i)}(x_i)=B_{1i}\cos k_i x_i+\frac{B_{2i}}{E_i k_i}\sin k_i x_i,\qquad k_i=\omega/c_i,\quad c_i=(E_i/\rho_i)^{1/2}. \tag{15}$$

Substituting (5b) and (15) in (13a) and (13b) yields

$$N_a=b_f\sum_{i=1}^{N}\frac{1}{k_i}\left[B_{1i}\sin k_i h_i+\frac{B_{2i}}{E_i k_i}(1-\cos k_i h_i)\right] \tag{15a}$$

$$N_s=\sum_{i=1}^{N}\frac{\rho_i u_{si}(0)}{k_i}\left[B_{1i}\sin k_i h_i+\frac{B_{2i}}{E_i k_i}(1-\cos k_i h_i)\right]+\sum_{i=1}^{N}\frac{\rho_i\sigma_{si}(0)}{E_i k_i}\left\{B_{1i}\left[h_i\sin k_i h_i+\frac{1}{k_i}(\cos k_i h_i-1)\right]\right.$$
$$+\frac{B_{2i}}{E_i k_i}\left[-h_i\cos k_i h_i+\frac{\sin k_i h_i}{k_i}\right]\Bigg\}+\frac{b_f}{2}\sum_{i=1}^{N}\frac{\rho_i}{E_i k_i}\left\{B_{1i}\left[h_i^2\sin k_i h_i+\frac{2}{k_i}\left(h_i\cos k_i h_i-\frac{\sin k_i h_i}{k_i}\right)\right]\right.$$
$$\left.\left.+\frac{B_{2i}}{E_i k_i}\left[-h_i^2\cos k_i h_i+\frac{2}{k_i}\left(h_i\sin k_i h_i+\frac{(\cos k_i h_i-1)}{k_i}\right)\right]\right\}. \tag{15b}$$

For the special case when $f(t)$ is piecewise linear with n_c corner points

$$f(t)=\sum_{n=1}^{n_c-1}(f_n+s_n(t-t_n))[H(t-t_n)-H(t-t_{n+1})], \tag{16a}$$

$$\dot{f}(t)=\sum_{n=1}^{n_c-1}s_n[H(t-t_n)-H(t-t_{n+1})], \tag{16b}$$

$$\ddot{f}(t)=\sum_{n=1}^{n_c-1}s_n[\delta(t-t_n)-\delta(t-t_{n+1})]=\sum[(s_{n+1}-s_n)\delta(t-t_n)], \tag{16c}$$

where $s_n=(f_{n+1}-f_n)/(t_{n+1}-t_n)$ for $1\leqslant n\leqslant(n_c-1)$ and $f_1=f_{nc}=0$, $s_{n_c}=t_1=0$. After substituting (16a)–(16c) in (14) and noting that $f(0^-)=\dot{f}(0^-)=0$, the coefficient of each mode is determined to be

$$a(t)=-\frac{b_f N_a}{N\omega^2}\sum_{n=1}^{n_t}\left[(f_n+s_n(t-t_n))\cos\omega\zeta+s_n\left(-\zeta\cos\omega\zeta+\frac{1}{\omega}\sin\omega\zeta\right)\right]_{t-t_n}^{t-\hat{t}}$$
$$-\frac{N_s}{N\omega}\sum_{n=1}^{n_c}s_n[\sin\omega(t-t_n)H(t-t_n)-\sin\omega(t-t_{n+1})H(t-t_{n+1})], \tag{17}$$

where $\hat{t}=\min(t,t_{n+1})$ and n_t satisfies $t_{n_t}<t\leqslant t_{n_t+1}$.

Call the use of body force method 1, and call the static/dynamic superposition method 2. We now compare stress histories of the basic stack with six periodic sets discussed in Ref. 1, as computed according to the two methods. The stress history that appears is the normalized quantity $\tilde{\sigma}(t)=\sigma(t)/\sigma_0$. The first column in Fig. 1(a)–(e) was computed according to method 1 using five elastic modes per set ($m=5$), i.e., 30 modes for the stack. The histories in the second column of Fig. 1(f)–(j), also according to method 1, used 40 elastic modes per set ($m=40$), i.e., 240 modes for

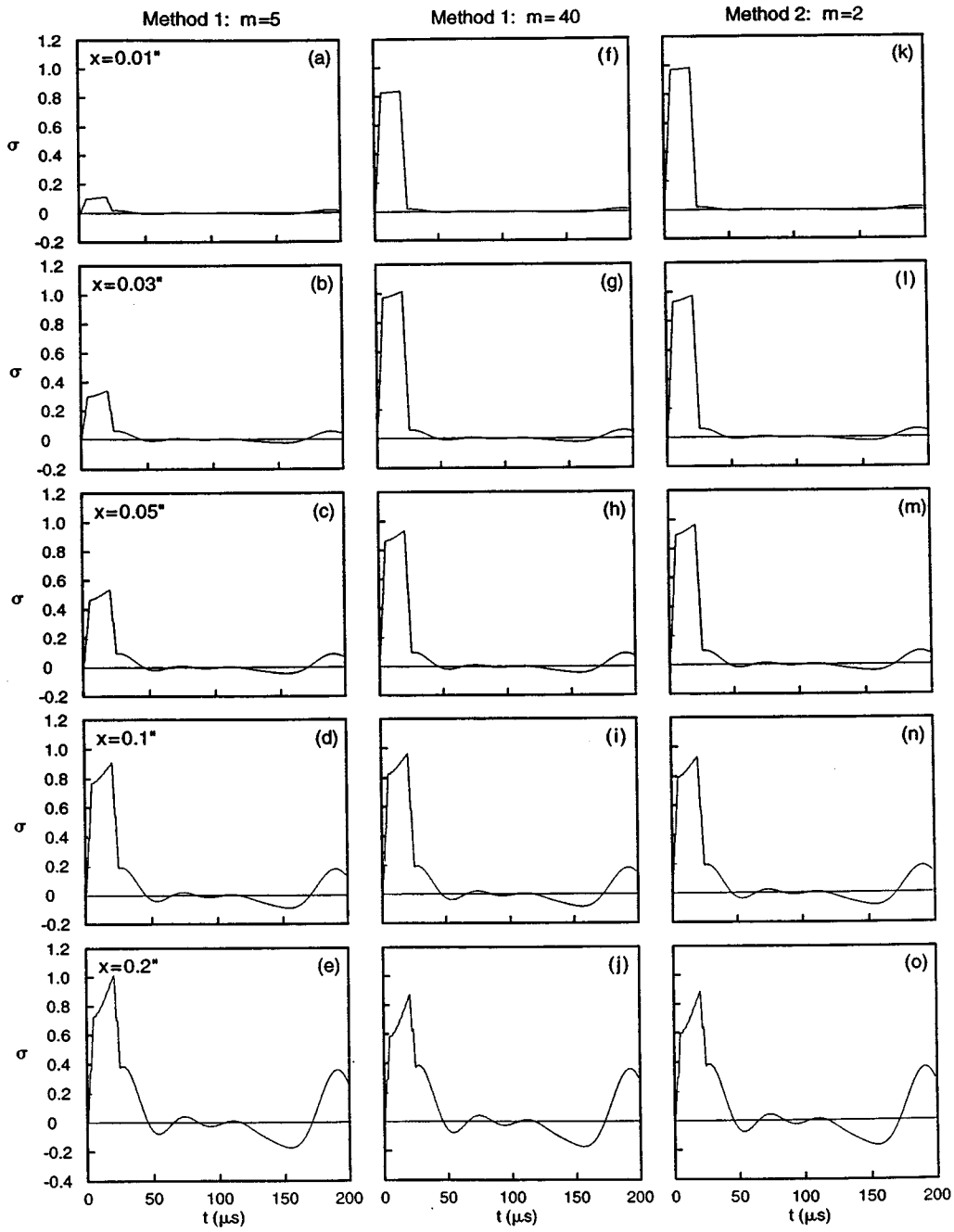


FIG. 1. Histories of $\sigma(t)$ along first layer: (a)–(e) method 1 with $m=5$; (f)–(j) method 1 with $m=40$; (k)–(o) method 2 with $m=2$.

the stack. Those in the third column in Fig. 1(k)–(o) were computed according to method 2, using two elastic modes per set ($m=2$). According to method 1, $\tilde{\sigma}(t)$ vanishes at the boundary, rises sharply with ''x'' reaching a peak at a magnitude exceeding unity exhibiting the Gibbs effect, and then converges with the $\tilde{\sigma}(t)$ computed by method 2. This behavior typifies a Fourier expansion of a delta function. The region near the excited boundary where the results of methods 1 and 2 differ narrows with the number of modes per set ''m.'' For example, when $m=5$, $\tilde{\sigma}(t)$ peaks at $s=0.2''$ and converges for $s>0.25''$. In column 2 ($m=40$) $\tilde{\sigma}(t)$ peaks at $s=0.02''$ and converges for $s>0.04''$. Clearly, method 2 requires far fewer modes per set than method 1 for similar convergence of $\tilde{\sigma}(t)$ near the excited boundary.

[1] M. El-Raheb, ''Transient elastic waves in finite layered media: One-dimensional analysis,'' J. Acoust. Soc. Am. **94**, 172–184 (1993).
[2] J. Berry and P. Naghdi, ''On the vibration of elastic bodies having time dependent boundary conditions,'' Q. Appl. Math. **14**, 43–50 (1956).

The modeling of a continuous structure as a fuzzy

Joel Garrelick
Cambridge Acoustical Associates, Inc., 200 Boston Avenue, Medford, Massachusetts 02155

(Received 14 August 1995; revised 19 April 1996; accepted 19 August 1996)

For a continuous structure with a deterministic drive, it is suggested by example that (1) the loss factor invariance of its fuzzy limit is realized only when boundaries/discontinuities away from the master structure measure many characteristic wavelengths so that the influence of their reflecting waves is negligible at the interface, and (2) the value of the fuzzy limit depends on structural details as well as averaged or ''smoothed'' characteristics. © *1997 Acoustical Society of America.* [S0001-4966(97)04712-7]

PACS numbers: 43.40.At, 43.40.Qi [PJR]

INTRODUCTION

In a number of recent papers it has been suggested that complex structural subsystems attached to a ''master'' structure may be analyzed as adjuncts and described in terms of a ''structural fuzzy.''[1–4] The simplest form of structural fuzzy is an array of point oscillators each affixed to the master through its spring. The relevant parameters then become the fuzzy mass density per cyclic frequency and modal density, both per characteristic dimension, or wavelength, of the master.

Viewed in this fashion it is found that considerable vibration damping may be realized over a reasonably broadband with a fuzzy of modest relative mass. Moreover, the asymptotic performance of a fuzzy with high modal density is invariant to the value of its characteristic damping parameter, e.g., loss factor. Also, in contrast to similar findings using the methods of statistical energy analysis (SEA), these results may be arrived at without invoking ensemble, spatial, or frequency averaging.[2,3] In other words, they appear to be deterministic. This note concerns the realization of this picture by generic substructures although the issues are described in the context of a simple rod and beam.

I. ANALYSIS

Consider the basic fuzzy structure representation of an array of J oscillators of equal mass and natural frequencies spaced from ω_1 to ω_2, with individual oscillators communicating only through a master structure. The array impedance and its fuzzy limit may be expressed as[2,3]

$$Z_{\text{fuzzy}} = -i\omega M\left(1-\omega^2\,\frac{m}{M}\sum_{j=1}^{J}\left[\omega^2-(\omega_j^*)^2\right]^{-1}\right) \tag{1a}$$

$$\Rightarrow \omega^2\mu\pi/2 - i\omega M(1+\chi), \tag{1b}$$

where m is the individual oscillator mass, ω is circular frequency, ω_j is the natural frequency of the jth oscillator, $\omega_j^* = \omega_j(1-i\eta)^{-1/2}$ with η structural loss factor, μ is the oscillator array mass density per unit circular frequency, $M = Jm$ is the total mass of the array, and χ an inertial factor that is of little present interest. [Dissipation in Eq. (1a) is expressed in terms of a structural damping model as it is in Ref. 3, rather than the Voigt model of Ref. 2.]

In reality, each oscillator may represent a separate substructure. However, for a generic substructure this is difficult to visualize for all but small values of J. The interface with the master must be effectively compact in terms of its characteristic wavelengths for the array to behave as a unit. This requirement can be quite stringent. For example, with a flexurally vibrating master structure of plating thickness h and speed c_p, the in-phase length of plating is $L_m = (\lambda_f/2) = \pi\sqrt{c_p h/3.46\omega}$. Thus, say at 250 Hz, and using typical material properties, this span measures only about 0.48 m in 0.025-m steel plating typical of ships, 0.93 m in 0.15-m concrete construction decking, or only 0.25 m in 6.35×10^{-3} m-thick aluminum fuselage plating.

Alternatively, the array represents J vibration modes of a single multimodal substructure. To explore this scenario further we consider a simple but continuous substructure with constant modal density, a rod free at one end and interfacing with the master at the other (Fig. 1). Here the drive impedance presented to the master is given by

$$Z_{\text{rod}} = Y_{\text{rod}}^{-1} = -i\omega M\left(1+2\omega^2\sum_{n=1}^{\infty}\left[\omega^2-(\omega_n^*)^2\right]^{-1}\right)^{-1} \tag{2a}$$

$$\Rightarrow \rho cA, \quad (\omega L/c)\,\eta/2 = \eta/2\Delta\omega \gg 1, \tag{2b}$$

where A is the cross-sectional area, $c = \sqrt{E/\rho}$ is the material compressional speed, $c^* = c(1-i\eta)^{1/2}$, E is the material compressional modulus, L is the bar length, $\Delta\omega = \pi c/L$ (inverse modal density), ρ is the material mass density, $\omega_n = n\pi c/L$, $n = 0,1,2...$ (rod natural frequencies), and $\omega_n^* = \omega_n(1-i\eta)^{1/2}$. Equation (2b) may be also be expressed as

$$Z_{\text{rod}} \Rightarrow \overline{\omega^2}(M/\Delta\omega)(\pi/2), \quad (\omega L/c)\,\eta/2 = \eta/2\Delta\omega \gg 1, \tag{2c}$$

where $\overline{\omega^2} = K/M$ with $K = 2\rho c^2 A/L$, the static stiffness of the bar when subjected to equal compressional forces at each end. For convenience, in Eqs. (1) and (2) we have assumed the same (constant) loss factor. Also, the rigid body component of the impedance has been isolated in both cases. We now compare Eqs. (1) and (2).

With the fuzzy oscillators being mechanically in parallel, their impedances add. This is in contrast to the modes of the rod which are mechanically in series and therefore their modal admittances add. This yields the overall reciprocal

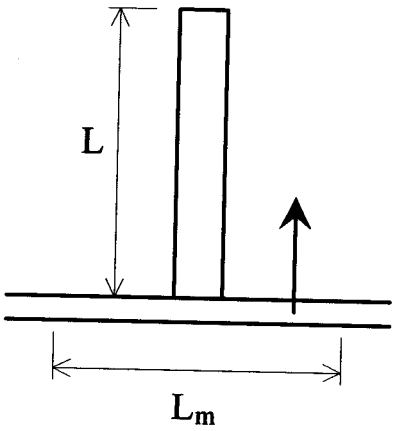


FIG. 1. Substructure (free rod) attached to master structure vibrating in-phase over length L_m.

form of Eqs. (1a) and (2a). However, our primary interest is in the high modal density or fuzzy limits given by Eqs. (1b) and (2b) or (2c), specifically the fact that the resistances of both the bar and the array are independent of loss factor.

This has traditionally been explained in terms of high modal density and modal overlap. However, deterministic solutions exhibiting high modal density generally imply wave motion, and for our purposes it is insightful to consider the equivalent traveling wave series solution for the rod impedance,

$$Z_{\text{rod}} = \rho c^* A \sum_{q=-\infty}^{\infty} \exp\left(\frac{i2q\omega L}{c^*}\right). \tag{3}$$

Equation (3) converges to the modal solution given by Eq. (1a), provided only that $\omega, \eta \neq 0$.

The high-frequency asymptotic limit, $q=0$, represents the direct field. It is the solution for the equivalent semi-infinite bar. As such it contains no contributions from waves that have propagated to the free end and been reflected back. Therefore, it is invariant to the attenuation of such waves as they propagate and consequently to the loss factor. Extrapolating, the (harmonic) resistance of a generic substructure may be deterministically modeled as a fuzzy only if it is effectively semi-infinite; that is, only if waves reflected from discontinuities anywhere along the substructure are of little consequence at the master interface. This may be expressed generally in the form $\alpha(\eta l/\lambda_c) \gg 1$ with λ_c the (longest) characteristic wavelength, l the distance to the boundary/discontinuity, and α a numerical constant depending on wave type. (This view is consistent with the semi-infinite and infinite canonical models that are commonly used with SEA for deriving coupling loss factors among subsystems.[5])

The above equivalence notwithstanding, the difference in the frequency dependence of the resistances given by Eq. (1b) and Eq. (2b) or (2c) cannot be reconciled. However, they may be equated by postulating a fuzzy oscillator array with the frequency-dependent mass density $\mu(\omega) = (2/\pi)\rho c A/\omega^2$. The question now becomes, how much

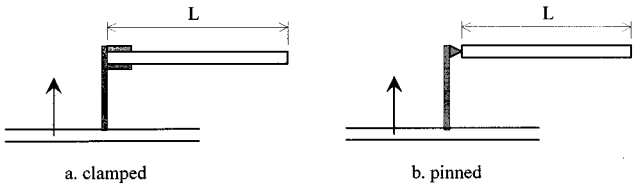


FIG. 2. Substructure (beam) attached to vibrating master structure with (a) clamped and (b) pinned connections.

substructure detail is required for its simulation as a fuzzy? The answer, unfortunately, appears to be that a rather precise and complete description may be required.

Again illustrating the point by example, we now consider a flexurally vibrating Euler beam as our substructure rather than a rod. In one case [Fig. 2(a)] the beam is constrained against rotation, or clamped, at the master structure interface and in the other [Fig. 2(b)] it is ''pin'' connected, that is, free to rotate. The equivalent high-frequency/modal density translational impedance expressions are[6]

$$Z_{\text{beam}} = 2\pi\omega^2[M/(\Delta\omega)_f](1-i) \quad \text{clamped,} \tag{4a}$$

$$Z_{\text{beam}} = \pi\omega^2[M/(\Delta\omega)_f](1-i) \quad \text{pinned,} \tag{4b}$$

where k_f is the flexural wavelength at frequency ω, $(\Delta\omega)_f = 2\pi\omega/k_f L$ is the inverse flexural modal density, and M and L again denote substructure mass and length. Here a structural near field produces an inertial component that is equal in magnitude to the resistance. More to the point, here the value of the high-frequency limit of the resistance changes with the detail of the interface boundary.

II. CONCLUSION

In summary, it is suggested by example that for a continuous structure with a deterministic drive (1) the loss factor invariance of its fuzzy limit will be achieved (only) when boundaries/discontinuities away from the master interface measure many characteristic wavelengths (λ_c) so that, stipulating only a finite loss factor η, the influence of their reflecting waves is negligible at the interface; and (2) the value of the fuzzy limit itself depends on structural details as well as averaged or otherwise ''smoothed'' characteristics. The first item may be expressed in the form $\alpha(\eta l/\lambda_c) \gg 1$ with 1 the distance to the nearest boundary/discontinuity and α a numerical constant depending on wave type. The second item appears to require detailed rather than ''broad brush'' calculations and/or measurements to implement a fuzzy analysis.

ACKNOWLEDGMENTS

The author would like to thank Dr. Murray Strasberg and Dr. David Feit and the reviewers for sharing their insights into this problem.

[1] C. Soize, ''A model and numerical method in the medium frequency rang for vibroacoustic predictions using the theory of structural fuzzy,'' J. Acoust. Soc. Am. 24 (2), 849–865 (1993).
[2] D. Feit and M. Strasberg, ''The structural fuzzy as a source of vibration

damping,'' in *Applied Mechanics in the Americas*, edited by L. A. Godoy *et al.* (AAM and AMCA, Santa Fe, Argentina, 1995), Vol. II, pp. 60–65.
[3] M. Strasberg and D. Feit, ''Vibration damping of large structures induced by attached small resonant structures,'' J. Acoust. Soc. Am. **99**, 335–344 (1996).
[4] A. D. Pierce, V. W. Sparrow, and D. A. Russell, ''Fundamental structural-acoustic idealizations for structures with fuzzy internals,'' J. Vib. Acoust. **117**, 339–348 (1995).
[5] R. H. Lyon, *Statistical Energy Analysis of Dynamical Systems. Theory and Applications* (MIT, Cambridge, MA, 1975).
[6] L. Cremer and M. Heckl, *Structure-Borne Sound* (Springer-Verlag, Berlin, 1973), translated by E. Ungar, p. 281.

On the standard deviation of change-in-impedance due to fuzzy subsystems

Y. K. Lin
Center for Applied Stochastics Research, Florida Atlantic University, Boca Raton, Florida 33431-0991

(Received 14 March 1996; accepted for publication 9 August 1996)

The effects of attached multiple subsystems of uncertain properties on the behavior of a main system are investigated by modeling the natural frequencies of the subsystems as random points in the frequency domain. Both the mean and the standard deviation of the change-in-impedance, attributable to the subsystems, are computed without the usual assumption of ''modal overlap.'' It is shown that the standard deviation, which has been overlooked by most authors on this subject, can sometimes be greater than the mean. Thus the knowledge of the mean alone is inadequate to make engineering decisions. © *1997 Acoustical Society of America.* [S0001-4966(97)03212-8]

PACS numbers: 43.40.Dx [CBB]

INTRODUCTION

Recently, there has been considerable interest in the effects of multiple secondary systems (or appendages) on the motion of a master structure. It is assumed that the secondary systems are numerous, each of which is much lighter than the master structure, and that their physical properties are not precisely known; thus they are referred to collectively as fuzzy subsystems, or internal fuzzies.

Soize[1,2] and his associates called attention to the problem, and provided first analyses of the effects. Further advances were made by Strasberg and Feit,[3] Pierce,[4] Pierce *et al.*,[5] and others, mostly focusing on the average energy dissipation, attributable to randomly distributed subsystems. It was shown[5] that this average dissipation could be obtained by replacing the sum of contributions from all the fuzzies by an integral, which could then be evaluated with an asymptotic method. The results so obtained were found to be independent of the damping in the fuzzy subsystems so long as it was small. The nondissipative (reactive) components were not equally emphasized, perhaps, for either one of the two reasons: Both positive and negative components were present, and as a whole they did not consume energy in the long run.

Surprisingly, not enough attention was paid to the standard deviation from the mean. Indeed, if standard deviation is of a similar magnitude as the mean or even larger, then the mean alone is not very useful for prediction or design purposes. Furthermore, the replacement of a sum by an integral is equivalent to converting the number of subsystems to uncountable infinity. Thus the most important effect of resonance when the excitation frequency coincides with the natural frequency of any single subsystem is accorded with probability zero. The above two problems will now be examined in a simplified analysis.

I. ANALYSIS

Let the main system be represented by a single-degree-of-freedom linear oscillator with mass M, natural frequency ω_0, and the ratio of damping to the critical damping ζ_0, and subjected to a sinusoidal excitation of frequency ω at the mass. Let the fuzzies be represented by N small oscillators, each of which is attached separately to mass M. Then the impedance of the overall system may be expressed as

$$Z(\omega)=M\{\omega_0^2-\omega^2+R(\omega)+i[2\zeta_0\omega_0\omega+I(\omega)]\}, \tag{1}$$

where

$$R(\omega)=\sum_{j=1}^{N}\nu_j\gamma(\omega,\Omega_j,\zeta_j), \tag{2}$$

$$I(\omega)=\sum_{j=1}^{N}\nu_j\chi(\omega,\Omega_j,\zeta_j), \tag{3}$$

and where ν_j is the mass of the jth fuzzy, expressed as a fraction of M, and Ω_j and ζ_j are the natural frequency and the damping ratio of the jth fuzzy, respectively. The impedance here refers to the *displacement response* (instead of velocity response) to force excitations; therefore, the imaginary part $\mathrm{I}(\omega)$ corresponds to damping and the real part $\mathrm{R}(\omega)$ corresponds to stiffness, contributed by the fuzzies. One can show that

$$\gamma(\omega,\Omega,\zeta)=-\frac{\omega^2[\Omega^2(\Omega^2-\omega^2)+4\zeta^2\Omega^2\omega^2]}{(\Omega^2-\omega^2)^2+4\zeta^2\Omega^2\omega^2}, \tag{4}$$

$$\chi(\omega,\Omega,\zeta)=\frac{2\zeta\Omega\omega^5}{(\Omega^2-\omega^2)^2+4\zeta^2\Omega^2\omega^2}. \tag{5}$$

The following simplified assumptions are made: (1) the ν_j are independent and identically distributed random variables with known mean and mean-square values, (2) the Ω_j are independent random points on an interval $[\Omega_l,\Omega_u]$ with a nonuniform average density (average number of random points per unit bandwidth) $\lambda(\Omega)$, and (3) $\zeta_j=\zeta$ for all j. Use is made of the theory of random points (see, for example, Stratonovich[6] or Lin[7]), a sketch of which, as applied to the present problem, is given below.

A stochastic process, such as $R(\omega)$ given in (2), may be described in terms of its characteristic functional

$$M_{\{R\}}[\theta(\omega)]=E\left\{\exp\left[i\int_{\Omega_l}^{\Omega_u}R(\omega)\theta(\omega)d\omega\right]\right\}, \tag{6}$$

where $E\{\ \}$ denotes the ensemble average, and $\theta(\omega)$ belongs to the family of functions for which the integral shown on the right-hand side exists in a certain sense. Substituting (2) into (6),

$$\begin{aligned}M_{\{R\}}[\theta(\omega)]&=E\left\{\exp\left[i\int_{\Omega_l}^{\Omega_u}\sum_{j=1}^{N}\nu_j\gamma(\omega,\Omega_j)\theta(\omega)d\omega\right]\right\}\\&=E\left\{\prod_{j=1}^{N}\exp\left[i\nu_j\int_{\Omega_l}^{\Omega_u}\gamma(\omega,\Omega_j)\theta(\omega)d\omega\right]\right\}\\&=E\left\{\prod_{j=1}^{N}\left[1+\sum_{k=1}^{\infty}\frac{1}{k!}i^k\nu_j^k\int_{\Omega_l}^{\Omega_u}\cdots\int_{\Omega_l}^{\Omega_u}\gamma(\omega_1,\Omega_j)\right.\right.\\&\quad\cdots\gamma(\omega_k,\Omega_j)\theta(\omega_1)\cdots\\&\quad\left.\left.\theta(\omega_k)d\omega_1\cdots d\omega_k\right]\right\}.\end{aligned} \tag{7}$$

The dependence of $\gamma(\omega,\Omega_j,\zeta_j)$ on ζ_j is omitted here since it has been assumed that $\zeta_j=\zeta$ for all j. This assumption can be removed if so desired.

On the other hand, a system of random points Ω_j can be described by its generating functional[6]

$$L[u(\Omega)]=E\left\{\prod_{j=1}^{N}[1+u(\Omega_j)]\right\}. \tag{8}$$

The two functionals (7) and (8), become the same, if we let

$$\begin{aligned}u(\Omega)=\sum_{k=1}^{\infty}\frac{1}{k!}i^k\nu_j^k\int_{\Omega_l}^{\Omega_u}\cdots\int_{\Omega_l}^{\Omega_u}\gamma(\omega_1,\Omega)\cdots\\\gamma(\omega_k,\Omega)\theta(\omega_1)\cdots\theta(\omega_k)d\omega_1\cdots d\omega_k.\end{aligned} \tag{9}$$

Now the natural logarithm of (6), called log-characteristic functional of stochastic process $R(\omega)$, admits a series expansion[6]

$$\begin{aligned}\ln M_{\{R\}}[\theta(\omega)]=\sum_{r=1}^{\infty}\frac{1}{r!}i^r\int_{\Omega_l}^{\Omega_u}\cdots\int_{\Omega_l}^{\Omega_u}\kappa_r[R(\omega_1),\ldots,\\R(\omega_r)]\theta(\omega_1)\cdots\theta(\omega_r)d\omega_1\cdots d\omega_r,\end{aligned} \tag{10}$$

where $\kappa_r[Y_1,\ldots,Y_r]$ denotes the rth joint cumulant (or semi-invariant) of the random variables $Y_1,\ldots,Y_r$. Likewise, the natural logarithm of (8) has an expansion

$$\begin{aligned}\ln L[u(\Omega)]=\sum_{m=1}^{\infty}\frac{1}{m!}\int_{\Omega_l}^{\Omega_u}\cdots\int_{\Omega_l}^{\Omega_u}\kappa_m[\Lambda(\Omega_1),\ldots,\\\Lambda(\Omega_m)]u(\Omega_1)\cdots u(\Omega_m)d\Omega_1\cdots d\Omega_m,\end{aligned} \tag{11}$$

where each $\Lambda(\Omega_s)$ is a random variable, representing the ''random density'' of random points at Ω_s.

Substituting (9) into (11), and comparing the result with (10), we obtain a set of expressions for computing the cumulants of $R(\omega)$. The complexity of such expressions depends, of course, on the cumulants of the random density $\Lambda(\Omega)$. If the random points Ω_j are independent for different j, as what has been assumed earlier, then only the first cumulant of $\Lambda(\Omega)$ is nonzero, and we obtain the simplest expressions:

$$\begin{aligned}&\kappa_r[R(\omega_1),\ldots,R(\omega_r)]\\&\quad=E[\nu^r]\int_{\Omega_l}^{\Omega_u}\gamma(\omega_1,\Omega)\cdots\gamma(\omega_r,\Omega)\kappa_1[\Lambda(\Omega)]d\Omega,\end{aligned} \tag{12}$$

where $E[\nu^r]$ can be any $E[\nu_j^r]$ since all ν_j are identically distributed. In particular, the first cumulant of $R(\omega)$ is the mean, and the joint cumulant of $R(\omega_1)$ and $R(\omega_2)$ is the covariant. These are given specifically below:

$$E[R(\omega)]=E[\nu]\int_{\Omega_l}^{\Omega_u}\gamma(\omega,\Omega)\lambda(\Omega)d\Omega, \tag{13}$$

$$\begin{aligned}&\mathrm{cov}[R(\omega_1),R(\omega_2)]\\&\quad=E[\nu^2]\int_{\Omega_l}^{\Omega_u}\gamma(\omega_1,\Omega)\gamma(\omega_2,\Omega)\lambda(\Omega)d\Omega,\end{aligned} \tag{14}$$

where we have also made the substitution $\kappa_1[\Lambda(\Omega)]=E[\Lambda(\Omega)]=\lambda(\Omega)$. The covariance reduces to the variance (the standard deviation squared) when $\omega_1=\omega_2=\omega$.

The analysis for the imaginary part $I(\omega)$ of the impedance is similar.

Numerical results have been obtained for two cases: (1) $N=100$, the average density of the random points Ω_j in the interval $[\Omega_l,\Omega_u]$ is $\sin^2[\pi(\Omega-\Omega_l)/(\Omega_u-\Omega_l)]$, and the random variables ν_j have a common mean $E[\nu]$ and a common mean-square value $E[\nu^2]$ of 3×10^{-3} and 1.146×10^{-5}, respectively; (2) $N=10$, the average density of random points Ω_j is $0.1\sin^2[\pi(\Omega-\Omega_l)/(\Omega_u-\Omega_l)]$, and the mean and mean-square values of ν_j are 3×10^{-2} and 1.146×10^{-3}, respectively. In both cases, we let $\Omega_l=50$ rad/s, $\Omega_u=250$ rad/s, and $\zeta=0.01$. Note that the average total mass of fuzzies is the same 0.3 M for both cases, but the number of fuzzies is ten times larger in the first case than that in the second case.

The computed mean and standard deviation of the real (reactive) part $R(\omega)$ are shown in Fig. 1. The average $R(\omega)$, labeled μ, is the same for both cases, but the standard devia-

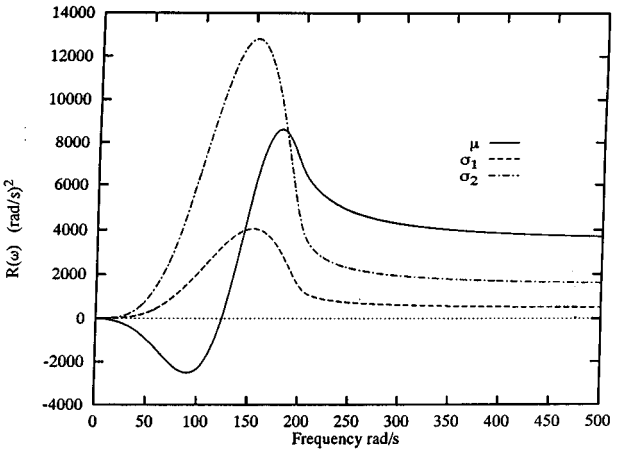


FIG. 1. Reactance part of fuzzies.

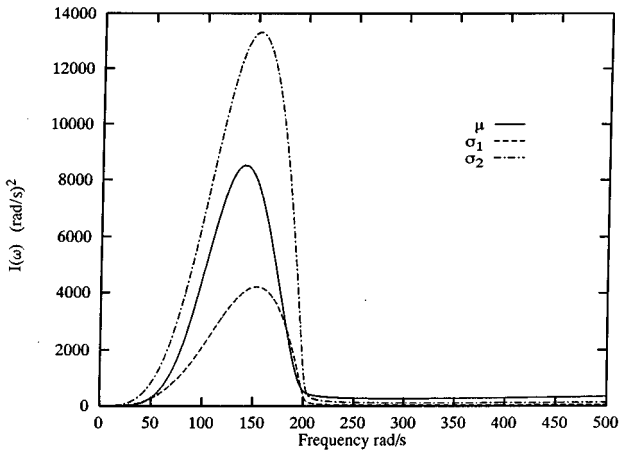


FIG. 2. Dissipation part of fuzzies.

tions of $R(\omega)$, labeled σ_1 and σ_2 for cases (1) and (2), respectively, are very different. The same remarks can be made of the results for the imaginary (dissipative) part $I(\omega)$, shown in Fig. 2.

The above results show that the standard deviations for $R(\omega)$ and $I(\omega)$ can indeed be greater than the mean, more so when fewer fuzzies are present in a given frequency band $[\Omega_l, \Omega_u]$. We also found that the magnitude of damping is important, and that it has a greater effect on the standard deviation than the mean. These can only be revealed if the fuzzies are treated realistically as discrete elements, not a *spreaded continuum* in the frequency domain (or equivalently in the mass domain).

We reiterate that the computed means for $R(\omega)$ and $I(\omega)$ are the same regardless of the total number of fuzzies assumed in the analysis ($N=100$ or $N=10$, as long as the total mass is the same). Therefore, the question of modal overlap is immaterial when computing the mean, a matter of main concern of several authors on the subject. This is, in fact, quite transparent from Eq. (1); a smaller number of fuzzies implies a larger average mass $E[\nu]$ of an individual fuzzy, but the average density of the random points is reduced proportionally. On the other hand, under the present independent random point assumption, the increase of the variance of $R(\omega)$ or $I(\omega)$ due to an increasing $E[\nu^2]$ is only partially compensated by a decreasing $\lambda(\Omega)$.

In closing, we note that if the assumption of independent random points is removed, then the variance of $R(\omega)$ or $I(\omega)$ also depends on $E[\nu]$ and on the covariance of $\Lambda(\Omega)$ at two Ω locations.[8]

[1] C. Soize, ''Probabilistic Structural Modeling in Linear Dynamic Analysis of Complex Mechanical Systems: I. Theoretical Elements,'' La Recherche Aerospatiale **5**, 23–48 (1986).
[2] C. Soize, ''A Model and Numerical Method in the Medium Frequency Range for Vibroacoustic Predictions Using the Theory of Structural Fuzzy,'' J. Acoust. Soc. Am. **94**, 849–865 (1993).
[3] M. Strasberg and D. Feit, ''Vibration Damping of Large Structures Induced by Attached Small Resonant Structures,'' J. Acoust. Soc. Am. **99**, 335–344 (1996).
[4] A. D. Pierce, ''Mass Per Unit Natural Frequency as a Descriptor of Internal Fuzzy Structure,'' J. Acoust. Soc. Am. **95**, 2845(A) (1994).
[5] A. D. Pierce, V. W. Sparrow, and D. A. Russell, ''Fundamental Structural-Acoustic Idealization for Structures with Fuzzy Internals,'' ASME J. Vib. Acoust. **117**, 1–10 (1995).
[6] R. L. Stratonovich, *Topics in the Theory of Random Noise* (Gordon and Breach, New York, 1963), Vol. 1.
[7] Y. K. Lin, *Probabilistic Theory of Structural Dynamics* (McGraw-Hill, New York, 1967, reprint by Krieger, Melbourne, FL, 1976).
[8] Y. K. Lin, ''On Random Pulse Train and Its Evolutionary Spectral Representation,'' Probabilistic Eng. Mech. **1**(4), 219–223 (1986).

Stop bands for cubic arrays of spherical balloons

M. S. Kushwaha
Instituto de Física, Universidad Autónoma de Puebla, Apartado Postal J-48, Puebla 72570, Pue., México

P. Halevi
Instituto Nacional de Astrofísica, Optica y Electrónica, Apartado Postal 51, Puebla 72000, Pue., México

(Received 21 December 1995; revised 24 June 1996; accepted 4 September 1996)

The acoustic band structures have been computed for cubic arrays of spherical hydrogen balloons in air. Stop bands are obtained for the face-centered-cubic (fcc) and body-centered-cubic (bcc) lattices, however, there is no band gap for the simple cubic lattice. These gaps are largest for a volume fraction of ~38% and, at low balloon pressure (1.10 atm) and small latex wall thickness (10^{-3} times the inner radius), the corresponding gap/midgap ratios are about 0.2 for the fcc and 0.1 for the bcc structure. As the pressure or wall thickness increases, the stop bands diminish and ultimately vanish.
 [S0001-4966(97)05412-X]

PACS numbers: 43.40.Tm, 43.20.Fn, 43.20.Hq, 43.50.Gf [CBB]

Consider two elastic materials, a and b, in each of which vibrations and sound can propagate freely. Now imagine mixing up a and b in some clever way; is it possible to impede the wave propagation? If the materials are disordered then the answer is given by Anderson localization,[1] which, indeed, has been observed for bending waves in a steel plate decorated with Lucite blocks.[2] Here, on the contrary, we are concerned with perfect order, a periodic arrangement of a and b.

In order to answer the question posed above one must recourse to band structure (BS) calculations, namely the normal mode solutions $\omega(\mathbf{k})$ for a 3-D periodic system.[3] These were performed for several geometries of periodic elastic composites and for various types of waves. In the case of one-dimensional periodicity (a superlattice) longitudinal, transverse, and mixed-polarization waves can propagate.[4] More recently longitudinal,[5] transverse,[6–10] and mixed (longitudinal-transverse)[6,9,10] modes were studied in systems of two-dimensional periodicity (parallel, infinite cylinders). While these works are all theoretical, experiments in 2-D systems are also in progress.[11] For all these configurations, band gaps were indeed found irrespective of the direction of the wave vector, thus leading to stop bands for certain values of the material parameters and filling fractions of the inclusions.

In solid composites with 3-D periodicity the longitudinal and transverse vibrations are always coupled,[9,12] thus complicating greatly the nature of the eigenmodes and the corresponding computation. Considerable simplification arises in the case of liquids and/or gases, which admit only longitudinal (''acoustic'') waves.[5] In Ref. 5, Sigalas and Economou concluded that the optimum situation for the appearance of gaps is low-density, low-velocity spheres occupying a volume fraction of the order of 10% of the host material.[13] However, no specific material media were considered.

Here we report BS and modal density calculations for a patently low-tech system: a periodic, 3-D array of spherical balloons containing gas (hydrogen) and immersed in air. Due to its extreme simplicity, this system is attractive for experimentation and convenient for applications such as the creation of a vibrationless environment in a desired band of frequencies (acoustic filters). In the frequency range of a ''complete'' band gap—independent of the polarization of the wave, as well as its direction of propagation—vibration, sound, and phonons are forbidden. Thus a small vibrator, introduced into the periodic composite as a defect, would be unable to vibrate freely at a frequency within the gap range. Similarly, a real crystallite in the role of a defect would be unable to generate low-frequency acoustic phonons. This would lead to the interesting possibility of localized vibrations or localized phonons, much like an atom inside a *dielectric* composite, whose radiation within a photonic band gap is frustrated, becomes surrounded by an evanescent electromagnetic field.[14] It would be very difficult—if not impossible—to investigate such localized states in a solid composite; *air* as the host medium is ideally suited for such experimentation. Moreover, the positions (or other characteristics) of the gas balloons can be easily randomized in air, thus providing a convenient system for the study of Anderson localization of sound waves.[1,2]

One random system of importance for acoustic wave propagation is bubbly liquids. Interestingly, some studies of this topic employ the model of an equivalent *periodic* distribution of bubbles.[15] However, these calculations are based on the assumption that the wavelength is much greater than the lattice period (that is, the average interparticle separation), thus eliminating the standing wave (Bragg) resonances between adjacent bubbles. On the other hand, the main thrust of this Letter is precisely the band gaps resulting from the phenomenon of Bragg diffraction and Mie resonance.

Very recently we have presented a detailed theory of elastic band structure.[9] The calculation at hand simplifies because the transverse speed of sound c_t is zero in gases (and in liquids), and we neglect the shear rigidity of the latex wall ($c_t \ll c_l$). Nevertheless the classical wave equation is inapplicable to inhomogeneous media. The correct wave equation—simply the equation of motion in the absence of external forces—is

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = \nabla(\rho c_l^2 \nabla \cdot \mathbf{u}), \tag{1}$$

where $\rho(\mathbf{r})$ is the mass density and $c_l(\mathbf{r})$ is the longitudinal speed of sound. Only if ρc_l^2 is independent of the position $\mathbf{r}$, then all three components of $\mathbf{u}(\mathbf{r},t)$ satisfy the classical wave equation. In the general case, from Eq. (1) we observe that $\nabla \times (\rho \mathbf{u}) = 0$. Hence it is possible to define a scalar potential $\Phi(\mathbf{r},t)$, such that $\rho \mathbf{u} = \nabla \Phi$. Then Eq. (1) may be rewritten as a scalar equation,[16]

$$(\rho c_l^2)^{-1} \frac{\partial^2 \Phi}{\partial t^2} = \nabla \cdot (\rho^{-1} \nabla \Phi). \tag{2}$$

Taking advantage of the periodicity of the medium we expand the quantities $\rho^{-1}(\mathbf{r})$ and $[\rho(\mathbf{r}) c_l^2(\mathbf{r})]^{-1}$ in Fourier series:

$$\rho^{-1} = \sum_{\mathbf{G}} \sigma(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \quad (\rho c_l^2)^{-1} = \sum_{\mathbf{G}} \zeta(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{3}$$

where $\mathbf{G}$ is the reciprocal lattice vector. The solution of Eq. (2) is given by means of the Bloch theorem,

$$\Phi(\mathbf{r},t) = e^{i(\mathbf{K}\cdot\mathbf{r} - \omega t)} \sum_{\mathbf{G}} \Phi_{\mathbf{K}}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{4}$$

where $\mathbf{K}$ is the Bloch wave vector. Substitution of Eqs. (3) and (4) in Eq. (2) yields an infinite-set of equations for the eigenvalues $\omega(\mathbf{K})$ and the eigenvectors $\Phi_{\mathbf{K}}(\mathbf{G})$:

$$\sum_{\mathbf{G}'} [\omega^2 \zeta(\mathbf{G}-\mathbf{G}') - (\mathbf{K}+\mathbf{G}) \cdot (\mathbf{K}+\mathbf{G}') \sigma(\mathbf{G}-\mathbf{G}')] \times \Phi_{\mathbf{K}}(\mathbf{G}') = 0. \tag{5}$$

We apply this equation to a periodic system of spherical, gas-filled balloons in a background gaseous medium. The gas inside the balloons, the background gas, and the latex walls are labeled by a, b, and w, respectively. The corresponding densities are ρ_a, ρ_b, and ρ_w; the (longitudinal) speeds of sound are c_a, c_b, and c_w; and the material filling fractions are f_a, f_b, and f_w; where $f_a + f_b + f_w = 1$. Then inverting the Fourier series (3)

$$\sigma(\mathbf{G}) = \frac{1}{V_c} \int_{V_c} d\mathbf{r}\ \rho^{-1}(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} = \begin{cases} \rho_a^{-1} f_a + \rho_b^{-1}(1 - f_a - f_w) + \rho_w^{-1} f_w \equiv \overline{\rho^{-1}}, & \mathbf{G} = 0, \\ (\rho_a^{-1} - \rho_w^{-1}) F_{r_<}(\mathbf{G}) + (\rho_w^{-1} - \rho_b^{-1}) F_{r_>}(\mathbf{G}), & \mathbf{G} \neq 0, \end{cases} \tag{6}$$

where $r_<$ and $r_>$ are the inner and outer radii of the balloons and

$$F_R(G) = \frac{1}{V_c} \int_{r \leqslant R} d^3 r\ e^{-i\mathbf{G}\cdot\mathbf{r}} = \frac{3 f_a}{(GR)^3} [\sin(GR) - GR\ \cos(GR)]. \tag{7}$$

Here, V_c is the volume of the unit cell, the integration is limited to a sphere of radius R, and $f_a = n(4\pi/3) r_<^3 / V_c$, where n is the number of spheres in the unit cell. An equation analogous to Eq. (6) can be written for $\zeta(\mathbf{G})$ in terms of the average of $(\rho c_l^2)^{-1} \equiv \overline{C^{-1}}$ and the corresponding contrast parameters. With these substitutions, Eq. (5) is an infinite set of homogeneous, linear equations for the Fourier components $\Phi_{\mathbf{K}}(\mathbf{G})$. The nontrivial solution is obtained from the requirement that the determinant of the square bracket in Eq. (5) vanishes. This yields the eigenvalues $\omega(\mathbf{K})$ which, for a periodic system, is the band structure. We have computed the lowest 50 bands for simple-cubic (sc), body-centered-cubic (bcc), and face-centered-cubic (fcc) arrays of hydrogen balloons in air. We obtain real eigenvalues and good convergence (of better than 2%) by limiting the number of plane waves in Eq. (5) to 343. By increasing the number of plane waves to 729, our results change by no more than 2%.

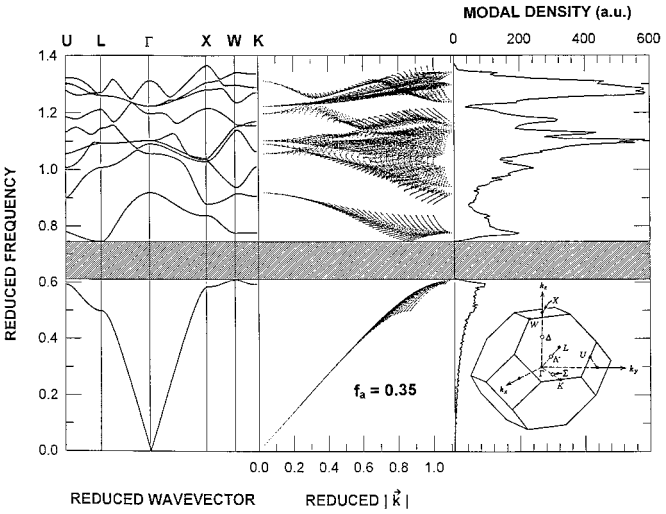


FIG. 1. Acoustic band structure for a fcc array of spherical balloons, containing hydrogen gas, in air (left part of triptych). Middle part: frequency eigenvalues as a function of the magnitude of the Bloch wave vector, scanned throughout the irreducible part of the Brillouin zone (shown in the inset on the right). The modal density (in arbitrary units), as a function of the reduced frequency $\omega a/2\pi\overline{c_l}$ (where a is the lattice constant and $\overline{c_l} = \overline{\rho^{-1}/C^{-1}}$) is graphed in the right part of the figure. The shaded band is forbidden for vibrations or sound. The parameter[17] contrasts are $\rho(H_2)/\rho(\text{air}) = 0.076$ and $c_l(H_2)/c_l(\text{air}) = 3.706$; the parameters for the latex material (rubber) are: $\rho = 1.07$ g/cc and $c_l = 1830$ m/s. The pressure inside the balloons in 1.10 atm and it is atmospheric outside. The thickness of the latex wall is 10^{-3} times the inner radius and the filling fraction is $f_a = 0.35$.

The left part of Fig. 1 is the band structure for a fcc lattice of balloons that occupy 35% of the total volume. The Brillouin zone of this lattice with the principal symmetry points in reciprocal space is shown in the inset on the right. Note that within any given segment (such as U-L and Γ-X), the vector $\mathbf{K}$ increases monotonously while keeping the same direction. There appears a complete acoustic gap between the first and second bands, and there are no other gaps at least as far as the 50th band. The middle part of Fig. 1 is the result of a detailed scan of $|\mathbf{K}|$ in the irreducible part of the first Brillouin zone—the interior of this zone and its surface, as well as the principal directions shown in the left part of the figure. Each curve here corresponds to some direction of $\mathbf{K}$. The modal density in the right part of the figure has been calculated on the basis of the scan in the middle part. It is obtained simply by computing the total number of $\mathbf{K}$ points in a small frequency interval $\delta\omega$. These parts of the figure prove that there is, indeed, a genuine, full gap between the first two

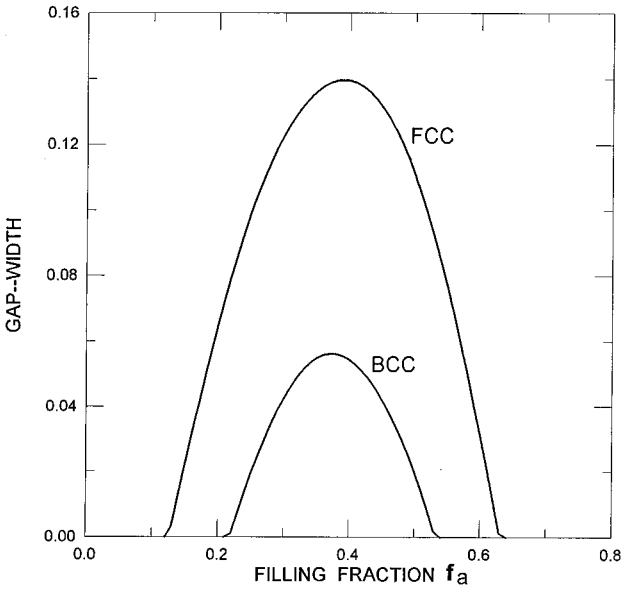


FIG. 2. Normalized gap width versus filling fraction f for the fcc and bcc structures (there is no gap for the sc structure). For both structures the largest gap is obtained for $f \simeq 0.38$, and, for a pressure $p = 1.10$ atm, the approximate gap/midgap ratios are 0.2 and 0.1 for the fcc and bcc lattices, respectively.

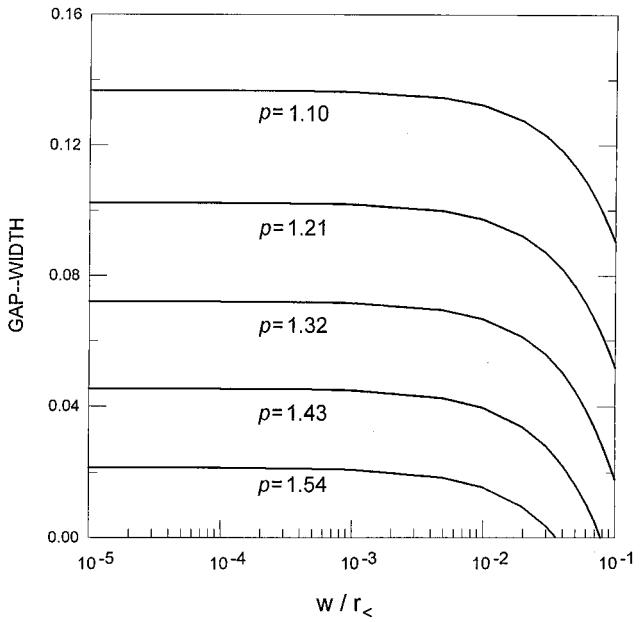


FIG. 3. Gap width versus latex wall thickness (w), normalized to its inner radius $r_<$ (fcc structure), with pressure p (in atm) inside the balloons as a parameter. The largest gaps are obtained for very thin walls and pressures just above atmospheric.

bands, and we consider such calculations as essential. Similar results are obtained for the bcc structure, however we did not find a stop band for the sc array—for any value of the filling fraction.

The dependence of the (lowest) gap width on the filling fraction, for the fcc and bcc structures, is summarized in Fig. 2. The filling fraction must exceed a certain minimum value, $f_{\min}$, for a gap to be obtained. If the pressure in the balloons is 1.1 atm, then for the fcc and bcc structures $f_{\min} \simeq 0.12$ and 0.21, respectively. The corresponding maximum values are $f_{\max} = 0.64$ and 0.54. For any given value of f the fcc structure gives a larger gap width than the bcc structure. For both lattices, the stop bands are widest when $f \simeq 0.38$, that is when the balloons occupy 38% of the space. However, the gap/midgap ratio is considerably larger for the fcc array: about 0.2 vs 0.1 for the bcc array.

In these calculations we assumed that the hydrogen and the air both satisfy the ideal gas equation of state; hence $\rho = pM/k_BT$. If the pressure p inside the balloons is only a little above atmospheric ($p = 1.10$ atm), then the density is essentially proportional to the molecular weight M. This being, on the average, about 29 for air, the density inside is roughly 14.5 times smaller than outside. The speed of sound is $c_l = (\gamma RT/M)^{1/2}$; the adiabatic constant γ is about the same for both gases, hence c_l is approximately inversely proportional to the square root of the molecular mass. Then the speed of sound inside is about $\sqrt{29/2} \simeq 3.8$ times greater than outside. Thus we see that *high*-velocity, low-density spheres in a low-velocity, high-density host can lead to substantial stop bands for (longitudinal) acoustic waves. This finding supplements the conclusion[5] that the bubble configuration (*low*-velocity, low-density spheres) is optimal for the appearance of gaps.

In Fig. 3 we examine the dependence of the gap width, for the fcc structure, on the pressure p inside the balloons and on the latex wall thickness ($w = r_> - r_<$) normalized to $r_<$. For $w/r_< \lesssim 10^{-3}$ the band gaps are quite independent of the wall thickness; however their widths decrease rapidly as the pressure is increased. The reason for this is, of course, that the density contrast between the two gases diminishes with growing p. When the latex thickness is less than 1% of the balloon (inner) radius the gap width quickly approaches zero. Thus we may conclude that a substantial stop band can be achieved only if the latex is quite thin and the pressure inside the balloons is not *much* above atmospheric. Clearly, for the gases considered here, the gap for the uppermost curve ($p = 1.10$ atm) in the limit of vanishing latex thickness is the largest one; it corresponds to the gap/midgap ratio of $\sim 20\%$.

Within the gaps of Figs. 1–3, the perfect and infinitely extended ''acoustic crystals'' investigated here stand still and total silence reigns. The situation is of comparable interest to the full photonic gaps obtained in periodic dielectric composites.[18] These, however, were realized only with certain complex structures of the fcc unit cell[18] not with the simple FCC lattice (one spherical ''atom'' per site), and much less for the bcc lattice. The periodic arrays of balloons discussed in this letter are probably the simplest physical systems that exhibit complete stop bands. For gas-filled balloons in air the conditions that lead to large band gaps are a gas with small molecular weight and pressure just slightly above atmospheric, very thin latex walls, and the fcc arrangement.

Note added in proof: After this paper had been submitted two papers (Refs. 19 and 20) were published on the existence of giant acoustic stop bands in two- (and three-) dimensional inhomogeneous liquid systems. In addition, we refer the readers to a long review on the subject (Ref. 21).

ACKNOWLEDGMENTS

Useful discussions with L. Dobrzynski, B. Djafari-Rouhani, and A. Krokhin are gratefully acknowledged. We thank G. Martínez-Montes for his considerable help with the figures. Part of the research leading to this letter was performed while P. H. was at the Universidad Autónoma de Puebla. This research was supported by CONACyT Grant No. 3923-E9402.

[1] R. L. Weaver, Wave Motion **12**, 129 (1990); see also, P. Sheng, Editor, *Scattering and Localization of Classical Waves in Random Media* (World Scientific, Singapore, 1990).
[2] L. Ye, G. Cody, M. Zhou, P. Sheng, and A. N. Norris, Phys. Rev. Lett. **69,** 3080 (1992).
[3] Such calculations are routinely performed in Solid State Physics. See, e.g., C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1986), 6th ed., and C. Kittel, *Quantum Theory of Solids* (Wiley, New York, 1963).
[4] See, for instance, R. Esquivel-Sirvent and G. H. Cocoletzi, J. Acoust. Soc. Am. **95,** 86 (1994) and references therein.
[5] M. M. Sigalas and E. N. Economou, J. Sound Vib. **158**, 377 (1992).
[6] M. Sigalas and E. N. Economou, Solid State Commun. **86,** 141 (1993).
[7] M. S. Kushwaha, P. Halevi, L. Dobrzynski, and B. Djafari-Rouhani, Phys. Rev. Lett. **71,** 2022 (1993).
[8] M. S. Kushwaha and P. Halevi, Appl. Phys. Lett. **64,** 1085 (1994).
[9] M. S. Kushwaha, P. Halevi, G. Martínez-Montes, L. Dobrzynski, and B. Djafari-Rouhani, Phys. Rev. B **49,** 2313 (1994).
[10] J. O. Vasseur, B. Djafari-Rouhani, L. Dobrzynski, M. S. Kushwaha, and P. Halevi, J. Phys., Condens. Matter. **6,** 8759 (1994).
[11] J. O. Vasseur and P. A. Deymier, J. Mater. Sci. (to be published).
[12] E. N. Economou and M. Sigalas, J. Acoust. Soc. Am. **95,** 1734 (1994).
[13] We calculated the DOS using the same parameters as in Fig. 1(c) of Ref. 5, namely, $\rho_{in}/\rho_{out}=1/15$, $c_{out}/c_{in}=2.65$, and $f=0.144$. We find band gaps that differ considerably from those in this figure (the second gap is more than twice larger than the corresponding one in Ref. 5).
[14] S. John and J. Wang, Phys. Rev. Lett. **64,** 2418 (1990); Phys. Rev. B **43,** 12 772 (1991).
[15] A. A. Ruffa, J. Acoust. Soc. Am. **91,** 1 (1992) and references therein.
[16] Exactly the same differential equation as Eq. (2) is also satisfied by the pressure $p(=-\rho c_l^2 \nabla\cdot\mathbf{u})$.[5] In fact, it is easy to show that $p=\omega^2\Phi$ for harmonic waves.
[17] *CRC Handbook of Chemistry and Physics* (CRC Press, Boca Raton, FL, 1985), 66th ed.
[18] For a review on photonic band structure see E. Yablonovitch, J. Phys., Condens. Matter. **5,** 2443 (1993).
[19] M. S. Kushwaha and P. Halevi, Appl. Phys. Lett. **69,** 31 (1996).
[20] M. S. Kushwaha and B. Djafari-Rouhani, J. Appl. Phys. **80,** 3191 (1996).
[21] M. S. Kushwaha, Int. J. Mod. Phys. B **10,** 977 (1996).

An error intensity spectral filtering method for active control of broadband structural intensity

David C. Swanson, Cassandra A. Gentry,[a] Sabih I. Hayek, and Scott D. Sommerfeldt[b]
The Applied Research Laboratory, The Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804-0030

(Received 14 August 1995; revised 22 April 1996; accepted 20 August 1996)

The implementation of structural intensity measurement using an accelerometer array allows active attenuation of broadband frequency vibration using a modified filtered-x algorithm. In this work, a multi-input channel accelerometer error filter outputs a single channel error signal proportional to the propagating power. The filter is specified in the frequency domain as a transfer function from a model of bending waves in a finite beam. Integrity of the beam model and active control simulations is verified by monitoring an acceleration in the far field. The best overall performance is achieved when the error sensor array is in the far field of the primary and adaptive control force actuators. Broadband application of the intensity error technique in physical model-based simulations showed 15- to 25-dB attenuation except at a few isolated frequencies. These preliminary simulation results support new approaches to active stuctural control through error sensor array processing. © *1997 Acoustical Society of America.* [S0001-4966(97)00101-X]

PACS numbers: 43.40.Vn [PJR]

INTRODUCTION

In this work we examine the application error sensor array filtering to simultaneously extract the shear and moment field components which lead to real stuctural power, and drive and adaptive filtered-x algorithm to minimize the resulting structural power radiation. The observed acceleration field in a structure can have many components due to standing waves where the actual propagating wave components are relatively small in amplitude. Weak propagating waves in the presence of strong standing waves is perhaps one of the more salient problems associated with active reduction of structural radiation from heavy fluid-loaded structures.[1] Long wavelength (supersonic wave number) structural vibrations couple well into the fluid and thus cause those structural vibration components to be highly damped relative to the unimportant subsonic vibration. By using an array of accelerometers to extract the shear and moment field components which contribute to real vibration power, one can filter the acceleration components which do not contribute to real power and implement an adaptive filtered-x algorithm for active cancellation of the propagating power in the structure. The bending wave intensity technique has become relatively mature using accelerometer arrays[2,3] as well as using mordern cross-spectral techniques.[4] For active control of the total bending wave power, it has been shown that two control actuators are needed to simultaneously control both the shear and moment forces.[5] The major issue in integrating an intensity error signal into an active vibration control adaptive system is that the error signal used in the adaptive algorithm must be a linear function of the control action on the system. One recent approach, although limited to single narrow-band frequencies, has been experimentally shown effective where the normal product of the error and filtered-x signals in the adaptive filter coefficient updates are replaced with the instantaneous intensity itself.[6] The design approach presented here[7] is to develop an algorithm for broadband structural intensity control using the filtered-x algorithm where one has an independent reference signal coherent with the noise to be canceled. Our design develops a multichannel error accelerometer filter which combines five accelerometer signals to produce a single output signal which is linearly proportional to the bending wave components which constitute real power.

I. FREQUENCY DOMAIN MODEL OF STRUCTURAL INTENSITY

This analytical study is based on a structural steel beam also used by Hayek *et al.*[8] and Schwenk[9] so that the error intensity simultations can be compared to actual experiment for select cases. Of particular interest is the effect of structural near fields on the intensity control performance. The beam has a length L of 1.22 m, excluding an additional 0.3 m length terminated in a bed of dry, loose sand, and a cross-sectional area A of 74 μm^2. The modulus of elasticity E is assumed to be 186.9 GPa, density ρ of 7700 kg/m^3, and moment of inertia I of 2.71 nm^4. The end of the beam near $x=L$ is free and we have a harmonic point force F_0 and position x_0 normal to the beam surface. The termination impedances over the range of frequencies of interest were measured experimentally by Tousi[10] based on measuring the two flexural standing wave ratios. In order to measure the flexural intensity using standard accelerometers, one simply estimates the spatial derivitives using finite difference approximations. In this work we use a five accelerometer array with uniform 4-cm spacing to estimate all the spatial derivitives at the location of the middle, or third, accelerometer. Consider the beam setup as seen in Fig. 1.

[a]Currently at Virginia State and Polytecnical Institute, Blacksburg, VA.
[b]Currently at Brigham Young University, Salt Lake City, UT.

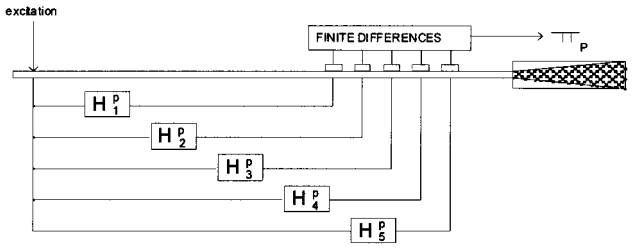


FIG. 1. The transfer function between the intensity output and the force input can be described as a weighted combination of five filters due to the finite difference derivitive approximations.

Applying well-known difference equations, one has the total bending intensity at the position of the middle accelerometer a_3 [dropping the (ω) notation for compactness], as seen in Eq. (1):

$$\Pi_p(\omega) \simeq j \frac{EI}{2\Delta^3 \omega^3} \{[-a_5 + 2a_4 - 2a_2 + a_1] \cdot a_3^* + [a_4 - 2a_3 + a_2] \cdot [a_4^* - a_2^*]\}. \quad (1)$$

Equation (1) is actually nonlinear due to the cross-spectra of accelerations. However, for steady-state or stationary vibration signals small changes in acceleration levels lead to approximately linear changes in intensity level. Since the intensity is expressed as a spectrum, it is already time averaged as well as orthogonal in frequency. Therefore, for stationary signals one can see that the approximate flexural intensity in Eq. (1) has been linearized with respect to the steady-state excitation force. The error spectral response is now simply a weighting function to drive the filtered-x adaptive control algorithm where the phase of the error plant is still included as part of the filtered-x operation on the reference signal.

II. SIMULATION RESULTS

A range of control actuator and error sensor locations was examined to minimize the intensity errors due to the finite difference approximation as well as to examine the near-field and far-field performances of intensity control in general. The spacing between the two control actuators is 10 cm and the spacing between the accelerometers in the intensity array is 4 cm. Figure 2 summarizes the control results

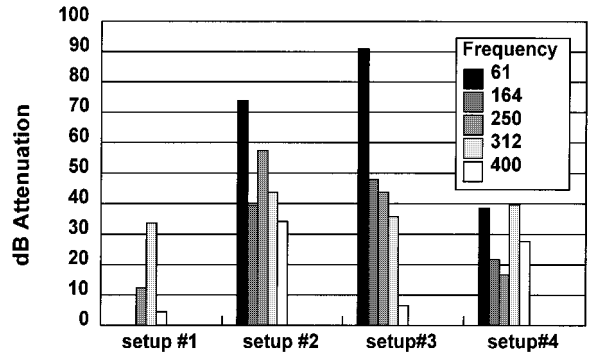


FIG. 2. Setup 3 with the control actuators in the near field of the primary excitation and error sensors in the farfield of the control actuators appears to have the best performance near resonances of the beam.

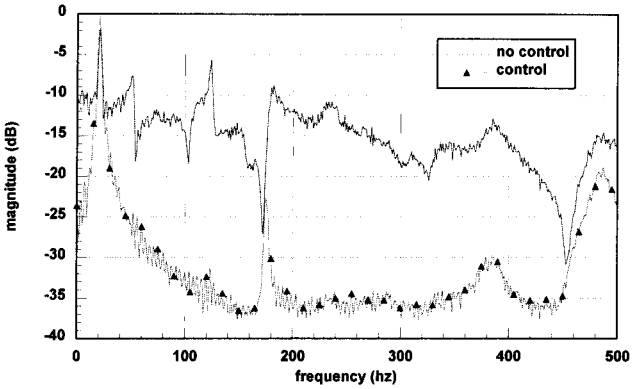


FIG. 3. Broadband control of flexural intensity from random vibrations is possible using the psuedo-intensity error filtering technique and a simple time-domain filtered-x controller.

using the psuedointensity error filtering algorithm for four actuator and error sensor array locations. Since the near-field/far-field boundary is frequency dependent, the beam termination impedance, precise locations of the control actuators and error sensor array is different for each frequency and setup and can be found in Ref. 7. Setups 1 and 2 have the active control sources in the far field of the excitation, but setup 1 has the error in the far field of the control actuators and setup 2 has the error array in the near field of the control actuators. Setups 3 and 4 have the control actuators in the near field of the primary excitation force, but setup 3 has the arror array in the far field and setup 4 in the near field of the control actuators.

Due to the limitations of the model length, one could not place the control actuators in the far field of the primary excitation for setup 1 at the lowest frequencies of 61 and 164 Hz. Both setups 2 and 3 appear to give good overall performance with setup 3 providing the best low frequency results. It is likely that finite difference approximations in the intensity calculation are contributing factors to this performance limit. The main benefit of the intensity error filtering algorithm in the filtered-x adaptive controller[11] is the ability to actively control broadband intensity using a time-domain adaptive control algorithm. The only limitation is that the excitation force should be stationary and ergodic. This allows the measured time-averaged intensity at the error array to be used to construct a filter which passes vibration signal components which are coherent with the propagating flexural intensity. Figure 3 shows the broadband performance simulation results for the intensity active control algorithm. The peaks and dips in the responses are due to the primary and secondary actuator locations on the beam and the resulting modal response as observed at the accelerometer array in the simulation.

III. CONCLUSIONS

A broadband flexural intensity adaptive control algorithm is presented which shows reasonable performance on single-frequency and broadband excitation for active intensity control. The force-to-acceleration transfer functions from point to point in the beam are modeled using the classic Euler–Bernoulli 4th order differential equation for a finite

beam. The transfer functions led us to develop a psuedo-intensity filter which essentially linearizes the intensity control problem by providing only the important error signal components which lead to propagating power. It can be seen as a broadband wave vector filter which suppresses the standing wave and near-field components while passing the propagating wave components. This filtering is however, signal dependent, meaning that one must on-line measure the flexural intensity and adaptively design a filter to pass the proper signal components. The single-frequency performance of the psuedo-intensity and linearized intensity gradient algorithms is quite comparible as are the near-field/far-field performances with the experiments of Sommerfeldt.[6,9] It can therefore be seen that the numerical modeling presented here is of reasonable enough accuracy to claim that the psuedo-intensity error filtering approach could be quite useful for the development of real-time steady-state flexural intensity control on finite beams.

[1] E. G. Williams, ''Supersonic Acoustic Intensity,'' J. Acoust. Soc. Am. **97**, 121–127 (1995).

[2] G. Pavic, ''Measurement of Structure-Borne Wave Intensity, Part I: Formulation of the methods,'' J. Sound Vib. **49**(2), 221–230 (1976).

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

[4] G. Rasmussen, ''Structural Dynamic Measurements Using Intensity Methods,'' Proc. 3rd International Modal Analysis Conference, pp. 558–564, 1985.

[5] J. Pan and C. Hansen, ''Active Control of Total Vibratory Power Flow in a Beam, I Physical System Analysis,'' J. Acoust. Soc. Am. **89**, 200–209 (1991).

[6] A. E. Schwenk, S. D. Sommerfeldt, and S. I. Hayek, ''Adaptive Control of Structural Intensity Associated with Bending Waves in a Beam,''J. Acoust. Soc. Am. **96**, 2826–2835 (1994).

[7] C. A. Gentry, ''Investigation of Adaptive Intensity Control Methods For a Finite Beam,'' M.S. thesis, The Pennsylvania State University, University Park, PA, 1993.

[8] S. I. Hayek, M. J. Perchesky, and B. Suen, ''Measurement and Analysis of Near and Far Field Structural Intensity by Scanning Laser Vibrometry,'' CETIM, Senlis, France, pp. 281–288, 1990.

[9] A. Schwenk, ''Adaptive Feedforward Control of Structural Intensity in a Beam,'' M.S. thesis, The Pennsylvania State University, University Park, PA, 1991.

[10] S. Tousi, ''Complex Impedances of Structural Terminations,'' Honors Baccalaureate thesis, The Pennsylvania State University, University Park, PA, 1991.

[11] B. Widrow and S. D. Sterns, *Adaptive Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1985).

Direct measurement of ultrasonic velocity of thin elastic layers

Mingxi Wan, Bei Jiang, and Wenwu Cao
Intercollege Materials Research Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802

(Received 2 May 1996; accepted for publication 24 August 1996)

This letter reports a simple new technique for the ultrasonic velocity measurement on thin elastic layers of fractional wavelength thickness using broadband transducers. An external trigger to the oscilloscope with continuously variable frequency and an intensity enhancing strobe signal allow direct measurement of the wavefront time shift down to 13 ns. The technique makes it possible to measure the sound velocity of thin layers with the thickness down to $10^{-2}\lambda$. Using a pair of 2.25-MHz transducers, we have accurately measured the sound velocity of a 26-μm-thick aluminum foil. © *1997 Acoustical Society of America.* [S0001-4966(97)04112-X]

PACS numbers: 43.58.Dj, 43.20.Jr, 43.35.Zc, 43.35.Cg [SLE]

INTRODUCTION

There are many situations of technological importance in which one wishes to carry out the elastic property characterization of thin layer materials; for example, bonding layers in structural materials, layers in composite structures, and thin film materials on thick substrate. Over the past 30 years, a wide variety of ultrasonic techniques have been reported for the measurements of wave velocities or the thickness of a thin single elastic layer. Among all classical methods are pulse echo,[1] resonance testing,[2] pulse interference methods,[3] and laser ultrasonic method.[4] These traditional techniques measure the time delay in well-separated echoes. However, as the layer thickness decreases, the time interval between two successive echoes from the front and back surfaces of the specimen decreases. Eventually, all of the classical methods fail when the echoes become inseparable.

In many applications, the thickness of the elastic layer, h, is of the order of 10–100 μm, and may even be imbedded inside a structure or blocked by another object. If using the well-separated pulse method to characterize such thin layers, the frequency would have to be larger than 150 MHz. The use of such a high frequency not only drastically increases the cost, but also limits the application to low loss materials because of the very short penetration depth of high-frequency waves.

Several years ago, a frequency domain method[5] for ultrasonic nondestructive evaluation of thin specimens was developed and improved by means of combining the standard FFT methods with conventional ultrasonic method. This method is based on the fact that a thin specimen in the time domain takes the form of a thick specimen in the frequency domain. A time domain method for the measurement of ultrathin specimen was also developed in 1993 in which low-frequency ultrasonic wave was also used.[6] However, for these frequency and time domain methods, digital oscilloscopes with a very high sampling rate and specialized data analyses are required. Sub-half-wavelength thickness thin layers can be measured using these methods down to $10^{-1}\lambda$.

In this letter we describe a new approach for the measurement of a thin elastic layer with the thickness in the order of $10^{-2}\lambda$. The idea is to manipulate the triggering rate and the strobe sweeping to fully utilize the maximum sensitivity of the oscilloscope. We call the technique variable trigger and strobe (VTS) method. Using this technique, the time shift caused by placing and removing the sample from the acoustic pathway between transmitting and receiving transducers can be accurately measured down to the maximum capacity of the oscilloscope (0.01 ns for the one available to us). Since only the starting point of the transmitted wave is used, a low-frequency transducer can be used to measure these samples. This gives us the advantage of deep penetration and the ability to measure thin layers that are sandwiched in lossy materials.

I. THE VTS METHOD AND EXPERIMENTAL PROCEDURE

A schematic of the experimental apparatus is shown in Fig. 1. A pair of accurately matched broadband water immersion longitudinal wave transducers with a center frequency of 2.25 MHz were used for generating and receiving the ultrasonic waves. The distance between the two transducers is adjusted to be equal to twice the focal length, and the specimen is placed in the focal region of the transducers. The incident wave travels in the direction normal to the specimen. For a single-layer specimen, the displayed pulse signal from the time t_0, at which the received signal appears, to the time t_0+2h/c will not have the effect of multireflection interference. Therefore, we can get the longitudinal velocity c by measuring the time shift Δt of the starting point of the transmitted waves caused by including the specimen in the acoustic pathway,

$$c=\frac{c_w h}{h-c_w\Delta t}, \tag{1}$$

where h is the sample thickness and c_w is the sound velocity of water (reference medium). The time shift Δt can be either positive or negative depending on if the c is faster or slower than c_w.

Because the technique is to compare the travel time with and without the sample in the acoustic pathway, it is insensitive to other objects in the pathway since their contributions can be included in the travel time without the sample.

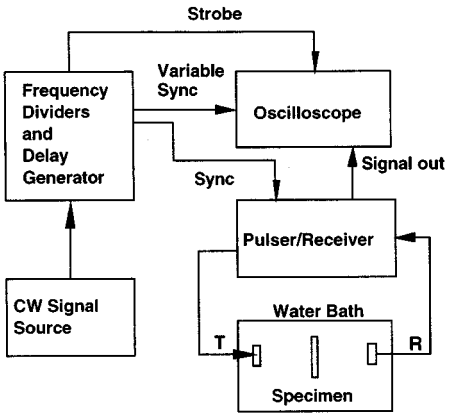


FIG. 1. Schematic diagram of the experimental setup.

The only difference between having water as the base medium or including other objects in the base medium is the intensity level. Since the technique allows the use of relatively low frequencies, it is possible to penetrate even lossy materials in the pathway. This characteristic makes the technique very useful for characterizing a thin layer in a sandwich structure, such as the bonding layer.

Although most oscilloscopes have subnanosecond resolution in the horizontal scale, it is not possible to use the highest resolution in most cases because of the limited display of time interval on the display window. Most of the analog oscilloscopes, such as the one used in this experiment, do not have time delay function; therefore, it is impossible to follow the wavefront when changing the horizontal resolution. One way to utilize higher resolution of the oscilloscope is to use higher external triggering frequency; however, it will mix up the received pulse signals, making it impossible to distinguish different echoes. Considering the analog oscilloscope has an intensity regulating input, the selected echo is highlighted by using a strobe signal from a Matec 122B decade dividers and delay generator. By reducing the background intensity, this selected signal can be singled out for measurement using the high resolution scale of the oscilloscope with high external triggering frequency.

General procedure for the VTS method: As shown in Fig. 1, the Matec 110-CW signal source is integrated with a highly stable oscillator tunable over a range of approximately from 11 to 52 MHz. The frequency output of the oscillator is divided by 100, and then is further divided by 100 in the Matec 122B decade dividers and delay generator. A more stable pulse signal with a frequency range of 1.1–5.2 kHz is obtained. This frequency is then used to trigger the JSR DPR 35 pulser/receiver and the Tektronix 2465 oscilloscope. The received pulse signals are sent from the DPR 35 to the oscilloscope for display. The strobe from the Matec 122B is coupled to the Z axis (intensity) of the oscilloscope and the front portion of the first transmitted wave signal is brightened by adjusting the strobe delay, width, and amplitude. The intensity of the oscilloscope must be carefully adjusted together with the strobe amplitude to obtain a proper display of the front portion of the transmitted wave signal so that the unwanted portion of the wave signal and the background sweep line will disappear. The front portion of the transmitted wave is then expanded horizontally to the proper resolution of the oscilloscope according to the level of time shift to be expected. The signal will be out of the screen from the right-hand side as the resolution of the oscilloscope increases. Now the external trigger is switched from the Matec 110 signal source to another scale which has 100 times higher triggering frequency. The brightened front portion of the first transmitted wave (or wave train) will reappear on the screen. By adjusting the frequency of the signal source to change the triggering frequency of the whole system, we can place the starting point of the front portion of the first transmitted wave anywhere on the screen. As a rule of thumb, place the signal on the left-hand side of the screen if the wave speed in the specimen is faster than that in water, and to the right-hand side if the compressional wave velocity in the sample is slower than that in water. After the thin layer sample is taken out of the ultrasonic field, the starting point of the front portion of the first transmitted wave shifts to a new position. This shift can be directly measured by using the scale bars on the screen of the scope. Then the compressional velocity in the thin layer of known thickness can be calculated using Eq. (1). Vice versa, if the velocity is known, the time shift can be used to measure the sample thickness.

II. RESULTS AND DISCUSSIONS

The main objective of this work is to develop a simple technique suitable for the characterization of ultrathin elastic layers of only a fraction of a wavelength in thickness. We

TABLE I. Wave speed measurement data for single elastic layers.

Specimen No.	Material	Thickness (mm) ± 3 μm	Time shift (ns)	Velocity (mm/μs)	h/λ	Sweep speed (ns/Div.)
1	Glass	1.247	613.4	5.579	0.503	100
2	PZT	0.236	91.70	3.539	0.150	20
3	Transparency #1	0.104	24.55	2.298	0.102	20
4	Transparency #2	0.096	18.45	2.088	0.103	20
5	Plastic shopping bag	0.058	16.55	2.592	0.050	10
6	Aluminum foil	0.026	13.26	6.206	0.009	5

TABLE II. Wave speed measurement data for elastic thin layers on substrates.

Specimen No.	Coating material	Thickness (mm) ± 3 μm	Time shift (ns)	Velocity (mm/μs)	h/λ	Sweep speed (ns/Div.)
3	Transparency #1	0.104	25.00	2.321	0.101	20
4	Transparency #2	0.096	19.70	2.146	0.101	20

have measured a number of specimens using the newly developed VTS method. The results for a single layer and for a thin layer on a thick substrate are presented in Tables I and II, respectively. All results were obtained under constant temperature of 20 °C; thus the value of c_w is taken as 1.49 mm/μs. The substrate used for testing the coating specimen are glass plates of 1.247 mm in thickness.

As shown in Tables I and II, the measurements were performed on several materials with thickness ranging from 0.026 mm to 1.24 mm. The central frequency of the broadband transducer used in the experiments is 2.25 MHz, hence, the thickness range corresponds to 0.009 λ–0.503 λ. From a systematic error analysis (to be reported later), we found that the total error in the velocity measurement is less than 3%. Since most of the materials measured do not have known velocity values, we can only compare the longitudinal velocity of the aluminum foil which has been measured in bulk form.[7] Our measured value of 6.206 mm/μs is 3% smaller than the bulk value of 6.400 mm/μs. Since this difference is within our experimental accuracy, we cannot attribute it to the difference between bulk and foil materials. Generally speaking, the VTS method works better for high sound velocity materials which will lead to larger time shift, as shown in Table I. In other words, it is more sensitive for metal, ceramic, and glass than for plastic and polymer products.

As the thickness of the thin layer or the velocity difference between the sample and water decreases, we must increase the horizontal magnification on the scope. For the ultrathin aluminum foil (specimen #6 in Table I) with $h/\lambda = 0.009$, the vertical scale also must be enlarged in order for the starting point to be located more accurately on the screen.

The signal source and decade dividers used are very stable; no drifting of the signal was observed. However, water disturbance caused by removing the specimen out of the ultrasonic field does cause the signal to oscillate. Therefore, it is necessary to remove the sample slowly and to wait a few minutes before taking the comparison data.

For the ultrathin layers with $h/\lambda < 10^{-2}$, the noise in the receiver output could affect the accuracy. Especially when the vertical amplitude has to be enlarged enough on the screen together with time axis expansion, the ambiguity of the starting point resulting from noise is the main error source of wave velocity measurement for ultrathin layers of $h/\lambda < 10^{-2}$ using this VTS technique. It is expected that the VTS technique could be further improved by using digitizing oscilloscope combined with numerical signal average techniques. This technique could be used in reflection mode if the coupling between the transducer and the sample is accurately controlled.

ACKNOWLEDGMENT

This work is supported by the Office of Naval Research under Contract No. N00014-93-1-0340.

[1] E. P. Papadakis, ''Ultrasonic velocity and attenuation: Measurement methods with scientific and industrial applications,'' in *Physical Acoustics Principles and Methods*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1976), Vol. XII, pp. 277–374.

[2] D. I. Bolef and M. Menes, ''Measurement of elastic constants of RbBr, RbI, CsBr, and CsI by an ultrasonic CW resonance technique,'' J. Appl. Phys. **31**, 1010–1017 (1960).

[3] H. L. Mcskimmin, ''Pulse superposition method for measuring the velocity of sound in solids,'' J. Acoust. Soc. Am. **33**, 12–16 (1961).

[4] A. C. Tam and H. Coufal, ''Photoacoustic generation and detection of 10-ns acoustic pulse in solids,'' Appl. Phys. Lett. **42**, 33–35 (1983).

[5] V. K. Kinra and V. Dayal, ''A new technique for ultrasonic-nondestructive evaluation of thin specimen,'' J. Exp. Mech. **28**(3), 288–297 (1988).

[6] V. K. Kinra and C. Zhu, ''Time-domain ultrasonic NDE of the wave velocity of a sub-half-wavelength elastic layer,'' J. Test. Eval. **21**(1), 29–35 (1993).

[7] W. P. Mason, ''Acoustic Properties of Solids,'' in *American Institute of Physics Handbook* (McGraw-Hill, New York, 1972), 3rd ed.